



HAL
open science

Asymptotiques en grande population et temps long de perturbations aléatoires de modèles de champ moyen

Eric Luçon

► **To cite this version:**

Eric Luçon. Asymptotiques en grande population et temps long de perturbations aléatoires de modèles de champ moyen. Mathematics [math]. Université Paris Cité, 2024. tel-04451781

HAL Id: tel-04451781

<https://hal.science/tel-04451781>

Submitted on 11 Feb 2024

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

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

Université Paris Cité

MÉMOIRE
en vue de l'obtention de
**L'HABILITATION À DIRIGER DES
RECHERCHES**

Spécialité : Mathématiques Appliquées

Eric LUÇON

**Asymptotiques en grande population et temps long
de perturbations aléatoires de modèles de
champ-moyen**

Rapporteurs : François DELARUE Université Côte d'Azur
 Arnaud GUILLIN Université Clermont Auvergne
 Michael SCHEUTZOW Technische Universität Berlin

Soutenue publiquement le 31 janvier 2024 devant le jury composé de :

François DELARUE	Professeur, Université Côte d'Azur
Giambattista GIACOMIN	Professeur, Université Paris Cité
Arnaud GUILLIN	Professeur, Université Clermont Auvergne
Eva LÖCHERBACH	Professeure, Université Paris 1
Sylvie MÉLÉARD	Professeure, Ecole Polytechnique
Ellen SAADA	Directrice de recherche, Université Paris Cité
Michael SCHEUTZOW	Professeur, Technische Universität Berlin
Lorenzo ZAMBOTTI	Professeur, Sorbonne Université

- *Toujours à la recherche de nouvelles aventures ? Qu'allez-vous choisir cette fois-ci ?*
- *Jusqu'ici je n'ai guère eu le choix.*
- *C'est vrai. Eh bien moi, je vous propose un choix. Vous connaissez le jeu de Pile ou Face ? Je vous laisse choisir : Pile vous allez dans un sens, et Face vous allez dans l'autre.*
- *Cela ne revient-il pas au même ?*
- *Pas toujours... Alors ? Que choisissez-vous ?*
- *Heu.. Pile. C'est Face. Evidemment.*
- *Ah, ah, ah, ne vous inquiétez pas, Monsieur Acquefacques. A ce jeu là, on ne perd jamais. Mais on ne gagne pas non plus.*

Marc-Antoine Mathieu - *L'épaisseur du miroir*

Contents

Remerciements	7
Avant-propos	9
List of publications	11
1 Introduction	13
1.1 Mean-field particle systems in large population	13
1.2 Towards non-exchangeability	15
1.3 Long-time periodicity for mean-field systems	21
1.4 Inhibition and oscillations for mean-field Hawkes processes	23
2 Inhomogeneous particles in large population	25
2.1 Quenched limit for the empirical flow	25
2.2 Quenched fluctuations	28
2.3 Quenched LDP for the empirical measure	32
2.4 The case of singular spatial interactions	36
2.5 Comments and perspectives	41
3 Spatially-structured models and random graphs	43
3.1 Homogeneous graphs: Law of Large Numbers	44
3.2 Homogeneous graphs: Central Limit Theorem	50
3.3 Inhomogeneous graphs	56
3.4 Perspectives	62
4 Periodicity for nonlinear Fokker-Planck equations	67
4.1 About Stable Normally Hyperbolic Manifolds (SNHM)	67
4.2 Traveling waves in the asymmetric Kuramoto model	70
4.3 Excitable systems and the FitzHugh-Nagumo model	72
4.4 FHN model: the kinetic case	78
4.5 On regularity of limit cycles and isochrons	81
4.6 Perspectives	84
5 Long-time analysis of inhomogeneous diffusions	89
5.1 Random traveling waves in the disordered Kuramoto model	89
5.2 Long-time diffusive behavior for general mean-field diffusions	91
5.3 Perspectives	93
6 Hawkes processes with multiplicative inhibition	95
6.1 Hawkes processes with excitation and inhibition	95
6.2 Our model: multiplicative inhibition	96
6.3 Perspectives	100
Bibliography	105

Remerciements

Exercer le métier d'enseignant-chercheur est une aventure décidément collective et tout ce travail ne serait rien sans de nombreuses personnes que ces quelques lignes ne remercieront qu'imparfaitement.

J'adresse tout d'abord mes remerciements les plus chaleureux à François Delarue, Arnaud Guillin et Michael Scheutzow pour leur lecture attentive et leurs commentaires bienveillants et constructifs sur ce manuscrit. Les articles fondateurs de Sylvie Méléard sur les fluctuations de modèles de champ-moyen constituent une sorte de fil rouge de mes travaux et ceux d'Eva Löcherbach sur les processus de Hawkes ont aussi été particulièrement importants pour moi. Je la remercie d'ailleurs vivement d'avoir été à l'initiative de l'encadrement de thèse de Zoé Agathe-Nerine. Des remerciements particuliers vont à Ellen Saada pour avoir accepté de co-encadrer Zoé, ce fut un mentorat efficace et productif! Je mesure bien le chemin parcouru depuis ma soutenance de thèse et la présence dans ce jury de Giambattista Giacomini et Lorenzo Zambotti me fait très plaisir. Je suis très honoré de la participation de toutes et tous à mon jury.

Il se dit que l'atout principal du laboratoire MAP5 est sa terrasse du 7^{ème} étage. Ce n'est que partiellement vrai. C'est avant tout l'atmosphère chaleureuse et conviviale qui y règne qui fait de cet endroit un lieu d'échanges et de travail exceptionnel (et c'est la raison pour laquelle j'ai la satisfaction quotidienne d'affronter les périls du RER B et des ascenseurs défaillants des Saints-Pères pour y venir). Je remercie ici vivement tous les membres du laboratoire et de l'UFR, permanents, doctorants, personnels administratifs (je n'oublie pas la scolarité) que j'ai eu et que j'ai le plaisir de côtoyer depuis mon arrivée. Cette atmosphère bienveillante doit certainement beaucoup aux directrices et directeur successif.ive.s du laboratoire, Annie Raoult, Fabienne Comte, Anne Estrade et Antoine Chambaz, qu'elles et il en soient ici grandement remercié.e.s. Des vifs remerciements vont évidemment à Marie-Hélène Gbaguidi sans qui pas grand chose ne fonctionnerait dans le laboratoire. Merci à Nathalie Eisenbaum pour avoir diligemment orienté mon manuscrit dans les méandres administratifs de la faculté. Un grand merci à mes co-bureaux successifs Marc, Jonathan et plus particulièrement à Céline pour avoir partagé dès le début l'expérience du métier d'enseignant-chercheur (ainsi que thé et chocolat).

Je remercie vivement tous les collègues avec qui j'ai eu la chance de discuter et de travailler, en particulier Zoé Agathe-Nerine, Fabio Coppini, Sylvain Delattre, Céline Duval, Christophe Poquet, Christophe Pouzat et Wilhelm Stannat dont les discussions mathématiques m'ont beaucoup apporté. Je tiens en particulier à souligner le compagnonnage au long cours avec Christophe Poquet, à qui les travaux de ce manuscrit doivent beaucoup. A tou.te.s : si une idée intelligente y subsiste, c'est très certainement la vôtre! Je remercie aussi vivement les membres des projets ANR *Peristoch*, *Chamane* et *Happy* pour les rencontres et discussions stimulantes.

La tentation a été grande d'ajouter à ce manuscrit un dernier chapitre intitulé *Doliprane et Amoxicilline*. La rédaction d'un manuscrit d'habilitation en période post-Covid avec des enfants en bas-âge est un processus aléatoire, discontinu en tout point, non-prévisible dont la convergence à l'équilibre fut à l'évidence sous-optimale. Si, dans ce contexte sanitaire compliqué, on a souvent évoqué les dites *professions de première ligne* (et celle d'enseignant-chercheur n'en a certainement pas fait partie, cachés que nous étions derrière nos sessions zoom), celle d'assistante maternelle l'a été indéniablement : qu'il me soit permis ici de remercier chaleureusement Odile J. et Martine A. pour leur disponibilité et leur aide précieuse. Un grand merci à mes beaux-parents pour leur soutien permanent et indispensable, sans lequel ce manuscrit serait encore dans les limbes.

Ce manuscrit est dédié à Claire, Maël et Camille.

Avant-propos

Ce mémoire d'habilitation est une synthèse des travaux que j'ai effectués depuis ma thèse de doctorat, soutenue en juin 2012 à l'Université Pierre et Marie Curie. Ma recherche s'inscrit dans le domaine des systèmes de particules en champ-moyen et leurs perturbations (dynamique en environnement et sur des graphes aléatoires) et les applications de ces modèles à des systèmes physiques (e.g. synchronisation de modèles d'oscillateurs) ou biologiques (populations de neurones en interaction). Deux asymptotiques de ces systèmes sont étudiées : en grande population et en temps long.

Ces travaux ont donné lieu aux publications référencées ci-dessous, dont on donne ici un bref descriptif :

- Les articles suivants s'intéressent au comportement en grande population sur un intervalle de temps borné de systèmes en champ-moyen soumis à diverses perturbations :
 1. En présence d'un milieu aléatoire, que ce soit du point de vue de la convergence (loi des grands nombres) de la mesure empirique vers la solution d'une équation de Fokker-Planck non linéaire et des fluctuations associées [L1, L4] mais aussi des Grandes Déviations (quenched) [L8] autour de cette limite.
 2. Pour des modèles avec interactions spatiales, possiblement singulières, là encore du point de vue de la loi des grands nombres [L3] et du Théorème Central Limite [L6] de la mesure empirique.
 3. Pour des modèles interagissant sur des graphes aléatoires possiblement dilués (lois des grands nombres [L5, L10] et fluctuations [L14]).
- Un second aspect des travaux exposés dans ce manuscrit concerne le comportement en temps long de processus de McKean-Vlasov non linéaires associés à ces modèles de champ-moyen. L'accent est mis sur l'émergence, la stabilité et la régularité de trajectoires périodiques pour ces processus :
 1. L'article [L2] traite de l'existence de solutions périodiques stables pour le modèle de Kuramoto soumis à un désordre asymétrique,
 2. Les articles [L9, L11] s'intéressent à l'influence du bruit et de l'interaction sur l'émergence de solutions périodiques pour des systèmes excitables et en particulier pour le modèle de FitzHugh-Nagumo,
 3. La question générale de la stabilité et de la régularité de ces solutions périodiques est abordée dans [L12].
- Un troisième aspect des travaux exposés ici concerne la stabilité des comportements périodiques évoqués plus haut vis-à-vis de la mesure empirique du système microscopique de taille finie. Le comportement en temps long de la mesure empirique du modèle de Kuramoto en présence de désordre symétrique est abordé dans [L7] et le problème similaire dans un cadre général de systèmes excitables est traité dans [L15].
- Un dernier aspect traité ici concerne un intérêt récent pour des dynamiques à sauts de type Hawkes en présence d'inhibition. L'apport de l'article [L13] est de proposer une modélisation nouvelle de l'inhibition dans de tels modèles et d'en étudier l'émergence de solutions macroscopiques périodiques.

Par ailleurs, je co-encadre, avec Ellen Saada, depuis septembre 2020 la thèse de Zoé Agathe-Nerine portant sur la dynamique en grande population et en temps longs de processus de Hawkes

en interaction sur des graphes aléatoires, thématique présentant des liens forts avec les outils développés dans ce manuscrit.

Pour faciliter la lecture du manuscrit, les résultats évoqués sont souvent non exhaustifs, parfois mentionnés sous des hypothèses simplificatrices et seules les lignes directrices des preuves seront données. Pour un exposé plus complet des résultats, on convie le lecteur à se référer aux articles eux-mêmes.

List of publications

Published papers

- [L1] E. Luçon. Quenched limits and fluctuations of the empirical measure for plane rotators in random media. *Electronic Journal of Probability*, 16:792–829, 2011.
- [L2] Giambattista Giacomin, Eric Luçon, and Christophe Poquet. Coherence Stability and Effect of Random Natural Frequencies in Populations of Coupled Oscillators. *J. Dynam. Differential Equations*, 26(2):333–367, 2014.
- [L3] Eric Luçon and Wilhelm Stannat. Mean field limit for disordered diffusions with singular interactions. *Ann. Appl. Probab.*, 24(5):1946–1993, 2014.
- [L4] Eric Luçon. Large time asymptotics for the fluctuation SPDE in the Kuramoto synchronization model. *Journal of Functional Analysis*, 266(11):6372 – 6417, 2014.
- [L5] Sylvain Delattre, Giambattista Giacomin, and Eric Luçon. A Note on Dynamical Models on Random Graphs and Fokker–Planck Equations. *J. Stat. Phys.*, 165(4):785–798, 2016.
- [L6] Eric Luçon and Wilhelm Stannat. Transition from Gaussian to non-Gaussian fluctuations for mean-field diffusions in spatial interaction. *Ann. Appl. Probab.*, 26(6):3840–3909, 2016.
- [L7] E. Luçon and C. Poquet. Long time dynamics and disorder-induced traveling waves in the stochastic Kuramoto model. *Ann. Inst. Henri Poincaré Probab. Stat.*, 53(3):1196–1240, 2017.
- [L8] Eric Luçon. Quenched large deviations for interacting diffusions in random media. *Journal of Statistical Physics*, 166(6):1405–1440, Jan 2017.
- [L9] Eric Luçon and Christophe Poquet. Emergence of Oscillatory Behaviors for Excitable Systems with Noise and Mean-Field Interaction: A Slow-Fast Dynamics Approach. *Communications in Mathematical Physics*, 373(3):907–969, December 2019.
- [L10] Eric Luçon. Quenched asymptotics for interacting diffusions on inhomogeneous random graphs. *Stochastic Processes and their Applications*, 130(11):6783–6842, November 2020.
- [L11] Eric Luçon and Christophe Poquet. Periodicity induced by noise and interaction in the kinetic mean-field FitzHugh–Nagumo model. *The Annals of Applied Probability*, 31(2), April 2021.
- [L12] Eric Luçon and Christophe Poquet. Existence, stability and regularity of periodic solutions for nonlinear Fokker–Planck equations. *Journal of Dynamics and Differential Equations*, March 2022.
- [L13] Céline Duval, Eric Luçon, and Christophe Pouzat. Interacting Hawkes processes with multiplicative inhibition. *Stochastic Processes and their Applications*, 148:180–226, June 2022.
- [L14] F. Coppini, E. Luçon, and C. Poquet. Central limit theorems for global and local empirical measures of diffusions on Erdős–Rényi graphs. *Electronic Journal of Probability*, 28, Jan. 2023.

Preprints

- [L15] Eric Luçon and Christophe Poquet. Periodicity and longtime diffusion for mean field systems in \mathbb{R}^d . *ArXiv e-print 2107.02473*, July 2021.

Chapter 1

Introduction: perturbations of mean-field interacting particle systems

The purpose of this manuscript is to present in a coherent way the research I have been conducting since my PhD thesis on the topic of mean field interacting particle systems subject to (possibly random) inhomogeneities. The main attention will be put on the influence of these perturbations (seen as a random environment) on the behavior of the system for two asymptotics: in large population and in long time. This introductory chapter is the opportunity to present the main objects we will consider in the manuscript as well as the main issues on these models.

1.1 Mean-field particle systems in large population

1.1.1 The mean-field framework

A fairly generic instance of mean-field interacting diffusions may be given as follows: let $n \geq 1$ be the size of the population, $T > 0$ the time horizon and $d \geq 1$ the dimension of the state space and consider the solution $(\theta_t^{1,n}, \dots, \theta_t^{n,n})$ in $(\mathbb{R}^d)^n$ to the following system of coupled Stochastic Differential Equations (SDEs)

$$d\theta_t^{i,n} = F(\theta_t^{i,n})dt + \Gamma(\theta_t^{i,n}, \mu_{n,t})dt + \sigma(\theta_t^{i,n}, \mu_{n,t}) \cdot dB_t^i, \quad i = 1, \dots, n, \quad t \in [0, T], \quad (1.1)$$

where μ_n is the empirical measure of the particles $(\theta^{1,n}, \dots, \theta^{n,n})$

$$\mu_{n,t} := \frac{1}{n} \sum_{i=1}^n \delta_{\theta_t^{i,n}}, \quad t \in [0, T], \quad (1.2)$$

element of the space $\mathcal{C}([0, T], \mathcal{P}(\mathbb{R}^d))$ of continuous functionals with values in the set of probability measures on \mathbb{R}^d . The structure of (1.1) accounts for various dynamical features: a local drift $F(\cdot)$ representing the intrinsic dynamics of each particle, a global coupling term $\Gamma(\theta_t^{i,n}, \mu_{n,t})$ modelling all-to-all interaction between particles through the empirical measure of the system (1.2) and (possibly multiplicative) noise, in the presence of i.i.d. standard Brownian motions (B_t^i) on \mathbb{R}^d . Note that the mean-field formalism does not reduce to diffusions, as it is for example possible to include various dynamics with jumps, see e.g. [APF18, DIRT15, Cor21]. An instance of mean-field interacting point processes modelling neuronal activity (Hawkes processes) will be analysed in Chapter 6. The applications of system (1.1) are numerous and particularly well-covered in the literature (see [CD22] for a review): to list only a few, statistical physics [DdH96, CR16, CDF15, CFT16, GP15], biological models of synchrony [DPGR14, CG19, AM19a], neuroscience [BFFT12, nT20, GPPP12, GP15, MQT16, THF12, DL17, CTV21], life sciences [JTT18, DFL12], social sciences [CDS10], fluid dynamics [AM19b], finance [JR13], this list being by no means exhaustive. Note that the possibility that the interaction kernel Γ may have singularities often raises serious questions concerning the well-posedness of (1.1) and its

behavior as $T \rightarrow \infty$, see e.g. [JTT18, TT20, BCD20]. This is not a point that will concern us mostly in this manuscript: the appropriate smoothness hypotheses on Γ and σ will be adopted when required. In fact, there is already for us sufficient difficulty in analysing the following simpler instance of (1.1): with an obvious abuse of notation, choose a linear interaction in μ , $\Gamma(\theta, \mu) = \int \Gamma(\theta, \theta') \mu(d\theta')$ and some constant diffusion $\sigma(\theta, \mu) = \sigma$

$$d\theta_t^{i,n} = F(\theta_t^{i,n})dt + \frac{1}{n} \sum_{j=1}^n \Gamma(\theta_t^{i,n}, \theta_t^{j,n}) dt + \sigma dB_t^i, \quad i = 1, \dots, n, \quad t \in [0, T]. \quad (1.3)$$

Extending the results of this manuscript from (1.3) to the general (1.1) may present technical difficulties that we do not address here (although most of the present results extend readily to the case of multiplicative noise provided σ is nondegenerate).

1.1.2 Propagation of chaos and McKean-Vlasov diffusions

The first standard question concerns the behavior of the empirical measure (1.2) in large population, that is as $n \rightarrow \infty$ on some bounded time interval $[0, T]$. Applying Ito formula to (1.1) (see (1.8) below), and passing formally to the limit as $n \rightarrow \infty$, we see that the limit μ_t of $\mu_{n,t}$ is the weak solution to the following nonlinear Fokker-Planck equation, with $a(\theta, \mu) := \sigma \sigma^\dagger(\theta, \mu)$, (where σ^\dagger denotes the transpose of σ)

$$\partial_t \mu_t = \frac{1}{2} \sum_{k,l=1}^d \partial_{k,l}^2 (a_{k,l}(\theta, \mu_t) \mu_t) - \nabla \cdot \{(F(\theta) + \Gamma(\theta, \mu_t)) \mu_t\}. \quad (1.4)$$

The question of convergence of the empirical measure (1.2) to its mean-field limit (1.4) has raised a considerable interest since the seminal works of Kac [Kac56] and McKean [McK67]. It would be pointless to review here the vast literature on such matter, but let us nonetheless briefly comment on the main issues at stake. A key point in the analysis of (1.1) relies on the notion of *exchangeability*: the law f_t^n of $(\theta_t^{1,n}, \dots, \theta_t^{n,n})$ at time t is invariant by permutation, provided that the same holds at time $t = 0$ (this is for example true when $(\theta_0^{1,n}, \dots, \theta_0^{n,n})$ are i.i.d.). In such a case, the k -th marginal $f_t^{k,n}$ of f_t^n for any $k \geq 1$ is defined without ambiguity as the law of any finite collection of k variables among $(\theta_t^{1,n}, \dots, \theta_t^{n,n})$. Intuitively, as the interaction between k particles in (1.1) is weak (of order $\frac{1}{n} \xrightarrow{n \rightarrow \infty} 0$), these particles decorrelate as $n \rightarrow \infty$. This is formalised in terms of the k -chaoticity of f^n : for all fixed $k \geq 1$, $f^{k,n}$ converges weakly to $\mu^{\otimes k}$ where μ solves (1.4). This convergence can be considered for various topologies (see [HM14, CD22] for extensive reviews on the subject). A classical result (see e.g. [Szn91, Prop. 2.2]) is that k -chaoticity of f^n for some k (itself equivalent to k -chaoticity for $k = 2$) is equivalent to the convergence of the empirical measure μ_n to μ . The celebrated notion of *propagation of chaos* states that chaoticity for (1.1) at time $t = 0$ propagates at later times $t > 0$. Once again, we do not aim to be exhaustive on this point, but let us at least mention the main strategies that have been followed to prove such propagation of chaos in mean-field systems.

Coupling arguments

A first approach is to proceed via *coupling arguments*: the proper limit candidate as $n \rightarrow \infty$ for one particle, say $\theta_t^{1,n}$, on $[0, T]$ is given by the nonlinear McKean-Vlasov process (where $\mathcal{L}(\theta)$ is the law of θ)

$$d\bar{\theta}_t = F(\bar{\theta}_t)dt + \Gamma(\bar{\theta}_t, \mu_t) dt + \sigma(\bar{\theta}_t, \mu_t) dB_t, \quad t \in [0, T], \quad \mu_t = \mathcal{L}(\bar{\theta}_t). \quad (1.5)$$

The nonlinear character of the diffusion (1.5) lies in the fact that $\bar{\theta}$ interacts with its own law μ solution to (1.4). The well-posedness of (1.5) is an issue in itself, especially in case of singular coefficients (a standard way is to proceed with some fixed-point arguments [Szn91] in case of regular (Lipschitz) coefficients). A standard coupling between the particle system (1.1) and its mean-field limit (1.5) is the simplest *synchronous* one, defined as follows: consider the collection of i.i.d. $(\bar{\theta}^i)_{i=1, \dots, n}$ solutions to (1.5), driven by the *same* Brownian motions (B_i) as for (1.1)

with the *same* initial condition $\bar{\theta}_0^i = \theta_0^{i,n}$. Then, some easy Grönwall arguments [Szn91] show that under reasonable (Lipschitz) conditions on F , Γ and σ , the following estimate holds

$$\mathbf{E} \left(\sup_{t \in [0, T]} |\theta_t^{i,n} - \bar{\theta}_t^i| \right) \leq \frac{C(T)}{\sqrt{n}}, \quad (1.6)$$

hence implying the k -chaoticity of $f^{k,n}$ for all $k \geq 1$ as $n \rightarrow \infty$. Note that an alternative approach using *reflective coupling* (that we will not use here) has been successfully used to estimate rates of convergence to equilibrium for nonlinear McKean-Vlasov SDEs [Ebe16, EGZ16] as well as uniform propagation of chaos for mean-field systems (1.1) (see [DEGZ20] and references therein). Estimate (1.6) directly implies the weak convergence of the empirical measure μ_n of the particles (1.2) towards the solution μ to the nonlinear Fokker-Planck equation (1.4). This can be expressed simply in terms of the bounded-Lipchitz distance on the set $\mathcal{P}(\mathbb{R}^d)$ of probability measures on \mathbb{R}^d , i.e. $d_{BL}(\mu, \nu) = \sup_{f, \|f\|_\infty \leq 1, \|f\|_{Lip} \leq 1} |\int f d\mu - \int f d\nu|$ as

$$\mathbf{E} \left(\sup_{t \in [0, T]} d_{BL}(\mu_{n,t}, \mu_t) \right) \leq \frac{C(T)}{\sqrt{n}}, \quad (1.7)$$

but numerous alternative results that encode the convergence in law are possible e.g. in terms of Wasserstein metrics. We refer e.g. to [CD22] for more references on the subject.

Compactness arguments

A second approach is to tackle the convergence of μ_n to μ directly from the semimartingale decomposition of μ_n , easily derived from Ito's formula: for any regular test function f (use the usual duality notation $\langle \mu, f \rangle = \int f d\mu$),

$$\begin{aligned} \langle \mu_{n,t}, f \rangle &= \langle \mu_{n,0}, f \rangle + \int_0^t \left\langle \mu_{n,s}, \frac{1}{2} \operatorname{div}_\theta (\sigma \sigma^\dagger \nabla_\theta f) + \nabla_\theta f \cdot F \right\rangle ds \\ &\quad + \int_0^t \left\langle \mu_{n,s}, \nabla_\theta f \cdot \int \Gamma(\cdot, \theta') \mu_{n,s}(d\theta', d\omega') \right\rangle ds + M_{n,t}(f), \end{aligned} \quad (1.8)$$

where $M_{n,t}(f)$ is a martingale. A common strategy in this context is (i) to prove tightness of (μ_n) (exploiting classical criteria in $\mathcal{P}(\mathcal{C}([0, T], \mathbb{R}^d))$ see e.g. [Bil99]), (ii) to characterise any limit point as a solution to a martingale problem and (iii) to prove that this problem has a unique solution. We refer e.g. to [Szn84, Gär88, Oel84] for results in this direction. This kind of techniques have proven to adapt easily to different dynamics such as *moderately interacting diffusions* [JM98].

Remark 1.1.1. *In comparison with coupling arguments, one drawback of tightness/martingale formulations is that they do not provide with rate of convergence for the empirical measure μ_n towards μ . On the contrary, this approach allows for more general assumptions on the initial condition $\mu_{n,0}$: with coupling arguments, one generically needs to have i.i.d. (or at least independent) initial conditions $(\theta_0^{1,n}, \dots, \theta_0^{n,n})$ whereas tightness arguments only require that the initial empirical measure $\mu_{n,0}$ converges to μ_0 (see e.g. [Gär88, Th. 1.5]). This remark, which may not be critical in the context of homogeneous mean-field systems as (1.1), will be crucial in the case of inhomogeneous versions of (1.1) that will be discussed below, as the exchangeability of the system will not be preserved by the dynamics. Therefore it will be essential to consider initial conditions that are not a priori exchangeable.*

1.2 Towards non-exchangeability

From a modelling point of view, system (1.1) suffers from several limitations.

1.2.1 Diffusions in random environment

The first limitation is that, by definition, the system (1.1) is exchangeable. In particular, one assumes the local dynamics F to be the same along the entire population. As far as applications are concerned, this assumption is questionable: to take only one biological example, there is evidence [Luo20] of a large diversity in the nature, morphology and behavior within a mesoscopic population of neurons (e.g. excitable vs inhibitory neurons within a same population). In this respect, it is natural to consider the following generalization: take first a sequence $(\omega_i)_{i \geq 1}$ in \mathbb{R}^p , ($p \geq 1$) and consider the system

$$d\theta_t^{i,n} = F(\theta_t^{i,n}, \omega_i)dt + \frac{1}{n} \sum_{j=1}^n \Gamma(\theta_t^{i,n}, \theta_t^{j,n}, \omega_i, \omega_j) dt + \sigma \cdot dB_t^i, \quad i = 1, \dots, n, \quad t \in [0, T]. \quad (1.9)$$

Here, ω_i has to be considered as some *disorder*, some local intrinsic inhomogeneity for each particle $\theta^{i,n}$, so that the local dynamics $F(\cdot, \omega_i)$ may vary with i . Although one convenient way to define the sequence $(\omega_i)_{i \geq 1}$ is through a random procedure (a *random environment*, e.g. $(\omega_i)_{i \geq 1}$ i.i.d. with law ν), the point of view will always be to see $(\omega_i)_{i \geq 1}$ as much as possible as a deterministic sequence (*quenched model*), in opposition to the *annealed model* where one integrates w.r.t. the law of the disorder. Whereas the quenched system (1.9) is no longer exchangeable, it nonetheless remains within the mean-field framework, up to the additional cost of extending the definition of μ_u : the interaction in (1.9) is still a functional of the *double-layer empirical measure* [DdH96]

$$\mu_{n,t}^{(\omega)} := \frac{1}{n} \sum_{j=1}^n \delta_{(\theta_t^{j,n}, \omega_j)}, \quad (1.10)$$

the superscript (ω) being to stress that this is defined for a fixed realisation of $(\omega) := (\omega_i)_{i \geq 1}$. Note here that this not a matter of simply extending the dimension of the problem: one could of course think that solving the system (1.9) in $(\mathbb{R}^d)^n$ is formally equivalent to considering the system solved now by $(\theta_t^i, \omega_{t,i})_i \in (\mathbb{R}^{d+1})^n$ with the additional constraint that $d\omega_{i,t} = 0$. But then, the difficulty lies on the initial condition: in the quenched model, $(\theta_0^i, \omega_i)_{i=1, \dots, n}$ is by no means a sequence of i.i.d. random variables. Hence, there is a crucial difficulty in understanding the effect of non-exchangeability on systems like (1.9).

1.2.2 A crucial example: phase oscillators

Definitions

One of the main example considered in the manuscript concerns the Kuramoto model and its extensions, firstly introduced in [Kur75] as idealized models for synchronization. Synchrony is an ubiquitous phenomenon in physics and biology (social behavior of insects [Buc88], lasers arrays [KLN⁺95], circadian rhythms [THHI07], pacemaker cells [MMJ87], neuronal networks [Wai10]). Phase oscillators models (that is, interacting angles, i.e. diffusions on the torus $\mathbb{T} := \mathbb{R}/[0, 2\pi)$) have met a considerable interest in the physics literature [Kur75, ABPV⁺05, Str00], since, despite their simplicity, they have been able to reproduce several key features of synchronicity (incoherence, partial synchronization, [SM88a, SM88b, BS00], traveling waves, chimera states [MVS⁺14, AS06]). Take $\sigma = 1$ in the following for simplicity.

Definition 1.2.1 (Phase oscillators). *Let \mathbb{T} be the one dimensional torus and consider the family of diffusions in \mathbb{T}*

$$d\theta_t^{i,n} = \delta F(\theta_t^{i,n}, \omega_i)dt + \frac{1}{n} \sum_{j=1}^n J(\theta_t^{i,n} - \theta_t^{j,n}) dt + dB_t^i, \quad i = 1, \dots, n, \quad t \in [0, T]. \quad (1.11)$$

Here, $\delta \geq 0$ is a scaling parameter and, following the original definition of [Kur75], J is given as

$$J(\theta) = -K \sin(\theta), \quad \theta \in \mathbb{T}, \quad (1.12)$$

where $K \geq 0$ is the strength of interaction between particles. The corresponding nonlinear Fokker-Planck equation is described as $\mu_t(d\theta, d\omega) = p_t(\theta, \omega)d\theta\nu(d\omega)$, where p_t solves

$$\partial_t p_t(\theta, \omega) = \frac{1}{2} \partial_\theta^2 p_t(\theta, \omega) - \partial_\theta \left(p_t(\theta, \omega) \left\{ \int J * p_t(\cdot, \omega') \nu(d\omega') + \delta F(\theta, \omega) \right\} \right), \quad \theta \in \mathbb{T}, \quad \omega \in \text{Supp}(\nu) \quad (1.13)$$

where $h * \mu(\cdot) = \int h(\cdot - \theta') \mu(d\theta')$ denotes the convolution. Equation (1.13) is nothing else than a disordered version of the general nonlinear Fokker-Planck equation (1.4), where $p_t(\cdot, \omega)$ stands for the density (in a infinite population) of oscillators having the same disorder ω .

We will concentrate mostly on the following particular cases of Definition 1.2.1:

Example 1.2.2 (Kuramoto model, [Kur75, ABPV⁺05]). Take $F(\theta, \omega) := \omega$ in (1.11). The Kuramoto model is described as

$$d\theta_t^{i,n} = \delta\omega_i dt + \frac{1}{n} \sum_{j=1}^n J(\theta_t^{i,n} - \theta_t^{j,n}) dt + dB_t^i, \quad i = 1, \dots, n, \quad t \in [0, T]. \quad (1.14)$$

and the nonlinear Fokker-Planck equation reads

$$\partial_t p_t(\theta, \omega) = \frac{1}{2} \partial_\theta^2 p_t(\theta, \omega) - \partial_\theta \left(p_t(\theta, \omega) \left\{ \int J * p_t(\cdot, \omega') \nu(d\omega') + \delta\omega \right\} \right), \quad \theta \in \mathbb{T}, \quad \omega \in \text{Supp}(\nu). \quad (1.15)$$

In this setting, (1.14) models two contradictory effects: the interaction kernel $J(\cdot)$, that tends to make all phases $\theta^1, \dots, \theta^n$ equal (which is the basic meaning that one would give to the notion of synchronization) and the presence of a local frequency ω_i for each particle, that tends to deviate each particle from the common center of synchronisation. Note that this model is intrinsically invariant by rotation: one only needs to know the behavior of the particles $(\theta^1, \dots, \theta^n)$ up to a global phase translation on the torus \mathbb{T} .

Example 1.2.3 (Active rotators, [SK86a]). Suppose for simplicity that $\omega \equiv 0$ and take $F(\theta, \omega) = -\partial_\theta V(\theta)$, for some smooth potential V on \mathbb{T} in (1.11). In this case, the nonlinear Fokker-Planck equation reads

$$\partial_t p_t(\theta) = \frac{1}{2} \partial_\theta^2 p_t(\theta) - \partial_\theta (p_t(\theta) (J * p_t(\theta))) + \delta \partial_\theta (p_t(\theta) \partial_\theta V(\theta)), \quad \theta \in \mathbb{T}. \quad (1.16)$$

This models is a slight generalization of the Kuramoto model in the sense that one isolated particle has now a nontrivial dynamics on \mathbb{T} , driven by the potential V .

Synchronisation in the Kuramoto model

The first step towards a rigorous meaning of synchronization for (1.14) was carried out by Kuramoto [Kur75] and Sakaguchi [Sak88] (see also [ABPV⁺05, dH00] and references therein). Synchronization reads in terms of the existence of nontrivial stationary solutions q to (1.15). Crucial features of evolution (1.15) are captured by order parameters $r_t \geq 0$ and $\psi_t \in \mathbb{T}$ defined by:

$$r_t e^{i\psi_t} = \int_{\mathbb{T}} e^{i\theta} p_t(\theta, \omega) d\theta d\nu(\omega). \quad (1.17)$$

The quantity $r_t \in [0, 1]$ captures the intensity of synchronization of a solution, the two extreme examples being (i) $r_t = 0$ when $p_t = \frac{1}{2\pi}$ is the uniform measure on \mathbb{T} (representing total lack of synchronisation, or incoherence) and (ii) $r_t = 1$ when p_t is a Dirac (representing complete synchronisation within the population). The variable ψ_t identifies the center of synchronization (at least for unimodal profiles). Reformulating the interaction in (1.15) in terms of (r_t, ψ_t) in (1.17) and noting that (1.15) is invariant by rotation, we see that any stationary solution of (1.15) is necessarily of the form $q_{r,\psi}(\theta, \omega) = q_{r,0}(\theta - \psi, \omega)$ where $r \geq 0$ solves the stationary version of (1.17):

$$r = \int_{\mathbb{T}} \cos(\theta) q_{r,0}(\theta, \omega) d\theta d\nu(\omega) \quad \text{and} \quad 0 = \int_{\mathbb{T}} \sin(\theta) q_{r,0}(\theta, \omega) d\theta d\nu(\omega). \quad (1.18)$$

Any solution r to the nonlinear fixed-point relation (1.18) provides a stationary solution to (1.15) (e.g. $r = 0$ always solves (1.18) which corresponds to the *incoherent profile* $q(\cdot) \equiv \frac{1}{2\pi}$, see Figure 1.1).

The reversible case $\delta = 0$. When $\delta = 0$, the dynamics of (1.11) is *reversible* with respect to the Gibbs measure $\rho_n(d\theta_1, \dots, d\theta_n) := \frac{1}{Z_n} \exp\left(-\frac{K}{\sigma^2 n} \sum_{i,j=1}^n \cos(\theta_i - \theta_j)\right) d\theta_1 \dots d\theta_n$ (that is, the classical mean-field XY model [CR16]). One is left with the following version of the Fokker-Planck equation

$$\partial_t p_t(\theta) = \frac{1}{2} \partial_\theta^2 p_t(\theta) - \partial_\theta (p_t(\theta) (J * p_t)(\theta)), \quad \theta \in \mathbb{T}. \quad (1.19)$$

The fixed-point problem (1.18) reduces to

$$r_0 = \Psi_0(2Kr_0) \quad \text{where} \quad \Psi_0(x) := \frac{I_1(x)}{I_0(x)}, \quad (1.20)$$

where $I_k(x) = \frac{1}{2\pi} \int_{\mathbb{T}} \cos(\theta)^k e^{x \cos(\theta)} d\theta$ is the modified Bessel function of order k . The mapping Ψ_0 is increasing, concave [Pea81], with derivative at 0 equal to $\frac{1}{2}$. Consequently if $K \leq K_c := 1$, $r_0 = 0$ is the unique solution of (1.20) and $q(\cdot) \equiv \frac{1}{2\pi}$ (*incoherent solution*) is the only stationary solution of (1.19). If $K > 1$, we get in addition a circle (by rotation invariance) of *synchronized solutions* (Figure 1.1)

$$M_0 := \left\{ q_{r,\psi}(\theta) = q_{r,0}(\theta - \psi), \quad \psi \in \mathbb{T} \right\}, \quad \text{with} \quad q_{r,0}(\theta) = \frac{e^{2Kr \cos(\theta)}}{\int_{\mathbb{T}} e^{2Kr \cos(\theta)} d\theta}, \quad (1.21)$$

where $r > 0$ is the unique positive solution to (1.20).

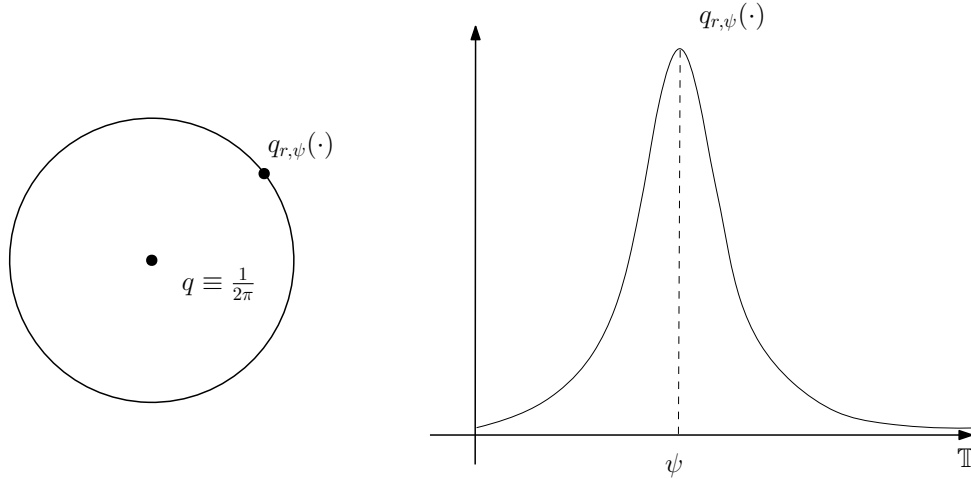


Figure 1.1 – Synchronisation in the mean-field Kuramoto model: transition from incoherence to synchrony in (1.19) reads in term of the unique existence of the incoherent solution $q \equiv \frac{1}{2\pi}$ when $K \leq 1$ to its coexistence with the circle M_0 of synchronized solutions, all translations of the same nontrivial pattern given in (1.21), when $K > 1$.

The disordered case $\delta > 0$. Adding some non-trivial disorder to (1.14) makes the system no longer reversible: we intrinsically deal with a nonequilibrium system. For a symmetric law ν , stationary solutions to (1.15) can be still written in a semi-explicit way [Sak88, dH00] as

$$q_{r,\psi}(\theta, \omega) := q_{r,0}(\theta + \psi, \omega) \quad \text{with} \quad q_{r,0}(\theta, \omega) := \frac{S(\theta, \omega, 2Kr)}{Z(\omega, 2Kr)} \quad (1.22)$$

with $S(\theta, \omega, x) = e^{G(\theta, \omega, x)} \left[(1 - e^{4\pi\omega}) \int_0^\theta e^{-G(u, \omega, x)} du + e^{4\pi\omega} \int_0^{2\pi} e^{-G(u, \omega, x)} du \right]$ and $G(u, y, x) = x \cos(u) + 2yu$, $Z(\omega, x) = \int_{\mathbb{T}} S(\theta, \omega, x) d\theta$ so that (1.18) turns into

$$r = \Psi_\nu(2Kr), \quad \text{where} \quad \Psi_\nu(x) := \int_{\mathbb{R}} \frac{\int_{\mathbb{T}} \cos(\theta) S(\theta, \omega, x) d\theta}{Z(\omega, x)} \nu(d\omega). \quad (1.23)$$

Solving (1.23) gets much more complicated even for simples instances of the disorder ν , as Ψ_ν may no longer be concave (see e.g. [Luç12, Prop. 3.11]). Uniqueness of a synchronized

circle is expected when ν is unimodal, but this has not been established. Nonetheless, a phase transition from incoherence to a single circle of synchronized patterns can be proven at least when $\text{Supp}(\nu) \subseteq [-\bar{\omega}, \bar{\omega}]$ for $\bar{\omega} \geq 0$ sufficiently small (see [L2, Lem. 2.3] or [Luç12, Prop. 3.13]): there is some $K_c > 0$ such that if $K \leq K_c$, the only stationary solution to (1.15) is $\frac{1}{2\pi}$ whereas for $K > K_c$, there is in addition a circle a stationary measures $M_\nu := \{q_{r,\psi}(\theta, \omega) = q_{r,0}(\theta - \psi, \omega), \psi \in \mathbb{T}, \omega \in \text{Supp}(\nu)\}$.

Stability of synchronisation in the reversible case $\delta = 0$

A considerable step further in the mathematical analysis of (1.19) was carried out in the seminal work of Bertini, Giacomin and Pakdaman in [BGP10]. As it is at the core of the analysis of the present manuscript, we briefly review the results of [BGP10] here. A crucial observation in [BGP10] is that, for $\delta = 0$, (1.19) is in gradient form: $\partial_t p_t(\theta) = \partial_\theta \left[p_t(\theta) \partial_\theta \left(\frac{\delta \mathcal{F}(p_t)}{\delta p_t(\theta)} \right) \right]$, where $\delta \mathcal{G}(q)/\delta q(\theta)$ is the L^2 -Fréchet derivative of the functional \mathcal{G} and $\mathcal{F}(q) := \frac{1}{2} \int_{\mathbb{T}} q(\theta) \ln q(\theta) d\theta - \frac{K}{2} \int_{\mathbb{T}^2} \cos(\theta - \theta') q(\theta) q(\theta') d\theta d\theta'$. Consequently, the energy decreases along the trajectories of (1.19): $\partial_t \mathcal{F}(p_t) = - \int_{\mathbb{T}} p_t(\theta) \left(\partial_\theta \frac{\delta \mathcal{F}(p_t)}{\delta p_t(\theta)} \right)^2 d\theta \leq 0$ (which is a weak form of stability of synchronisation). The main object of [BGP10] is to study the linear stability of nontrivial synchronised solutions (1.21) in the supercritical case $K > 1$. Setting $q := q_{r,0}$ (recall (1.21)) the linearised operator L_q^0 around the stationary solution q is defined as

$$L_q^0 u(\theta) = \frac{1}{2} \partial_\theta^2 u(\theta) + K \partial_\theta \left[q(\theta) \int_{\mathbb{T}} \sin(\theta - \theta') u(\theta') d\theta' + u(\theta) \int_{\mathbb{T}} \sin(\theta - \theta') q(\theta') d\theta' \right]. \quad (1.24)$$

By invariance by translation, one has $L_q^0 \partial_\theta q = 0$. The main result of [BGP10] is then

Theorem 1.2.4 (Th. 1.8 in [BGP10]). *The operator L_q^0 is essentially selfadjoint in $H_{1/q}^{-1}$, its spectrum is pure point and lies in $(-\infty, 0]$. The eigenvalue 0 is one-dimensional, spanned by $\partial_\theta q$ and the distance between 0 and the rest of the spectrum is $\lambda(K) > 0$.*

Here, $H_{1/q}^{-1}$ is the weighted Hilbert space of distributions u such that $u = \mathcal{U}'$, with $\mathcal{U} \in L^2$, such that $\int_{\mathbb{T}} \mathcal{U}(\theta)/q(\theta) d\theta = 0$, endowed with the norm

$$\|u\|_{-1,q} := \left(\int_{\mathbb{T}} \frac{\mathcal{U}(\theta)^2}{q(\theta)} d\theta \right)^{1/2}. \quad (1.25)$$

Explicit lower bounds on the spectral gap $\lambda(K)$ (in terms of K and r solving (1.18)) are also given. We will rely heavily on this result, which gives local stability of the manifold M_0 in (1.21), as it will be at the core of many perturbation arguments. The simplicity of the sine interaction in (1.19) has allowed for further detailed analysis of the nonlinear stability of M_0 , see [GPP12, GPPP12, FGVG14, DFGV16] for further details.

Phase oscillators and generalisations

In comparison with more elaborate oscillator systems in \mathbb{R}^d (e.g. the FitzHugh-Nagumo model in Example 1.3.2), models of phase oscillators on \mathbb{T} possess significant mathematical advantages. A crucial simplifying point of Example 1.2.2 is its *rotation invariance*: the fact that periodicity naturally appears in such models is essentially true *by construction*, whereas proving the existence of limit cycles in the FitzHugh-Nagumo model (Example 1.3.2) is a highly nontrivial problem which in itself requires a proof (we will address this point in Chapter 4). A second nice feature of Definition 1.2.1 comes from the simplicity of the sine interaction in (1.11), which allows for a direct parameterisation of the synchronised solutions as well as most of the time explicit computations. Thirdly, the fact that the state space \mathbb{T} in Definition 1.2.1 is compact is a considerable simplifying technical assumption, in comparison with the case of oscillators in \mathbb{R}^d (in the analysis of FitzHugh-Nagumo oscillators in [L9, L11], a considerable amount of work is spent to carefully control the behavior of the system at infinity). This last point (which is not a minor point in the analysis) is not only technical, it is also an illustration

of the difficulty of rigorously implementing a *phase reduction method* for a dynamical system with some (locally) stable limit cycle (that is how to simplify the analysis of its trajectory by projecting it along this limit cycle) in situation where both non-vanishing noise and mean-field interaction are present. This subject has attracted much attention in recent years (mostly in the physics literature, see e.g. [GTNE10, TNE09, Wai10], see also [Nak16] for a formal derivation of the Kuramoto model from general mean-field oscillators but also from a rigorous point of view [Poq14, GPS18]). We will discuss below the contributions of [L9, L12, L15] on this matter.

1.2.3 Spatially-structured interactions

In the definition of (1.9), the interaction between particles $\theta^{i,n}$ and $\theta^{j,n}$ may also depend on the respective value of the disorder ω_i and ω_j . A relevant interpretation is to consider $\omega_i = x_i$ as a fixed *spatial position* of the particle $\theta^{i,n}$ so that the nature and intensity of interaction between $\theta^{i,n}$ and $\theta^{j,n}$ depend on x_i and x_j : a generic choice would be to take $\Gamma(\theta, \theta', x, x') = \Gamma(\theta, \theta')\Psi(x, x')$ where Ψ is a nontrivial spatial kernel. Spatial extensions of mean-field dynamics are particularly relevant in a context of neuroscience where one accounts for the spatial organisation of neurons in the cortex (see [Mül17, CT18a, CDLO19, MSSZ20] and references therein). The corresponding system becomes

$$d\theta_t^{i,n} = F(\theta_t^{i,n})dt + \frac{1}{n} \sum_{j=1}^n \Gamma(\theta_t^{i,n}, \theta_t^{j,n}) \Psi(x_i, x_j) dt + \sigma \cdot dB_t^i, \quad i = 1, \dots, n, \quad t \in [0, T]. \quad (1.26)$$

Having in mind applications to neuroscience, the random variable θ_i would incorporate the value of the potential V_i of a neuron (as well as auxiliary variables accounting e.g. for the proportion of open channels along the axons of the neuron, as in the Hodgkin-Huxley model, see e.g. [BFFT12] for precise definitions). The variable x_i would then account for the position in \mathbb{R}^d of the neuron i within the cortex (of within cortical columns, see e.g. [Tou14] for more details). Another interpretation of this spatial variable comes from the modelling of the visual cortex: supposing that $x \in \mathbb{T}$, the spatial variable represents the susceptibility of a neuron in the visual cortex to accommodate for a given orientation [KE13]. We have been interested here to cases where Ψ presents singularities, the main examples being $x, y \in \mathbb{T}$ and $\Psi(x, y) = \mathbf{1}_{d(x,y) \leq R}$, $R \in [0, 1)$ (P nearest neighbor model) and $\Psi(x, y) = \frac{1}{d(x,y)^\alpha}$ for $\alpha \in (0, 1)$. We analyze in [L3, L6] the influence of the singularities of the kernel Ψ on the law of large numbers and the fluctuations of the empirical measure around its mean-field limit.

Content of Chapter 2

The purpose of Chapter 2 is to address the behavior of both (1.9) and (1.26) as $n \rightarrow \infty$, on any bounded time interval $[0, T]$. We first review [L1, L8] where the Law of Large Numbers, Central Limit Theorem and Large Deviation Principle for the empirical measure (1.10) are addressed, in a situation where the disorder (ω_i) has been frozen (quenched regime). We also briefly review the work [L4], taken from my PhD Thesis, where the long-time behavior of the resulting fluctuation process is analyzed in the context of the Kuramoto model with disorder. In a second time, the purpose of Chapter 2 will be to comment on the works [L3, L6] where the influence of singular spatial interactions is addressed at the level of Law of Large Numbers and fluctuations for the empirical measure of the system.

1.2.4 Perturbing the graph of interaction

The second limitation of (1.1) is that one assumes all-to-all coupling between particles: the graph of interaction is by definition the complete graph $K_n := (\{1, \dots, n\}, \{1, \dots, n\}^2)$, with homogeneous uniform strength of interaction of order $\frac{1}{n}$ along the whole population. This is again both questionable from a modelling point of view (and not true in the case of neuronal networks) and a strong mathematical limitation: if one removes any finite number of edges from the interaction graph K_n in (1.1), it is intuitively true that the asymptotic behavior of (1.2) as $n \rightarrow \infty$ remains unchanged (and one point of what follows will be to show that one can actually

remove far more than a finite number of edges). A natural extension of (1.1) is then to consider a generic graph $\mathcal{G}_n = (\mathcal{E}_n, \mathcal{V}_n)$ with set of vertices $\mathcal{E}_n := \{1, \dots, n\}$ and set of edges $\mathcal{V}_n \subset \{1, \dots, n\}^2$. Encoding the presence (resp. absence) of the edge $i \rightarrow j$ in \mathcal{V}_n by setting $\xi_{i,j}^n := 1$ (resp. $\xi_{i,j}^n := 0$), define the particle system $(\theta^{1,n}, \dots, \theta^{n,n})$ interacting on the graph \mathcal{G}_n by

$$d\theta_t^{i,n} = F(\theta_t^{i,n})dt + \frac{1}{d_i^n} \sum_{j=1}^n \xi_{i,j}^n \Gamma(\theta_t^{i,n}, \theta_t^{j,n}) dt + \sigma \cdot dB_t^i, \quad i = 1, \dots, n, \quad t \in [0, T], \quad (1.27)$$

Here, the interaction is renormalised by $d_i^n := \sum_{j=1}^n \xi_{i,j}^n$, the degree of vertex $i \in \mathcal{E}_n$ so that the interaction remains of order 1 as $n \rightarrow \infty$. Informally speaking, if the graph \mathcal{G}_n is sufficiently close to the complete graph K_n (in a way to be made precise), one naturally expects the same asymptotics as $n \rightarrow \infty$ as for the pure mean-field case. On a general level, the main questions are

- Question 1.2.5.** *1. How universal the mean-field framework is? How much can we perturb the complete graph of interaction K_n of (1.1) into some graph \mathcal{G}_n and nonetheless conserve similar asymptotics (in particular the same mean-field limit (1.4)) for the empirical measure of (1.27) as $n \rightarrow \infty$? At which level is this universality true? law of large numbers, fluctuations, large deviations?*
- 2. Is this universality annealed or quenched? Are the previous claimed asymptotics true for any fixed realisation of the possibly random graph \mathcal{G}_n ?*
- 3. Is it possible to quantify the proximity of μ_n to its mean-field limit μ in terms of the proximity between \mathcal{G}_n and K_n ? for which graph topology? what does it imply on the local or global structure of the graph \mathcal{G}_n ? To which extent can we consider diluted graphs?*

Again, contrary to the classical mean-field setting (1.1), exchangeability is lost for (1.27): the main difficulty regarding Question 1.2.5 (which is already present once only one edge is removed from K_n) comes from the observation that the interaction in (1.27) is now no longer a functional of the empirical measure (1.2), but rather of a collection of *local* empirical measures

$$\mu_{n,t}^i := \frac{1}{d_i^n} \sum_{j=1}^n \xi_{i,j}^n \delta_{\theta_t^{j,n}}, \quad i = 1, \dots, n, \quad (1.28)$$

Each μ_n^i accounts for particles within distance 1 of node i in the graph \mathcal{G}_n . The problem is that the dynamics of these local empirical measures depend themselves on higher order empirical measures (that involve particles at distance 2, etc.): a whole hierarchy of empirical measures arises and the difficulty is to find a way to properly close this expansion as $n \rightarrow \infty$.

Content of Chapter 3

The purpose of Chapter 3 will be to address this question of universality of the mean-field framework for several general classes of graphs \mathcal{G}_n : firstly homogeneous graphs, both at the level of the law of large numbers [L5] and fluctuations [L14] of the empirical measure and secondly in the case of inhomogeneous graphs [L10].

1.3 Long-time periodicity for mean-field systems

All the previous results concerned the asymptotics of (perturbations of) the particle system (1.1) as $n \rightarrow \infty$, on a bounded time interval $[0, T]$. A crucial remark (that is already relevant for the homogenous case (1.1)) concerns the dependence in T in the estimate (1.7). Without prior knowledge on F , Γ and σ , the Grönwall argument leading to (1.7) gives an exponential constant $C(T) = e^{CT}$ for some $C = C(F, \Gamma, \sigma) > 0$ so that this estimate remains relevant only for times that are logarithmic in n , i.e. $T \sim c \ln n$ for some $c > 0$ sufficiently small. A natural and important question is then

Question 1.3.1. *Is it possible to obtain a better constant $C(T)$ in (1.7), even ultimately independent of T ? In other words, does μ_t remain a proper approximation of $\mu_{n,t}$ on time scales that go beyond $t \sim c \ln n$?*

The answer to Question 1.3.1 is intrinsically model-dependent. A case that has received particular attention in the recent years concerns *granular type models*, where $F = -\nabla V$ and $\Gamma = -\nabla W * \cdot$, (see e.g. [Tug13]). Under appropriate hypotheses on regularity and convexity at infinity of the potentials V and W , the answer to Question 1.3.1 is mostly positive: due to the gradient structure of the model, the use of functional inequalities has proven to be a powerful way to obtain exponential rate of convergence to equilibrium for such models as well as uniform in time propagation estimates, see e.g. [BGM10, CGM08, Mal03, MT19, GM21, CMV03, BGG12] and references therein. Note however that this class of dynamics is intrinsically *reversible* and may not be transposed without difficulty to the examples we have in mind, which are non reversible.

However, and this is one of the major difficulties here, it is easy to see that the answer to Question 1.3.1 is negative in general as the limits as $n \rightarrow \infty$ and $T \rightarrow \infty$ do not commute (see the discussion in [L5]): noise has a cumulative effect on the empirical measure of the system (1.2) on time scales beyond $\ln n$, so that the dynamics of the microscopic system (1.1) and its nonlinear counterpart (1.5) generically diverge on a time-scale that is not bounded. One crucial point is to understand how to take into account this cumulative influence of noise so as to derive the correct asymptotic for the empirical measure (1.2) from (1.4).

1.3.1 Periodicity in noisy mean-field systems

The next point of this manuscript is to question this absence of uniform propagation of chaos in case (1.4) exhibits synchronized periodic behaviors. Collective oscillations reflects a common feature of self-organization in complex systems (e.g. circadian rhythms [KYW⁺10], synchrony in neural networks in motor systems or pacemaker cells [Luo20, Buz06] or pathologies of the brain [Buz04, TPVE20, Wan10]). Several papers in physics have studied the existence of coherent structures (e.g. coherence resonance [LSG99], pattern formation [GaOS99] or wave propagation [LGONSG04]) for systems exhibiting *excitability properties* [LGONSG04, Bre14a, Izh07]. Excitability is a key feature of neuronal dynamics, as it is at the basis of signal processing and transmission in neuronal systems. Roughly speaking, excitability refers to the ability of a neuron to emit spikes (oscillations) in the presence of perturbations (noise and/or external input) whereas this neuron would be at rest (steady state) without perturbation. A crucial question, that is still poorly understood [OBH08, BH99], concerns the role of noise and interaction in the emergence and stability of periodic behaviors. Noise can be intrinsic to each neuron (e.g. the random switching of ion channels [GT12, PTW10]) or can come from the random input from other neurons.

A prominent example of excitable dynamics (and one of the main examples that we will consider here) concerns the FitzHugh-Nagumo model [Fit61, NAY62, RsGGt00, BFFT12, BPG⁺03], introduced as a two-dimensional idealization of neuronal dynamics:

Example 1.3.2 (FitzHugh-Nagumo model). *Take $d = 2$, $\theta = (v, w)$ and*

$$F(v, w) = \left(v - \frac{v^3}{3} - w, \frac{1}{c} (v + a - bw) \right) \quad (1.29)$$

with chosen constants $a \in \mathbb{R}$ and $b, c > 0$. Here, v stands for the voltage whereas w is the recovery variable.

The existence of random attractors and the effect of noise on random dynamical systems has a longstanding history (see e.g. [FGS17, SV18, LMS22] for further references on the subject). Talking about periodic behaviors for mean-field systems, it is important to distinguish the microscopic system (1.1) from its mean-field limit (1.5). One should not expect the microscopic system (1.1) to have a periodic law for any (possibly large) but finite n : for fixed $n \geq 1$, system (1.1) is a time-homogeneous Markov process and hence either possesses an invariant

probability measure that is globally asymptotically stable or escapes each compact set for $t \rightarrow \infty$ with probability 1 [Kha12]. This is however no longer true for the non Markovian nonlinear diffusions (1.5). This was originally observed by Scheutzow in [Sch85] who provided examples of nonlinear diffusions (in dimension 1) exhibiting oscillatory behaviors in presence of noise (whereas no oscillations are present without noise). It is also shown in [Sch85] that noise may induce the existence of a globally stable invariant measure for a nonlinear diffusion, that would not exist in absence of noise. Existence of periodic behaviors for (1.5) was also proven in [Sch86] in the 2-dimensional Brusselator model. Note here that the *Active rotators model* mentioned in Example 1.2.3 possesses similar excitability features in a mathematically simpler context [SSK88, GPPP12, EK86]. Emergence of periodic behaviors induced by noise and interaction for this model has been proven in the seminal work [GPPP12]. The crucial argument of [GPPP12] (and a central notion for us) relies on the concept of *Stable Normally Hyperbolic Manifold* and on the stability of these dynamical structures under perturbation.

Content of Chapter 4

The strategy of proof of [Sch85] relies on the fact that the nonlinear diffusion considered in the paper admits Gaussian solution so that the dynamics of the whole process boils down to the analysis of its mean and variance (see [THF12] for similar ideas), which is not a path that one can follow here. We review in Chapter 4 several works proving the emergence of rotating waves for various instances of (1.4): we first deal in [L2] with the simpler case of Kuramoto type oscillators (Example 1.2.2) in the presence of asymmetric frequencies. In the case of the FitzHugh-Nagumo model (Example 1.3.2), the existence of periodic solutions to (1.4) was only remarked based on numerical simulations [BFFT12, nT20] but a rigorous proof was lacking. The main contribution of [L9, L11] is precisely to prove the emergence of periodic solutions to (1.4) for Example 1.3.2, in both cases of full connectivity (elliptic case [L9]) and also in the more difficult kinetic case where noise and interaction are only present on the voltage variable v in [L11]. We complement these existence results in [L12] with stability and regularity estimates that will be necessary for Chapter 5.

1.3.2 Going beyond the finite time scale

As already mentioned, even if the nonlinear dynamics (1.4) might possess periodic solutions, no such periodic behavior is possible for the finite particle system (1.1). However, taking advantage of the stability of these limit cycles, it is possible to show that the empirical measure (1.2) stays for a long time in the vicinity of such structures. The second step is then to understand on which time scale one observes a nontrivial dynamics for (1.2) along these limit cycles and how to characterize this dynamics. The point is therefore to look at the empirical measure (1.2) when both n and t go to ∞ simultaneously, that is, at the process $(\mu_{n,\alpha_n t})_{t \in [0,T]}$, where $\alpha_n \rightarrow \infty$ is an appropriate scaling parameter. This program was first initiated by Bertini, Giacomin and Poquet in [BGP14] for the Kuramoto model without disorder. We have further developed and extended this strategy for various instances of mean-field models in [L7] and [L15].

Content of Chapter 5

We address this question in Chapter 5, first about the Kuramoto model [L7] in presence of inhomogeneities (where we show the emergence of traveling waves induced by the quenched environment on the time scale $\alpha_n \sim \sqrt{n}$), and secondly in a general case [L15], where we show that, for $\alpha_n = n$, the empirical measure $\mu_{n,nt}$ follows the periodic solution of the Fokker-Planck equation (1.4), up to a diffusive correction term that scales as a Brownian motion as $n \rightarrow \infty$.

1.4 Inhibition and oscillations for mean-field Hawkes processes

The last point of this manuscript concerns a recent interest in mean-field interacting point processes, with direct applications to neuroscience. In opposition to Hodgkin-Huxley or FitzHugh-

Nagumo models which directly model the evolution of the membrane potential of a neuron, a simplifying and mathematically convenient point of view is to consider action potentials as stereotyped in their shape and intensities [Bri88], so that the only remaining relevant information is the sequence of instants of spikes. This formalism has led to a large variety of models with jump dynamics, where any jump or reset encodes for a spiking event. The model we consider enters into the framework of Hawkes processes [HO74, DFH16], that models directly the stochastic intensity of a neuron through the following general formalism

$$\lambda_t^i = \Phi_i \left(\sum_{j \rightarrow i} \int_0^t h_{j \rightarrow i}(t-u) dZ_u^j \right), \quad (1.30)$$

where λ_t^i is the intensity of neuron i and Z_t^j is the counting process associated with neuron j . The form of the intensity λ_t^i in (1.30) accounts for the dependence of the activity of neuron i in the history of the whole system through the positive function Φ_i (the *spike rate function*) and $h_{j \rightarrow i}(t)$ (the *synaptic kernel* associated with the synapse between neurons j and i). This Hawkes formalism, originally introduced to model earthquake activity [HO74] has been sufficiently versatile to accommodate for various features such as age-dependent behaviors/refractory periods [Che17b, Che17a, RL20] or spatially-structured dynamics [CDLO19]. We refer to e.g. [BM96, DFH16, DL17, CDLO19, Che17b] and references therein for more details. A particularly important instance of (1.30) that is explicitly solvable concerns the linear case where $\Phi_i(x) = \mu_i + x$. Note that in this case one needs to assume that $h_{i \rightarrow j} \geq 0$ to make sure that the intensity λ_t^i remains nonnegative.

The main motivation here is to incorporate inhibition in Hawkes processes. Inhibition is a crucial feature in neuronal systems, as a key mechanism in the control and coordination of neuronal systems [Luo20]. The main strategy followed in the literature so far has been to model inhibition by using kernels $h_{j \rightarrow i}(\cdot)$ in (1.30) taking negative values [CGMT20, RL20, RBRTM13, BHS21] or positive kernels h (e.g. of Erlang type) multiplied by random and possibly negative coefficients [DL17, DLO19, PRS22]. This is what one could qualify as an *additive* inhibition, in the sense that the intensity of a neuron is the (temporal) additive superposition of several memory kernels $h_{i \rightarrow j}$. In order to keep λ_t^i positive, one crucially needs here to modulate this superposition by a positive (necessarily nonlinear) synaptic function Φ_i , (e.g. $\Phi_i(x) = \mu^i + x_+$ where $x_+ = \max(x, 0)$). Considering both signed $h_{i \rightarrow j}$ and nonlinear Φ_i is a major step forward in the difficulty as it breaks totally the monotonicity that one has for purely excitatory Hawkes processes. Understanding the temporal dynamics of (1.30) is very difficult and very few rigorous results exist on this matter (see [CGMT20, CCC22, RL20, DL17]).

Content of Chapter 6

We review in Chapter 6 the article [L13] where the main novelty is to propose an alternative mechanism for inhibition that is *multiplicative*. The main advantage of this strategy is to circumvent the difficulty mentioned above of models with *additive inhibition* as the multiplicative structure allows to keep the positivity of the kernels $h_{i \rightarrow j}$ and hence the monotonicity of the system. We prove in [L13] propagation of chaos for a system of n coupled Hawkes processes with multiplicative inhibition as $n \rightarrow \infty$ towards a system of inhomogeneous Poisson processes with coupled intensities. Secondly, we give in [L13] sufficient conditions for the convergence of the mean-field intensities as $t \rightarrow \infty$, revealing in particular biological features of inhibition (notably that inhibition prevents supercriticality). We give finally insights and numerical evidence that inhibition may lead to oscillations, leaving the ground for future proof of this phenomenon.

Chapter 2

Large population behavior of inhomogeneous particles

We analyse in this chapter the large n behavior, on a bounded time interval $[0, T]$ of the disordered system

$$d\theta_t^{i,n} = F(\theta_t^{i,n}, \omega_i)dt + \frac{1}{n} \sum_{j=1}^n \Gamma(\theta_t^{i,n}, \theta_t^{j,n}, \omega_i, \omega_j) dt + \sigma \cdot dB_t^i, \quad i = 1, \dots, n, \quad t \in [0, T], \quad (2.1)$$

where $n \geq 1$, $T > 0$, $\theta_t^{i,n} \in \mathcal{X}$, $\omega_i \in \mathcal{E}$ for each $i = 1, \dots, n$ and (B_i) is a sequence of i.i.d. Brownian motions in \mathcal{X} , defined on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$. Here, $\mathcal{X} = \mathbb{R}^d$ for some $d \geq 1$ but it is also possible to assume that $\mathcal{X} = \mathbb{T}^d$ is the compact torus. In a same way, $\mathcal{E} = \mathbb{R}^p$ for some $p \geq 1$ and $F : \mathcal{X} \times \mathcal{E} \rightarrow \mathcal{X}$ and $\Gamma : \mathcal{X}^2 \times \mathcal{E}^2$ are the local term and the interaction kernel respectively. The system (2.1) lives in a random environment, given by a sequence of random variables $(\omega_i)_{i \geq 1}$ defined on a common probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$. Two different models are possible: firstly *the annealed model*, where one analyses the behavior of (2.1) under the *joint law* $\mathbf{P} \otimes \mathbb{P}$ of the noise and disorder. We are rather interested in the *quenched model*, where we consider (2.1) under \mathbf{P} only, for a fixed realisation of the disorder $(\omega_i)_{i \geq 1}$, as the point of view will always be to see (ω_i) as much as possible as a deterministic sequence. Note that exchangeability of the particle system $(\theta^{1,n}, \dots, \theta^{n,n})$ is obviously lost. Following the original formalism of Dai Pra and den Hollander [DdH96, dH00], define the *double-layer empirical flow*

$$\mu_{n,t}^{(\omega)} := \frac{1}{n} \sum_{j=1}^n \delta_{(\theta_t^{j,n}, \omega_j)}, \quad t \in [0, T] \quad (2.2)$$

the superscript (ω) being to stress that μ_n (seen as an element of $\mathcal{C}([0, T], \mathcal{P}(\mathcal{X})) \times \mathcal{E}$) is defined for a fixed realisation of $(\omega) := (\omega_i)_{i \geq 1}$.

Remark 2.0.1. Note also that Dai Pra and den Hollander originally considered (2.1) for $\mathcal{X} = \mathcal{E} = \mathbb{R}$ and $F(\theta, \omega) = -\partial_\theta g(\theta, \omega)$ and $\Gamma(\theta, \theta', \omega, \omega') = \partial_\theta f(\theta - \theta', \omega, \omega')$. In the annealed setting, this corresponds to a reversible dynamics, with unique equilibrium the Gibbs measure $\frac{1}{Z_n} \exp(-H_n(\boldsymbol{\theta}, \boldsymbol{\omega}))$, with $H_n(\boldsymbol{\theta}, \boldsymbol{\omega}) = \frac{1}{2n} \sum_{i,j=1}^n f(\theta_i - \theta_j, \omega_i, \omega_j) + \sum_{i=1}^n g(\theta_i, \omega_i)$. This particular assumption, motivated by considerations of Large Deviations, is not needed at the level of the law of large numbers and fluctuations for the empirical measure.

2.1 Quenched limit for the empirical flow

2.1.1 The McKean-Vlasov process and the nonlinear Fokker-Planck equation: existence, uniqueness and regularity.

Suppose that there exists a probability measure μ_0 in $\mathcal{X} \times \mathcal{E}$ such that the sequence $(\theta_0^{i,n}, \omega_i)$ satisfies $\mu_{n,0} = \frac{1}{n} \sum_{i=1}^n \delta_{(\theta_0^{i,n}, \omega_i)} \xrightarrow[n \rightarrow \infty]{} \mu_0$ ($d\theta, d\omega$) = $\mu_0^\omega(d\theta)\nu(d\omega)$ for the topology of weak convergence on $\mathcal{P}(\mathcal{X} \times \mathcal{E})$, where ν is the marginal of μ_0 on \mathcal{E} . Note that this implies in particular the

weak convergence of the empirical measure of the disorder sequence $(\omega_i)_{i \geq 1}$ to ν :

$$\nu_n^{(\omega)} := \frac{1}{n} \sum_{i=1}^n \delta_{\omega_i} \xrightarrow{n \rightarrow \infty} \nu. \quad (2.3)$$

The nonlinear process associated to (2.1) is then

$$d\bar{\theta}_t^\omega = F(\bar{\theta}_t^\omega, \omega) dt + \int \Gamma(\bar{\theta}_t^\omega, \theta', \omega, \omega') \mu_t(d\theta', d\omega') dt + \sigma \cdot dB_t, \quad t \in [0, T], \quad (2.4)$$

where $\mu_t(d\theta, d\omega)$ is the *joint law* of $(\bar{\theta}_t^\omega, \omega)$. Contrary to the homogeneous case (1.3), the dependence in the local variable ω_i in both the dynamics term $F(\cdot, \omega_i)$ and the interaction term $\Gamma(\cdot)$ persists at the limit: the limit process $\bar{\theta}^\omega$ still depends on its local inhomogeneity variable ω . This $\mu_t \in \mathcal{P}(\mathcal{X} \times \mathcal{E})$ is then the natural limit of the empirical flow (2.2) and solves the nonlinear Fokker-Planck equation, written in a weak form:

$$\begin{aligned} \langle \mu_t, f \rangle &= \langle \mu_0, f \rangle + \int_0^t \frac{1}{2} \langle \mu_s, \operatorname{div}(\sigma \sigma^\dagger \nabla f) \rangle ds \\ &\quad + \int_0^t \left\langle \mu_s, \nabla f \cdot \left\{ \int \Gamma(\cdot, \theta', \cdot, \omega') \mu_s(d\theta', d\omega') + F \right\} \right\rangle ds \end{aligned} \quad (2.5)$$

Note that the marginal ν of μ_0 on \mathcal{E} is preserved by the dynamics (2.5) at any time $t > 0$.

Well-posedness result and regularity estimates

The question of wellposedness of (2.4) and (2.5) has arisen repeatedly in all of the works mentioned in this manuscript, under varying hypotheses on the state space \mathcal{X} or coefficients F and Γ . Hence, we take the opportunity of this paragraph to give a unifying set of hypotheses that are representative of the framework adopted throughout this manuscript. We require that both F and Γ are regular, that $\theta \mapsto F(\theta, \omega)$ is locally Lipschitz for all ω and that the following one-sided Lipschitz condition holds: there exists some $C_F > 0$ such that uniformly on ω ,

$$\langle F(\theta, \omega) - F(\theta', \omega), \theta - \theta' \rangle \leq C_F |\theta - \theta'|^2 \quad (2.6)$$

Note that (2.6) includes the case of the FitzHugh-Nagumo model (Example 1.3.2), which is not globally Lipschitz. Here, the inhomogeneity may lie in the choice of the parameters $\omega_i := (a_i, b_i, c_i)$ in (1.29), so that (2.6) holds as long as ω remains bounded. Suppose also some polynomial control on F

$$|F(\theta, \omega)| \leq C_F (1 + |\theta|^k + |\omega|^l) \quad (2.7)$$

Require finally that Γ is Lipschitz, uniformly on the environment

$$|\Gamma(\theta_1, \theta'_1, \omega, \omega') - \Gamma(\theta_2, \theta'_2, \omega, \omega')| \leq C_\Gamma (|\theta_1 - \theta'_1| + |\theta_2 - \theta'_2|) \quad (2.8)$$

together with some sublinear control $|\Gamma(\theta, \theta', \omega, \omega')| \leq C_\Gamma (1 + |\theta| + |\theta'|)$. Note that this set of assumptions on Γ not only include bounded interactions but also the linear case $\Gamma(\theta, \theta') = \theta - \theta'$.

Proposition 2.1.1. *Under the previous assumptions, there is a unique weak solution μ in $\mathcal{P}(\mathcal{C}([0, T], \mathcal{X}) \times \mathcal{E})$ to the nonlinear Fokker-Planck equation (2.5) with initial condition μ_0 . Moreover, this solution μ is such that for ν almost every ω , $\mu^\omega(d\theta)$ is the law of the nonlinear process $(\bar{\theta}_t^\omega)_{t \in [0, T]}$ given in (2.4).*

In the homogeneous case (1.3), a vast literature exists on similar well-posedness results (see e.g. [McK67, Oel84, Szn91, Gär88]). The first result in the inhomogeneous case was originally stated in [DdH96, Th. 2], identifying μ solution to (2.5) as the unique zero of the Large Deviation rate function concerning the empirical measure (2.19), under some regularity of the initial condition and restrictive hypotheses on the coefficients F and Γ and the disorder. Concerning the present manuscript, similar well-posedness results have appeared on multiple occasions under various hypotheses: we refer in particular to [L1, Th. 2.5] (in the case $\mathcal{X} = \mathbb{T}$ with bounded coefficients), in [L8] (as the unique zero of a rate function), in [L3, Prop. 2.19] and [L6] in the case of diffusions with spatial structure. The present Proposition 2.1.1 can be seen as an adaptation of [L10, Prop. 2.7] (see also [L9, Lem. 4.1, Prop. 4.3] for similar arguments). We reproduce here the main lines of proof.

Main lines of proof of Proposition 2.1.1. The existence part of a weak solution to (2.5) follows from an adaptation of a usual fixed-point argument (which goes back to Sznitman [Szn91] in the homogeneous case): consider \mathcal{M}_ν as the set of probability measures $\mu \in \mathcal{P}(\mathcal{C}([0, T], \mathcal{X}) \times \mathcal{E})$ with fixed marginal on \mathcal{E} equal to ν , equipped with the Wasserstein-like metric (see also another possible alternative distance in [L3, Def. 3.1]) $\delta_T(\mu_1, \mu_2) = \sup_{\omega \in \mathcal{E}} \inf_{\pi} \left(\sup_{s \leq T} \int_{\mathcal{X}} |\vartheta_{1,s}^\omega - \vartheta_{2,s}^\omega|^{2k} \pi(d\vartheta_1, d\vartheta_2) \right)^{\frac{1}{2k}}$, where the infimum is taken over all coupling π under which $\vartheta_1^\omega \sim \mu_1^\omega$ and $\vartheta_2^\omega \sim \mu_2^\omega$. Fix now some $m = (m_t(d\theta, d\omega))_{t \in [0, T]} \in \mathcal{M}_\nu$ and consider Θ_m as the law of $(\theta_m^\omega, \omega)$ where

$$d\theta_{m,t}^\omega = F(\theta_{m,t}^\omega, \omega)dt + \int \Gamma(\theta_{m,t}^\omega, \theta', \omega, \omega') m_t(d\theta', d\omega') dt + \sigma \cdot dB_t. \quad (2.9)$$

By the Lipschitz properties of the coefficients, $m \mapsto \Theta_m$ is a contraction for the above metric and its fixed-point $\mu = \Theta_\mu$ gives a weak solution to (2.5). The uniqueness part of the analysis has also been the subject of many works in the absence of disorder, see [Oel84, DdH96, Gär88]. In our context, a possibility is to use a propagator argument (see [L10, App. A.2] for details): for all regular test function f , and $s \leq t \leq T$, define as

$$P_{s,t}f(\theta, \omega) := \mathbf{E}f(\theta_s^t(\theta, \omega), \omega), \quad (2.10)$$

where $s \mapsto \theta_s^t(\theta, \omega)$ solves (2.9) for $m = \mu$, with initial condition $\theta_s^s = \theta$. Now, if one takes another solution μ' to (2.5) with the same initial condition, the point is to apply Ito's formula to $s \mapsto P_{s,t}f(\theta_{\mu',s}, \omega)$, where $\theta_{\mu',s}$ solves (2.9) for $m = \mu'$. Using the fact that, under the additional assumption that F is globally Lipschitz, $s \mapsto P_{s,t}$ satisfies a Backward Kolmogorov equation and $P_{s,s}f = f$, we obtain finally that

$$\int f(\theta, \omega) \{ \mu_t^\omega(d\theta) - \mu_t^{\omega'}(d\theta) \} = \int_0^t \left\{ \int \nabla P_{s,t}f(\theta, \omega) \cdot \Gamma(\theta, \theta', \omega, \omega') \mu_s^\omega(d\theta) \right\} \{ \mu_s^{\omega'}(d\theta') - \mu_s^{\omega'}(d\theta') \} \nu(d\omega') ds$$

The function integrated in the right hand side of the previous inequality is Lipschitz, uniformly on f with $\|f\|_{Lip} \leq 1$. Hence, it is controlled in terms of $C \int_0^t W_1(\mu_s^{\omega'}, \mu_s^{\omega'}) \nu(d\omega') ds$. Taking the supremum in f with $\|f\|_{Lip} \leq 1$, we retrieve also the Wasserstein distance $W_1(\mu_t^{\omega'}, \mu_t^{\omega'})$ on the lefthand side and a Grönwall lemma gives uniqueness. Note that one can remove the additional assumption that F is globally Lipschitz and work under (2.6) only by introducing the Yosida approximation of F , we refer to [L3, Section 7] and [L10, App. A.2] for further details. \square

Extension 2.1.2. *The well-posedness result of Proposition 2.1.1 is stated in the space of probability measures on trajectories. Based on the observation that, by Sobolev embeddings [AF03], $\mathcal{P}(\mathbb{R}^d) \subseteq H^{-r}(\mathbb{R}^d)$ when $r > \frac{d}{2}$ (here $H^{-r}(\mathbb{R}^d)$ is the dual of the set $H^r(\mathbb{R}^d)$ of test functions with derivatives of order r in L^2), an alternative approach would be to prove directly well-posedness of (2.5) in $H^{-r}(\mathbb{R}^d)$ (or more precisely in the affine space $\{u \in H^{-r}(\mathbb{R}^d), \langle u, 1 \rangle = 1\}$, noting that mass is preserved by (2.5)). We refer to [L12, Th. 1.2] where we follow such approach in the case of linear interaction, based on some fixed-point argument [SY13].*

When σ is nondegenerate, standard results for uniform parabolic PDEs [Aro68, Fri64] state that the unique solution μ_t to (2.5) is absolutely continuous w.r.t. $d\theta \otimes \nu(d\omega)$, $\mu_t(d\theta, d\omega) = p_t(\theta, \omega) d\theta \nu(d\omega)$, where p_t is the strong solution to

$$\begin{aligned} \partial_t p_t(\theta, \omega) &= \nabla \cdot (\sigma \sigma^\dagger \nabla p_t)(\theta, \omega) \\ &- \nabla \cdot \left(p_t(\theta, \omega) \left\{ \int \Gamma(\theta, \theta', \omega, \omega') p_t(\theta', \omega') d\theta' \nu(d\omega') + F(\theta, \omega) \right\} \right), \quad \theta \in \mathcal{X}, \quad \omega \in \text{Supp}(\nu) \end{aligned} \quad (2.11)$$

A convenient point of view is to see (2.11) as a system of McKean-Vlasov equations, whose unknown is $p_t(\cdot, \omega)$ for each $\omega \in \text{Supp}(\nu)$, which are all coupled through the integration w.r.t. ν . Smoothness results of $(t, \theta) \mapsto p_t(\theta, \omega)$ can be easily deduce from standard bootstrap methods (see [L2, Prop. 7.1] and [L9, Prop. 4.10]; see also [GPP12, Th 2.1] for more refined analyticity estimates on p_t in the case of the Kuramoto model). Note that it is also possible to obtain regularity estimates on $\omega \mapsto p_t(\theta, \omega)$, either directly from parabolic estimates [L2, Lem. 7.2] or with propagator techniques similar to the proof of Proposition 2.1.1, even in the case Γ is singular in ω, ω' , see [L10, App. A.3].

2.1.2 Quenched law of large numbers

The next point concerns the convergence of the empirical flow (2.2) towards the solution μ to (2.5). A very simple tightness argument in the case where $\mathcal{X} = \mathbb{T}$ is given in [L1]:

Proposition 2.1.3 (Theorem 2.5 in [L1]). *Assume that $\mathcal{X} = \mathbb{T}$, $\mathcal{E} = \mathbb{R}$, F , Lipschitz continuous in θ (uniformly in ω), uniformly continuous on $\mathbb{T} \times \mathbb{R}$, Γ continuous and bounded and suppose that the sequence $(\theta_0^{i,n}, \omega_i)_{i=1, \dots, n}$ is such that $\mu_{n,0}^{(\omega)}$ converges in law to μ_0 for the topology of weak convergence in $\mathcal{P}(\mathbb{T} \times \mathbb{R})$. Suppose also that (recall the definition of ν in (2.3)) $\int_{\mathbb{R}} \sup_{x \in \mathbb{T}} |F(x, \omega)| \nu(d\omega) < \infty$, then the empirical flow $\mu_n^{(\omega)}$ converges in law, in $\mathcal{C}([0, T], \mathcal{P}(\mathbb{T} \times \mathbb{R}))$ to μ unique solution to (2.5).*

This convergence is a quenched result, i.e. valid for a fixed realisation of the environment (ω_i) . Note that we do not require the initial condition $(\theta_0^{1,n}, \dots, \theta_0^{n,n})$ to be i.i.d. (in particular they may very well depend on the environment ω_i), all we need is the convergence of the empirical measure $\mu_{n,0}$ as $n \rightarrow \infty$. Proof of Proposition 2.1.3 follows a very simple strategy: first prove tightness in $\mathcal{C}([0, T], \mathcal{P}(\mathbb{T} \times \mathbb{R}))$ endowed with the topology of vague convergence [RC86, Bil99], then prove that the mass $\langle \mu_n, 1 \rangle$ is conserved as $n \rightarrow \infty$ and finally identify any limit point μ as the unique solution to (2.5).

Extensions: Proposition 2.1.3 has been stated for simplicity in [L1] in the case $\mathcal{X} = \mathbb{T}$. Direct extensions of this method are possible to $\mathcal{X} = \mathbb{R}^d$, provided one assumes further that $(\theta, \theta') \mapsto \Gamma(\theta, \theta', \omega, \omega')$ is with compact support. An alternative to this result would be to use some propagator method as developed in [L3] and [L14]: an easy adaptation of [L14, Th. 2.4] also shows a similar quenched convergence of the empirical flow, under assumptions (2.6), (2.7) and (2.8) on F and Γ , under the additional assumption that Γ may be written as $\Gamma(\theta, \theta', \omega, \omega') = \sum_{p=0}^{+\infty} c_p \alpha_p(\theta, \omega) \beta_p(\theta', \omega')$, for some coefficients $(c_p)_{p \geq 0}$ such that $\sum_{p \geq 0} |c_p| < \infty$ and α_p and β_p functions bounded by 1 with Lipschitz constants smaller than 1 (see [DdH96, Eq. (1.31)] for the same hypothesis). In such a case, it is easy to prove that provided $\mu_{n,0}$ converges weakly to μ_0 , one has $\mathbf{E}[\sup_{s \leq T} d_{BL}(\mu_{n,s}, \mu_s)] \xrightarrow[n \rightarrow \infty]{} 0$, almost surely w.r.t. the disorder sequence (ω_i) . We refer to [L14, App. C] for more details on this approach.

Remark 2.1.4. *An alternative approach to Proposition 2.1.3 would be to proceed with a coupling argument: define the sequence $(\bar{\theta}^{\omega_1}, \dots, \bar{\theta}^{\omega_n})$ as independent copies of the nonlinear process (2.4), where for each $i = 1, \dots, n$, $\bar{\theta}^{\omega_i}$ is driven by the same Brownian motion B_i as for $\theta^{i,n}$ solving (2.1), with the same disorder ω_i and with identical initial condition $\bar{\theta}_0^{\omega_i} = \theta_0^{i,n}$. Then, an easy adaptation of [L10, Prop. 4.1] (which concerns the more difficult case where a random graph is present) shows that $\sup_{i=1, \dots, n} \mathbf{E} \left[\sup_{s \leq T} \left| \theta_s^{i,n} - \bar{\theta}_s^{\omega_i} \right| \right] \xrightarrow[n \rightarrow \infty]{} 0$, almost surely w.r.t. the disorder sequence, from which we deduce also the convergence of the empirical flow. Note that this result englobes also the case where Γ is linear. However, it is necessary for this last proof that one assumes $(\theta_0^{1,n}, \dots, \theta_0^{n,n})$ to be independent (although not necessarily identically distributed). Note that the coupling sequence $(\bar{\theta}^{\omega_1}, \dots, \bar{\theta}^{\omega_n})$ is made of independent variables, but not identically distributed (as each $\bar{\theta}^{\omega_i}$ obeys to its own intrinsic inhomogeneity ω_i). In comparison with the discussion in the homogeneous case in Chapter 1, one should rather speak here of propagation of independence [JPS21]: for fixed $k \geq 1$, if one denotes as $f^{n,(1, \dots, k)}$ as the law of the first k variables $(\theta^{1,n}, \dots, \theta^{k,n})$ (by lack of exchangeability, $f^{n,(1, \dots, k)}$ depends not only on k , but on the whole sequence of indices $(1, \dots, k)$), we have that, under the hypothesis that $f_0^{n,(1, \dots, k)}$ at time 0 is a product measure (this is what one could call as k -independent, instead of k -chaotic), then for $t > 0$, $f_t^{n,(1, \dots, k)}$ converges weakly as $n \rightarrow \infty$ to the product measure $\mu^{\omega_1} \otimes \dots \otimes \mu^{\omega_k}$ where μ^{ω} is the law of θ^{ω} solving (2.4). A crucial point here is that this notion of k -independence is no longer equivalent to the convergence of the empirical flow $\mu_n^{(\omega)}$, as it was the case for k -chaoticity.*

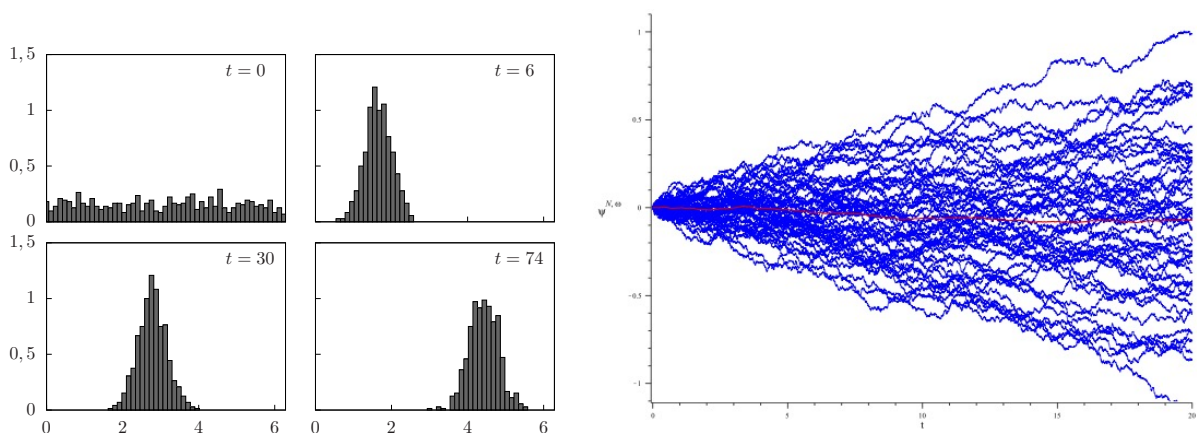
2.2 Quenched fluctuations

This part is taken from my PhD thesis, that is only briefly discussed here. We complement Proposition 2.1.3 with fluctuations results. The motivation comes from the Kuramoto model

(Example 1.2.2). A crucial aspect of the quenched convergence result of Proposition 2.1.3 is the *self-averaging* character of this limit: every typical disorder configuration $(\omega_i)_{i \geq 1}$ leads to the same deterministic evolution (2.11) as $n \rightarrow \infty$. However this may not be the case at the level of fluctuations: even if the distribution ν is symmetric (take the simple case where $\nu = \frac{1}{2}(\delta_{-1} + \delta_1)$ to fix ideas), the fluctuations of a fixed chosen sample of the disorder $(\omega_1, \dots, \omega_n)$ of spins in $\{\pm 1\}^n$ makes it *not symmetric*, resulting in a slow rotation of synchronized solutions in the direction of the majority, with a speed that depends randomly on the sample of the disorder (Fig. 2.1). This can be measured by computing the finite-size order parameters (i.e. the microscopic equivalents of (r_t, ψ_t) in (1.17)):

$$r_{n,t}^{(\omega)} e^{i\psi_{n,t}^{(\omega)}} = \frac{1}{n} \sum_{j=1}^n e^{i\theta_t^{j,n}} = \left\langle \mu_{n,t}^{(\omega)}, e^{i\theta} \right\rangle \quad (2.12)$$

where $\mu_n^{(\omega)}$ is defined in (2.2). In particular, $t \mapsto \psi_{n,t}^{(\omega)}$ which captures the position of the center of synchronization for finite n (see Figure 2.1a) has an approximately linear behavior whose slope depends on the choice of the disorder (Figure 2.1b).



(a) Evolution of the marginal on \mathbb{T} of $\mu_n^{(\omega)}$ for $n = 600$ oscillators ($\nu = \frac{1}{2}(\delta_{-1} + \delta_1)$, $K = 6$). The oscillators are initially chosen independently and uniformly on \mathbb{T} independently of the disorder. First the dynamics leads to synchronization of the oscillators ($t = 6$) to a profile which is close to a nontrivial stationary solution of (1.15). Secondly, the center $\psi_{n,t}^{(\omega)}$ of this density moves to the right with an approximately constant speed.

(b) Trajectories of the center of synchronization $\psi_{n,t}^{(\omega)}$ for different realisations of the disorder ($\mu = \frac{1}{2}(\delta_{-0.5} + \delta_{0.5})$, $K = 4$, $n = 400$). Direction and speed of the center of synchronization depend on the choice of the initial n -sample of the disorder. Moreover, these simulations are compatible with a speed of order $1/\sqrt{n}$. The red trajectory corresponds to the case where we average w.r.t. the disorder: non self-averaging does not hold in the averaged model.

Figure 2.1 – Non self-averaging for the quenched Kuramoto model (Example 1.2.2).

2.2.1 Quenched Central Limit Theorem for the empirical flow

The point of [L1] is to address this non self-averaging issue in presence of disorder at the level of fluctuations of the empirical flow. Consider

$$\eta_{n,t}^{(\omega)} := \sqrt{n} \left(\mu_{n,t}^{(\omega)} - \mu_t \right), \quad t \in [0, T]. \quad (2.13)$$

for a fixed realisation of the disorder $(\omega_i)_{i \geq 1}$. The fluctuation process $\eta_n^{(\omega)}$ has trajectories in the space of signed measures on $\mathcal{X} \times \mathcal{E}$. In a non-disordered context, similar fluctuation analysis goes back to [Szn85, Szn84, ST85], where it is proven that the fluctuation field $(\sqrt{n}(\langle \mu_n, f \rangle - \langle \mu, f \rangle), f \in \mathcal{F})$ (recall (1.2) and (1.4)) converges as $n \rightarrow \infty$ to some Gaussian field with prescribed covariance. Here, typically $\mathcal{F} = \{f \in L^2(P), \mathbf{E}_P(f) = 0\}$, where P is the law of the nonlinear process. The main strategy of proof is then twofold: first apply a Girsanov transform, which allows to get rid of the interaction in (1.1) so that the problem reduces to i.i.d. random processes and secondly apply classical asymptotic results for U -statistics [Bre83]. The covariance of the limiting process is then expressed in terms of integral operators appearing naturally within the limit of the

Radon-Nykodym derivative. In our context of disordered systems, the previous strategy presents several downsides: first, the above convergence is only stated in the sense of finite distribution, whereas one might expect a full convergence result of the fluctuation process. Secondly (and more importantly) the above techniques rely heavily on the exchangeability of (1.1), which is no longer true in our quenched context. And thirdly, the structure of covariance of the limiting process is expressed in terms on not so explicit integral operators, so that the analysis of the fluctuation process itself is not facilitated.

A strategy that is more amenable for our present situation concerns Hilbertian techniques developed in [FFG92] and used by Fernandez and Méléard [FM97] (see also [HM86, Mit83]) in the context of the fluctuation of the homogeneous system (1.1). The point is to start with the weak semimartingale decomposition of the fluctuation process $\eta_n = \sqrt{n}(\mu_n - \mu)$ written informally as

$$\langle \eta_{n,t}, f \rangle = \langle \eta_{n,0}, f \rangle + \int_0^t \langle \eta_{n,s}, \mathcal{L}_s^{\mu_n} f \rangle ds + M_{n,t}(f) \quad (2.14)$$

The difficulty here is to find a proper functional space in which one could pass to the limit $n \rightarrow \infty$ into the above formula resulting in a limit described in terms of a linear Stochastic PDE $\eta_t = \eta_0 + \int_0^t \mathcal{L}_s^{\mu,*} \eta_s ds + W_t$, where W is an appropriate Gaussian process. The point of [FM97] is to use weighted Sobolev spaces [Mét87] $W^{-j,\alpha}$, dual of test functions g such that $\left(\sum_{k \leq j} \int_{\mathbb{R}^d} \frac{|D^k g(x)|^2}{1+|x|^{2\alpha}} dx \right)^{\frac{1}{2}} < \infty$. The strategy follows from a tightness argument [Reb79, JM86] (based on compact Sobolev embeddings) and uniqueness of the limit (see e.g. [Oel87, JM98] for extensions of these techniques to moderately interacting diffusions).

In the present model, if now one fixes once and for all a realisation of the disorder (ω_i) , understanding $\langle \eta_n^{(\omega)}, f \rangle$ for all test functions f requires in particular to consider test functions of the form $f(\theta, \omega) = g(\omega)$, that is to look at $\sqrt{n}(\frac{1}{n} \sum_{i=1}^n g(\omega_i) - \int g(\omega) \nu(d\omega))$, quantity that *does not converge for a fixed realisation of the sequence* (ω_i) : there is no hope to prove that $\eta_n^{(\omega)}$ might be tight for a fixed realisation of (ω_i) . The strategy developed in [L1] to state a weak form of quenched convergence, that still captures the dependence in the disorder at the limit, is based on the following observation: even if, for fixed (ω) , the law $\mathcal{H}_n(\omega) \in \mathcal{P}(\mathcal{C}([0, T], \mathcal{S}'))$ of the process $(\eta_n^{(\omega)})$ is not tight, *the random variable* $(\omega) \mapsto \mathcal{H}_n^{(\omega)}$ is tight (we work here \mathcal{S}' is the Schwartz space of tempered distributions on $\mathcal{X} \times \mathcal{E}$, that can endowed with a structure of nuclear Fréchet space). The main result of [L1] is

Theorem 2.2.1 (Theorem 2.10 in [L1]). *Suppose that $\mathcal{X} = \mathbb{T}$, that (ω_i) are i.i.d. random variables with $\int |\omega|^k \nu(d\omega) < \infty$ for a sufficiently large k and that F and Γ are regular in θ, θ' with appropriate moment condition in ω, ω' . Then $(\omega) \mapsto \mathcal{H}_n(\omega)$ converges in law to a random variable $\omega \mapsto \mathcal{H}(\omega)$, where for fixed ω , $\mathcal{H}(\omega)$ is the law of the unique solution to the Ornstein-Uhlenbeck process η^ω in \mathcal{S}' :*

$$\eta_t^\omega = X(\omega) + \int_0^t \mathcal{L}_s^* \eta_s^\omega ds + W_t \quad (2.15)$$

where $X(\omega)$ is a Gaussian process with mean value $C(\omega)$, independent from W . As a random variable in ω , $\omega \mapsto C(\omega)$ is a Gaussian process.

The only remaining dependence in the disorder at the limit is in the initial condition $X(\omega)$ which has a nontrivial mean value for fixed ω : this is the signature of the non-selfaveraging phenomenon mentioned above, at the level of the CLT. Let us give briefly some lines of proof for Theorem 2.2.1. The difficulty part is the tightness of the law Θ_n of $(\omega) \mapsto \mathcal{H}_n^{(\omega)}$ (that is an element of $\mathcal{P}(\mathcal{P}(\mathcal{C}_{\mathcal{S}'}))$, where $\mathcal{C}_{\mathcal{S}'} := \mathcal{C}([0, T], \mathcal{S}')$). The first ingredient is Mitoma's criterion: a sequence (P_n) of probability measures on $\mathcal{C}_{\mathcal{S}'}$ is tight if for any φ in a suitable countable dense subset of \mathcal{S} , $(P_n \Pi_\varphi^{-1})$ is tight in $\mathcal{C} := \mathcal{C}([0, T], \mathbb{R})$, for $\Pi_\varphi : \psi(\cdot) \in \mathcal{C}_{\mathcal{S}'} \mapsto \langle \psi(\cdot), \varphi \rangle \in \mathcal{C}$. The second is Aldous criterion for tightness in \mathcal{C} : a sequence (Y^n) of processes with paths in \mathcal{C} is tight if the two conditions hold: (i) *Condition [T]*: for any values t in a dense set of $[0, T]$ and $q \geq 1$, there exists $C > 0$ such that $\sup_n \mathbf{P}(|Y_t^n| > C) \leq \frac{1}{q}$, (T, q, C) and (ii) *Condition [A]*: for all $p_1, p_2 \geq 1$, there exist $C > 0$, $n_0 \geq 1$ such that for all stopping time τ , $\sup_{n \geq n_0} \sup_{\theta \leq C} \mathbf{P}\left(|Y_\tau^n - Y_{\tau+\theta}^n| \geq \frac{1}{p_2}\right) \leq \frac{1}{p_1}$, $(A_{p_1, p_2, C})$.

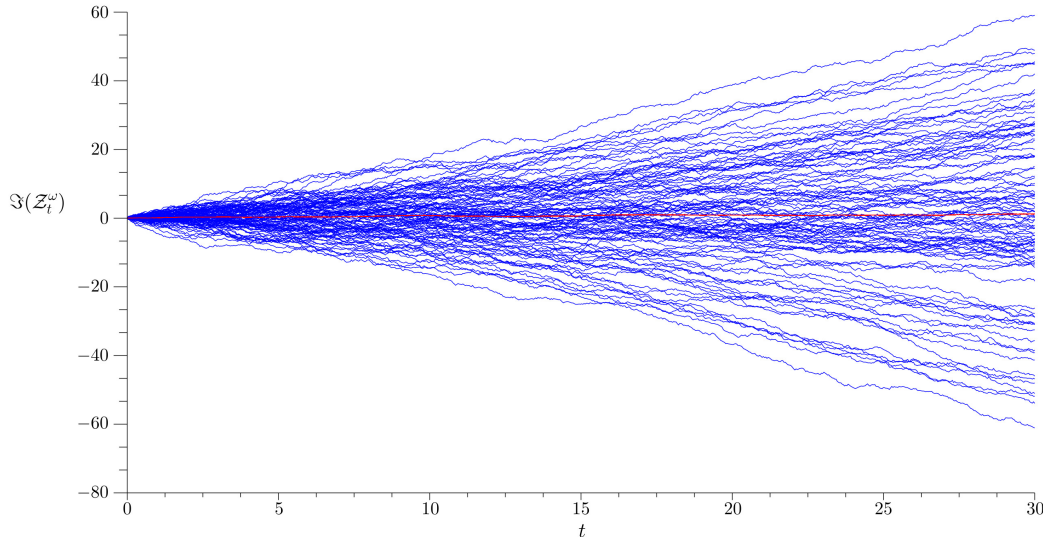


Figure 2.2 – Evolution of the process $\langle \eta_t^\omega, \sin \rangle$, for different realisations of ω ; the trajectories are sample-dependent and compatible with the ones observed in Fig. 2.1.

Then for a countable dense family $(\varphi_j)_{j \geq 1}$ in the nuclear space \mathcal{S} , define for all $\varepsilon > 0$, the subsets of $\mathcal{P}(\mathcal{C}_{\mathcal{S}'})$: $K_1^\varepsilon(\varphi_1, \dots, \varphi_J) := \left\{ P, \forall t, q, 1 \leq j \leq J, P\Pi_{\varphi_j}^{-1} \text{ satisfies } (T_{t,q,C_1}) \right\}$, $K_2^\varepsilon(\varphi_1, \dots, \varphi_J) := \left\{ P, \forall 1 \leq j \leq J, \forall p_1, p_2 > 0, P\Pi_{\varphi_j}^{-1} \text{ satisfies } (A_{p_1,p_2,C_2}) \right\}$ and finally $K^\varepsilon := \bigcap_{j \geq 1} K_1^\varepsilon(\varphi_1, \dots, \varphi_J) \cap K_2^\varepsilon(\varphi_1, \dots, \varphi_J)$. By construction K^ε is relatively compact in $\mathcal{P}(\mathcal{C}_{\mathcal{S}'})$ and the point is to prove that $\limsup_n \Theta_n(K^{\varepsilon,c}) \leq \varepsilon$ for a careful choice of values of $C_1 = C_1(q, \varepsilon)$ and $C_2 = C_2(\varepsilon, p_1, p_2)$ to conclude. The key to this last estimate is based on the crucial observation that the fluctuation process satisfies, for an adequate Sobolev norm with index r , $\mathbf{E} \left[\sup_{t \leq T} \|\eta_t^n\|_{-r}^2 \right] \leq C_n(\omega_1, \dots, \omega_n)$ for a quantity C_n depending only on the disorder, that is *not bounded* almost surely in (ω) but is such that $\lim_{A \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}(C_n > A) = 0$. This last bound is proven via a careful use of the Hilbertian techniques of [FM97].

2.2.2 Large time analysis of the fluctuation process

We briefly mention another result of my PhD thesis [L4], which concerns the analysis of the fluctuation SPDE (2.15) as $t \rightarrow \infty$, in the case of the Kuramoto model. One observes numerically that $\langle \eta_t^\omega, \sin \rangle$ where η_t^ω solves (2.15) shows non-self-averaging behavior as $t \rightarrow \infty$ similar to the ones observed for the microscopic center of synchronisation $\psi_{n,t}^{(\omega)}$ in Figure 2.1, see Figure 2.2. The main result of [L4] is

Theorem 2.2.2 (Theorem 2.10 in [L4]). *Suppose that $\nu = \frac{1}{2}(\delta_{-\omega_0} + \delta_{\omega_0})$ for some $\omega_0 > 0$. There exists a Sobolev space \mathbf{H} of distributions, such that for $K > 1$, if $\omega_0 > 0$ is sufficiently small, there exists a unique solution η to (2.15) in \mathbf{H} and for fixed initial condition η^ω , there exists $v(\omega) \in \mathbb{R}$ (explicit in terms of η_0^ω) such that*

$$\frac{\eta_t^\omega}{t} \xrightarrow[t \rightarrow \infty]{} v(\omega) \partial_\theta q \quad (2.16)$$

Moreover, $\omega \mapsto v(\omega)$ is a Gaussian random variable with explicit variance $\sigma^2 > 0$.

Here, $q = q_{r,0}$ is the synchronised stationary profile given in (1.22) and the linear operator \mathcal{L}_s^* driving (2.15) corresponds to the linearised operator around the solution $p_t(\theta, \omega)$ to (1.15):

$$\mathcal{L}_t^* h(\theta, \omega) = L_{p_t}^\nu h(\theta, \omega) = \frac{\sigma^2}{2} \partial_\theta^2 h(\theta, \omega) - \partial_\theta \left(h(\theta, \omega) \left(\int J * p_t d\nu + \omega \right) + p_t(\theta, \omega) \int (J * h) d\nu \right), \quad (2.17)$$

with domain $\mathcal{D} := \{h(\theta, \omega) \in \mathcal{C}^2(\mathbb{T}), \int_{\mathbb{T} \times \mathbb{R}} h(\theta, \omega) d\theta \nu(d\omega) = 0\}$. The space \mathbf{H} in Theorem 2.2.2 is essentially an extension of the weighted Sobolev space given in (1.25), but adapted to the context of test functions with disorder [L4, § 2.3]: we define as $\mathbf{H} = \mathbf{H}_{\nu, q}^{-1}$ the weighted spaces of distributions, defined as the closure of \mathcal{D} w.r.t. the norm

$$\|h\|_{-1, \nu, q} := \left(\int_{\mathbb{R}} \left(\int_{\mathbb{T}} h(\theta, \omega) d\theta \right)^2 \nu(d\omega) + \int_{\mathbb{R}} \int_{\mathbb{T}} \frac{\mathcal{H}_0^2(\theta, \omega)}{q(\theta, \omega)} d\theta \nu(d\omega) \right)^{\frac{1}{2}} \quad (2.18)$$

where for fixed ω , $\theta \mapsto \mathcal{H}_0(\theta, \omega)$ is the primitive of $h_0 := \theta \mapsto h(\theta, \omega) - (\int_{\mathbb{T}} h(u, \omega) du) q(\theta, \omega)$ (note that $\int h_0(\theta, \omega) d\theta = 0$) such that $\int_{\mathbb{T}} \frac{\mathcal{H}_0(\theta, \omega)}{q(\theta, \omega)} d\theta = 0$. The crucial point of Theorem 2.2.2 concerns the spectral properties of the operator (L_q^ν, \mathcal{D}) : it is proven in [L4, Th. 2.8] that (L_q^ν, \mathcal{D}) is densely defined, closable in \mathbf{H} , with compact resolvent and generates an analytic semigroup with spectrum in $\{z \in \mathbb{C}, \Re(z) \leq 0\}$. The dimension of the characteristic space of the eigenvalue 0 is exactly 2, spanned by $\partial_\theta q$ and p such that $L_q^\nu \partial_\theta q = 0$ and $L_q^\nu p = \partial_\theta q$. In words, this last identity states that there is a Jordan block of size 2 in 0, which is the crucial observation for the linear behavior (2.16) of η^ω as $t \rightarrow \infty$. Whereas the identity $L_q^\nu \partial_\theta q = 0$ is trivial (it is the invariance by rotation of the problem), the difficult part is the existence of some p such that $L_q^\nu p = \partial_\theta q$. This is based on coercivity estimates on the Dirichlet form associated to L_q^ν , $\mathcal{E}(h, l) := \langle L_q^\nu h, l \rangle_{-1, \nu, q}$. Here we take advantage of the symmetry of the system: one has that $\partial_\theta q \in \mathcal{O}$ and $L_q^\nu(\mathcal{O}) \subset \mathcal{O}$ so we work in $\mathcal{O} := \{h; \forall(\theta, \omega) \in \mathbb{T} \times \text{Supp}(\nu), h(-\theta, -\omega) = -h(\theta, \omega)\}$. Denote by λ the Lebesgue measure on \mathbb{T} . Then, integration by parts formulas and considerations of symmetry give that $\mathcal{E}(h, l) = \Gamma(h, l) + K \ell(h) \ell(l)$, for all $(h, l) \in \mathcal{D}(\mathcal{E}) := (L_\nu^2 \cap \mathcal{O}) \times (\mathbf{H}_{\nu, q}^{-1} \cap \mathcal{O})$, where $\ell(l) := \int_{\mathbb{T} \times \mathbb{R}} l \sin(\cdot) d\lambda d\mu$ and $\Gamma(h, l) := -\frac{1}{2} \int_{\mathbb{T} \times \mathbb{R}} \frac{hl}{q} d\lambda d\mu + \int_{\mathbb{T} \times \mathbb{R}} \kappa(\cdot) \frac{h\mathcal{L}}{q^2} d\lambda d\mu$, where $\kappa(\omega) := \frac{1-e^{4\pi\omega}}{2Z(\omega)}$ and \mathcal{L} is the primitive of l such that $\int_{\mathbb{T}} \frac{\mathcal{L}}{q} d\lambda = 0$. It is easy to see by direct calculations that there is some p_2 such that $\ell(l) = \langle L_q^\nu p_2, l \rangle_{-1, \nu, q}$ so that it remains to show the existence of some p_1 such that $\Gamma(p_1, l) = \langle \partial_\theta q, l \rangle_{-1, \nu, q}$. This is done by coercivity estimates: for the Hilbert space $\mathcal{H} := L_\nu^2 \cap \mathcal{O}$ and the normed space $\mathcal{G} := \mathbf{H}_{\nu, q}^{-1} \cap \mathcal{O}$, Γ is bilinear on $\mathcal{H} \times \mathcal{G}$ and $\Gamma(\cdot, \varphi)$ is continuous for each $\varphi \in \mathcal{G}$. Moreover choosing h of the form $h = qf\mathcal{L}$ with f some carefully chosen explicit function, we obtain $\Gamma(h, l) = \int_{\mathbb{T} \times \mathbb{R}} \left\{ \frac{1}{4} \partial_\theta f + \kappa(\cdot) \frac{f}{q} \right\} \mathcal{L}^2$. Choosing precisely f such that the term within brackets is equal to $\frac{1}{q}$, one obtains that $\Gamma(h, l) = \|l\|_{-1, \nu, q}^2$. We conclude from this that $\inf_{\|\varphi\|_{-1, \nu, q}=1} \sup_{\|h\|_{2, \nu} \leq 1} |\Gamma(h, \varphi)| \leq C$ for some universal constant $C > 0$. This is the desired coercivity estimate, which implies the result by the extension of the Lax-Milgram theorem in [Sho97, Chap. 3].

The rest of the proof follows from a perturbation argument (see [Kat95]) from the non-disordered case (Theorem 1.2.4), noting that the disordered operator L_q^ν is nothing else (up to technical details) than a relatively bounded perturbation of L_q^0 defined in (1.24). However, note that the combination of Theorem 2.2.1 and Theorem 2.2.2 does not give a rigorous proof of the non-self-averaging phenomenon observed in Figure 2.1, as one takes first the limit in $n \rightarrow \infty$ (for fixed t) in Theorem 2.2.1 and then $t \rightarrow \infty$ in Theorem 2.2.2, whereas the analysis of Section 2.2 would require to take a joint limit of (2.2) when both $(n, t) \rightarrow \infty$. We will address precisely this question later in Chapter 5.

2.3 Quenched Large Deviation Principles for empirical measure and flow

So far, we have addressed the case of the quenched Law of Large Numbers concerning the empirical flow (Proposition 2.1.3), as well as the quenched fluctuations around its limit (Theorem 2.2.1) [1]. A way to complement these results is to look at Large Deviations issues (the very motivation of [DdH96] was precisely to establish an *annealed* Large Deviations Principle for the flow (2.2) as $n \rightarrow \infty$). The picture would not be complete without some *quenched* Large

[1]. The annealed counterparts of these results are obviously true, definitely easier to prove and anyway ineffective in view of the quenched traveling waves experimented in Figure 2.1.

Deviations Principle: this is the purpose of the work [L8] that we address below. Note here that, as in [DdH96], we are not only interested in the empirical flow (2.2) but also in the more general [2] *double-layer empirical measure*

$$L_n^{(\omega)} := \frac{1}{n} \sum_{j=1}^n \delta_{(\theta^{j,n}, \omega_j)} \in \mathcal{P}(\mathcal{C}([0, T], \mathcal{X} \times \mathcal{E})). \quad (2.19)$$

We do not aim to enter into too much formalism on large deviations here, we refer to the classical references on the subject [DZ98, DS89]. Let us just recall informally that a sequence of probability measures (P_n) on a topological space \mathcal{X} satisfies a (strong) Large Deviation Principle (LDP) with speed n and good rate function I if the level sets $\{I \leq \alpha\}$ are compact and that the following holds: for all closed sets F (resp. open sets O) in \mathcal{X} ,

$$-\inf_{x \in O} I(x) \leq \liminf_{n \rightarrow \infty} \frac{1}{n} \ln P_n(O) \text{ and } \limsup_{n \rightarrow \infty} \frac{1}{n} \ln P_n(F) \leq -\inf_{x \in F} I(x). \quad (2.20)$$

A seminal work concerning LDP in the homogeneous case is the article of Dawson and Gärtner [DG87], where it is proven, using projective limits, that the rate function governing the LDP for the empirical measure (1.2) as $n \rightarrow \infty$ is expressed in terms of the *action functional* $S(\mu) := \int_0^T \|\partial_t \mu(t) - \mathcal{L}(\mu(t))^* \mu(t)\|_{\mu(t)}^2 dt$, where $\|\vartheta\|_{\mu}^2 = \frac{1}{2} \sup_{\phi} \frac{|\langle \vartheta, \phi \rangle|^2}{\langle \mu, |\nabla \phi|^2 \rangle}$ and \mathcal{L} is the generator of (1.4). This formalism has been further developed for a class of Curie-Weiss model in [DG89], where the quasi-potential associated to S has been related to the free energy functional associated with the equilibrium large deviations of the system. The techniques introduced in [DG87] have been successfully applied to extensions of mean-field models (see e.g. [Mül17] in the case with spatial dependence). Note that another approach using weak convergence methods have been developed in [BDF12].

2.3.1 Some hypotheses

We consider (2.1) in the case $\Gamma(\theta, \theta', \omega, \omega') = \partial_{\theta} f(\theta - \theta', \omega, \omega')$, for some symmetric and regular f : $f(\theta, \omega, \omega') = f(-\theta, \omega', \omega)$. Note that it is assumed in the original work of [DdH96] that the same gradient structure holds for $F(\theta, \omega) = -\partial_{\theta} g(\theta, \omega)$ and also that the initial conditions $(\theta_0^{1,n}, \dots, \theta_0^{n,n})$ are i.i.d with law γ and that $(\omega_i)_{i \geq 1}$ is a sequence of i.i.d. random variables with law ν : these last assumptions are actually not necessary and will be discarded in our analysis for the quenched model.

2.3.2 The annealed case

First recall the main arguments of [DdH96] in the annealed case: denote by $P_n^{(\omega)} \in \mathcal{P}(\mathcal{C}([0, T], \mathcal{X})^n)$ the law of $(\theta^{1,n}, \dots, \theta^{n,n})$ solving (2.1) in case both F and Γ are present. Denote in a same way by $W_{n,0}$ the law of $(\theta^{1,n}, \dots, \theta^{n,n})$ when there is no interaction and intrinsic dynamics ($F \equiv 0$ and $\Gamma \equiv 0$). The main line of arguments in [DdH96] is simple: when $F \equiv 0$ and $\Gamma \equiv 0$, in the case of i.i.d. initial conditions, *under the annealed law* $\mathbf{W}_n(\cdot) := \int W_{n,0}(L_n^{(\omega)} \in \cdot) \nu^{\otimes n}(d(\omega))$, $(\theta^{i,n}, \omega_i)_{i=1, \dots, n}$ is a simple sequence of i.i.d. random variables, so that, by Sanov's theorem, the empirical measure (2.19) satisfies a LDP governed by the entropy $\lambda \mapsto H(\lambda|W \otimes \nu)$, where W is the law of a standard Brownian motion. Next, by Girsanov transform and an application of Ito's formula (which uses the gradient structure of F and Γ), one sees that the Radon-Nykodym derivative $\frac{dP_n^{(\omega)}}{dW_{n,0}} = \exp(nR(L_n^{(\omega)}))$ can be expressed as a bounded and continuous (w.r.t. the weak topology) functional of $L_n^{(\omega)}$. One deduces immediately from Varadhan's Lemma the full LDP for the empirical measure (2.19) with rate function $\lambda \mapsto H(\lambda|W \otimes \nu) - R(\lambda)$. Additional calculations lead to the following final expression of the annealed rate function (see [DdH96,

[2]. The knowledge of (2.19) gives more information than (2.2), as one obtains $\mu_{n,t}^{(\omega)}$ from $L_n^{(\omega)}$ via the canonical projection $\pi_t : \mathcal{C}([0, T], \mathcal{X}) \times \mathcal{E} \rightarrow \mathcal{X} \times \mathcal{E}$, $(x, \omega) \mapsto (x_t, \omega)$, i.e. $\mu_{n,t}^{(\omega)} = \pi_t L_n^{(\omega)}$.

Cor. 1])

$$\mathcal{G}^{ann} := \lambda \mapsto \begin{cases} \int H(\lambda^\omega | P^{\lambda, \omega}) \lambda_2(d\omega) + H(\lambda_2 | \nu), & \text{if } \lambda_2 \ll \nu, \\ +\infty & \text{otherwise} \end{cases} \quad (2.21)$$

where $\lambda \in \mathcal{P}(\mathcal{C}([0, T], \mathcal{X}) \times \mathcal{E})$ disintegrates into $\lambda(d\theta, d\omega) = \lambda^\omega(d\theta) \lambda_2(d\omega)$ and where $P^{\lambda, \omega} \in \mathcal{P}(\mathcal{C}([0, T], \mathcal{X}))$ is the law of the unique strong solution to $d\theta_t = F(\theta_t, \omega)dt - \int \partial_\theta f(\theta_t - \theta', \omega, \omega') \lambda_t(d\theta', d\omega') dt + dB_t$, $\theta_0 \sim \gamma$.

2.3.3 Quenched LDP for the empirical measure

Now turn to the LDP for (2.19) in our quenched setting. The result is somehow technical and we do not give the full hypotheses here, we only highlight the relevant main points in comparison with the hypotheses in [DdH96]. A first important remark concerns F : we no longer assume that F is bounded, only that it satisfies appropriate polynomial bound $\sup_{\theta \in \mathcal{X}} |F(\theta, \omega)| \leq C(1 + |\omega|^k)$ as well as Lipschitz continuity w.r.t. both variables (θ, ω) . Secondly, we no longer suppose that F derives from a potential, as in [DdH96]. Thirdly, the i.i.d. assumption concerning both (ω_i) and the initial condition is discarded: we fix once and for all a (deterministic) sequence $(\omega_i)_{i \geq 1}$ satisfying (2.3) (together with some further asymptotics on moments). We suppose the initial conditions $(\theta_0^{1,n}, \dots, \theta_0^{n,n})$ to be independent (but not necessarily i.i.d.) with $\theta_0^{i,n} \sim \gamma^{\omega_i}$ with appropriate (Feller) regularity hypotheses and moment conditions on the family $\omega \mapsto \gamma^\omega$.

The strategy of [L8] in the quenched set-up follows similar lines as for [DdH96], with several significant bifurcations that we make explicit here. The fact that F no longer derives from a potential is not a real issue: the point is simply to modify slightly the Girsanov argument of [DdH96] by computing $\frac{dP_n^{(\omega)}}{dW_{n,F}^{(\omega)}}$ (to compare with $\frac{dP_n^{(\omega)}}{dW_{n,0}^{(\omega)}}$ in [DdH96]), where

$$W_{n,F}^{(\omega)} = W_F^{\omega_1} \otimes \dots \otimes W_F^{\omega_n} \in \mathcal{P}(\mathcal{C}([0, T], \mathcal{X})^n) \quad (2.22)$$

is the law of (2.1) when there is no interaction ($\Gamma \equiv 0$), but F is present. Note here that since $\Gamma = 0$, (2.1) consists of independent but non identically distributed diffusions: (2.22) is a product measure (here W_F^ω is the law of the diffusion $d\theta_t = F(\theta_t, \omega)dt + \sigma dB_t$). Doing so, one easily sees that the gradient structure on F is irrelevant to the calculations. This technical point apart, the major difficulty in the quenched setting is that the first argument of [DdH96] no longer applies: under the quenched law $W_{n,F}^{(\omega)}$, the particles $(\theta^{i,n})_{i, \dots, n}$ are independent, but no longer identically distributed: the usual Sanov's theorem does not apply. The first argument of the proof is then

Proposition 2.3.1 (Proposition 2.1 in [L8]). *Under the present hypotheses, the quenched law $\mathbf{W}_n^{(\omega)}(\cdot) := W_n^{(\omega)}(L_n^{(\omega)} \in \cdot)$ (that is the law of the empirical measure (2.19) with no interaction, $\Gamma \equiv 0$) satisfies a LDP in $\mathcal{P}(\mathcal{C}([0, T], \mathcal{X}) \times \mathcal{E})$ with good rate function*

$$\mathcal{I} := \lambda \mapsto \begin{cases} H(\lambda | W^\omega(d\theta) \nu(d\omega)), & \text{if } \lambda_2 = \nu, \\ +\infty & \text{otherwise} \end{cases} \quad (2.23)$$

where $W^\omega(d\theta)$ is the law of $d\theta_t = F(\theta_t, \omega)dt + \sigma dB_t$, $\theta_0 \sim \gamma^\omega$.

Proposition 2.3.1 is a quenched version of Sanov's Theorem (see e.g. [Léo07, CL95] and [DZ98, Th. 6.2.10] for variants of this result). Proof of Proposition 2.3.1 uses crucially large deviation techniques for projective limits introduced by Dawson and Gärtner [DG87, Th. 3.4]: the first point is to note that for any bounded continuous ϕ , the log-Laplace transform $\Lambda_n(\phi) := \frac{1}{n} \ln \int \exp(n \langle \lambda, \phi \rangle) \mathbf{W}_n^{(\omega)}(d\lambda)$ has a limit $\Lambda(\phi)$ as $n \rightarrow \infty$. Indeed, noting that the assumed Feller property of the initial condition $\omega \mapsto \gamma^\omega$ propagates to the whole law W^ω (see [L8, Lem. 2.4]), one can simply apply the convergence of the empirical measure of the disorder (2.3) to $\Lambda_n(\phi) = \frac{1}{n} \sum_{i=1}^n \ln \int_{\mathcal{C}([0, T], \mathcal{X})} \exp(\phi(x, \omega_i)) W^{\omega_i}(dx)$ to obtain that $\Lambda_n(\phi) \xrightarrow[n \rightarrow \infty]{} \Lambda(\phi) := \int_{\mathcal{E}} \ln \int_{\mathcal{C}([0, T], \mathcal{X})} \exp(\phi(x, u)) W^u(dx) \nu(du)$. It remains to show that Λ is Gâteaux differentiable

and that $\Lambda^*(\lambda) < \infty$ implies that $\lambda_2 = \nu$, where $\Lambda^*(\lambda) := \sup_{\phi} \{\langle \lambda, \phi \rangle - \Lambda(\phi)\}$ is the Fenchel-Legendre transform of Λ (see [L8, Th. 6.1] for further details). Then Proposition 2.3.1 is then a direct consequence of the abstract result of Dawson and Gärtner [DG87, Th. 3.4].

Following the same lines as for [DdH96], the next point is to compute the Radon-Nykodym derivative $\frac{dP_n^{(\omega)}}{dW_{n,F}^{(\omega)}} = \exp(nR(L_n^{(\omega)}))$ (where the functional R is explicit, see [L8, Prop. 3.1]). A second crucial difficulty here is that under our present hypotheses (recall in particular that F is no longer bounded), the function R is no longer continuous nor bounded for the weak topology so that Varadhan's Lemma does not apply directly. We proceed by a truncation argument of the disorder by replacing ω_i by $\chi_M(\omega_i) = (\omega_i \wedge M) \vee (-M)$ for some $M > 0$, apply Varadhan's Lemma for fixed $M > 0$ and proceed with $M \rightarrow \infty$, using exponential approximation techniques. This only gives a weak LDP (i.e. the upper-bound is only valid on compact sets in (2.20)) and the rest of the proof consists of properly identifying the rate function and prove that the LDP is strong. We refer to [L8] for further details on this point. All of this gives

Theorem 2.3.2 (Theorem 1.7 in [L8]). *Under the present hypotheses, the sequence $\mathbf{P}_n^{(\omega)}(\cdot) := P_n^{(\omega)}(L_n^{(\omega)} \in \cdot)$ satisfies a strong LDP in $\mathcal{P}(\mathcal{C}([0, T], \mathcal{X}) \times \mathcal{E})$, with speed n , governed by the good rate function*

$$\mathcal{G}^{quen} := \lambda \mapsto \begin{cases} \int H(\lambda^\omega | P^{\lambda, \omega}) \lambda_2(d\omega), & \text{if } \lambda_2 = \nu, \\ +\infty & \text{otherwise} \end{cases} \quad (2.24)$$

Note that $0 \leq \mathcal{G}^{ann}(\cdot) \leq \mathcal{G}^{quen}(\cdot)$ and that both rate functions have a unique zero λ^* that is the law of $(\bar{\theta}^\omega, \omega)$ where $\bar{\theta}^\omega$ is the nonlinear process (2.4). An easy corollary of Theorem 2.3.2 is the quenched convergence of the empirical measure (2.19) to λ^* , by an immediate Borel-Cantelli argument. Byproducts of Theorem 2.3.2 and the contraction principle are also LDP concerning the empirical measure on the particles only $\frac{1}{n} \sum_{i=1}^n \delta_{\theta^{i,n}}$ and local empirical measures conditioned on the value of the disorder. We refer to [L8, Prop. 1.15 and 1.19] for more details.

2.3.4 Quenched LDP for the empirical flow

By Theorem 2.3.2 and the contraction principle, the empirical flow (2.2) satisfies a strong LDP for the rate function

$$\mathcal{G}_1 : q \in \mathcal{C}([0, T], \mathcal{P}(\mathcal{X} \times \mathcal{E})) \mapsto \inf \{ \mathcal{G}^{quen}(\lambda), \pi \lambda = q \} \quad (2.25)$$

where π is the canonical projection. The main objective of [L8, Th. 1.12] is to identify the rate function \mathcal{G}_1 with

$$\mathcal{G} := q \mapsto \begin{cases} \int \mathcal{K}(q, \omega) \nu(d\omega), & \text{if } q \in \mathbb{A}, \\ +\infty & \text{otherwise} \end{cases} \quad (2.26)$$

where

$$\mathcal{K}(q, \omega) := \sup_{\phi \in \mathcal{C}_0^\infty([0, T] \times \mathcal{X})} \left\{ \int_0^T \langle \phi(t, \cdot), \partial_t q_t^\omega - \mathcal{L}^\omega q_t^\omega \rangle dt - \frac{1}{2} \int_0^T \langle (\partial_\theta \phi(t, \cdot))^2, q_t^\omega \rangle dt \right\} \quad (2.27)$$

and \mathbb{A} is the set of all admissible flows such that $t \mapsto q_t^\omega$ is weakly differentiable for ν -almost every ω and $q_2 = \nu$. Equality $\mathcal{G}_1(q) = \mathcal{G}(q)$ is standard (see [DdH96]), provided that one can make sure first that both quantities are finite: when $\mathcal{G}_1(q) < \infty$, since \mathcal{G}^{quen} is a good rate function, there exists some λ such that $\pi \lambda = q$ and $\mathcal{G}(\lambda) = \mathcal{G}_1(q) < \infty$. In particular $\lambda(d\theta, d\omega) = \lambda^\omega(d\theta) \nu(d\omega)$ and $\mathcal{G}^{quen}(\lambda) = \int \mathcal{H}(\lambda^\omega | W^\omega) \nu(d\omega) - \mathcal{J}(\lambda)$, where \mathcal{J} is the functional arising from the Girsanov transform, see [L8, eq. (3.3)]. $\mathcal{J}(\lambda)$ only depends on q . Therefore, for ν a.e. ω , λ^ω minimises $\mathcal{H}(\lambda^\omega | W^\omega)$ under the constraint $\pi \lambda^\omega = q^\omega$. This implies [Föl88] that λ^ω is the law of a diffusion $d\theta_t^\omega = b_t^\omega(\theta_t^\omega) dt + dw_t$. Plugging this expression into the expression of \mathcal{G}^{quen} and using the alternative form $\mathcal{K}(q, \omega) = \frac{1}{2} \sup_{\phi \in \mathcal{C}_0^\infty([0, T] \times \mathcal{X})} \frac{(\int_0^T \langle \phi(t, \cdot), \partial_t q_t^\omega - \mathcal{L}^\omega q_t^\omega \rangle dt)^2}{\int_0^T \langle (\partial_\theta \phi(t, \cdot))^2, q_t^\omega \rangle dt}$ gives the desired equality $\mathcal{G}_1(q) = \mathcal{G}(q)$. The main difficulty is to prove that $\mathcal{G}(q) < \infty$ implies that $\mathcal{G}_1(q) < \infty$ (this point is the object of a circular reasoning in [DdH96]). The proof of statement is carried out in [L8,

Prop. 5.2]) (note that another alternative proof could be to use projective limits as in [DG87, Mül17]) and is briefly discussed here: suppose that $\mathcal{G}(q) < \infty$, then for ν -a.e. ω , $\mathcal{K}(q, \omega) < \infty$ and one can alternatively write $\mathcal{K}(q, \omega) = \sup_{\phi \in \mathcal{C}_0^\infty([0, T] \times \mathcal{X})} \left\{ \mathcal{U}(\phi) - \frac{1}{2} \int_0^T \int (\partial_\theta \phi(\theta, t))^2 q_t^\omega(d\theta) dt \right\}$, with $\mathcal{U}(\phi) := \int \phi(\theta, t) q_T^\omega(d\theta) - \int \phi(\theta, 0) q_0^\omega(dx) - \int_0^T \int (\partial_t + (\mathcal{L}^\omega)^*) \phi(\theta, t) q_t^\omega(d\theta) dt$. One obtains then that for all $\gamma > 0$, $\left| \frac{\mathcal{U}(\phi)}{\gamma} \right| \leq \frac{1}{\gamma^2} \frac{1}{2} \int_0^T \langle (\partial_\theta \phi(t, \cdot))^2, q_t^\omega \rangle dt + \mathcal{K}(q, \omega)$ and choosing $\gamma = \|\partial_\theta \phi\|_{2, q^\omega}$, we have $|\mathcal{U}(\phi)| \leq \left(\frac{1}{2} + \mathcal{K}(q, \omega) \right) \|\partial_\theta \phi\|_{2, q^\omega}$. Hence \mathcal{U} can be extended to a continuous linear form on $H_0^{-1}(q^\omega)$, completion of smooth function under $\|\partial_\theta \phi\|_{2, q^\omega}$. By Riesz Theorem, there exists some $B^\omega \in H_0^{-1}(q^\omega)$ such that $\mathcal{U}(\phi) = \int_0^T \int (\partial_\theta \phi B^\omega) dq_t^\omega dt$. Plugging this expression into (2.27), one obtains finally $\mathcal{K}(q, \omega) = \sup_{\phi \in \mathcal{C}_0^\infty([0, T] \times \mathcal{X})} \left\{ \langle B^\omega, \partial_\theta \phi \rangle_{2, q^\omega} - \frac{1}{2} \|\partial_\theta \phi\|_{2, q^\omega}^2 \right\} = \frac{1}{2} \|B^\omega\|_{2, q^\omega}^2$. We then borrow from the deep result of [CL94, Th. 5.9]: there exists λ_*^ω such that $\pi \lambda_*^\omega = q^\omega$ and $\mathcal{H}(\lambda_*^\omega | P^{q, \omega}) = \frac{1}{2} \|B^\omega\|_{2, q^\omega}^2$ (this λ_*^ω is the law of the diffusion with generator $\frac{1}{2} \partial_\theta^2 + B^\omega \partial_\theta$, so-called *Nelson processes*, see [CL94, CL95] for further references). We immediately derive from the last equality and (2.25) and (2.26) that $\mathcal{G}_1(q) < \infty$ and we are done.

2.3.5 Perspectives

The fact that the interaction kernel Γ derives from a potential is mostly a convenient technical tool allowing to apply Ito's formula in the calculation of the Radon-Nykodym derivative, but this does not appear in the formal definition of the rate function. In the recent work [BPR22], the authors prove an annealed LDP on the empirical measure (2.19) without supposing that the interaction kernel is gradient (see also [DFMS18] for an alternative approach using rough paths methods). An extension to the quenched setting seems a natural perspective, although not straightforward.

2.4 The case of singular spatial interactions

The point of this section is to comment on the works [L3, L6] where the influence of the possible singularity in the interaction is analyzed. Here, the point of view is to consider the disorder as a spatial position $\omega_i = x_i$, so that we address the spatially-extended version of (2.1)

$$d\theta_t^{i,n} = F(\theta_t^{i,n}) dt + \frac{1}{n} \sum_{j=1}^n \Gamma(\theta_t^{i,n}, \theta_t^{j,n}) \Psi(x_i, x_j) dt + \sigma \cdot dB_t^i, \quad i = 1, \dots, n, \quad t \in [0, T], \quad (2.28)$$

where each position x_i lives in some space I (typically $I = [0, 1]^d$ or \mathbb{R}^d) and Ψ is a fixed spatial kernel on I^2 . Spatially-extended systems as (2.28) are particularly relevant in various contexts, especially in neuroscience as there is biological evidence of spatial organization of neurons in the brain (e.g. cortical columns in the visual cortex [Luo20]). The point of this paragraph is to question the influence of the spatial kernel Ψ on the dynamics of (2.28) as $n \rightarrow \infty$, at the level of both LLN and CLT results for the empirical measure. Of course, if Ψ were to be smooth, the results of the previous sections would apply: we are interested here in situations where Ψ is singular.

2.4.1 Law of large numbers

In [L3], we concentrate on the case $I := [0, 1]^d$ with deterministic regular positions $x_i = \frac{i}{n}$, $i \in \{1, \dots, n\}^d$ and aim at looking at the behavior as $n \rightarrow \infty$ of the spatially-extended empirical measure

$$\mu_t^n := \frac{1}{n} \sum_{i=1}^n \delta_{(\theta_t^{i,n}, x_i)} \quad (2.29)$$

Once again, typical assumptions are one-sided Lipschitz continuity for F (2.6) and boundedness and uniform Lipschitz concerning the interaction kernel Γ . It is supposed for simplicity in [L3] that the initial conditions in (2.28) are i.i.d. with appropriate moment conditions (but one

could very well relax this assumption towards non identically distributed initial conditions). The corresponding nonlinear Fokker-Planck equation is then $\mu_t(d\theta, dx)$ weak solution to

$$\begin{aligned} \langle \mu_t, f \rangle &= \langle \mu_0, f \rangle + \int_0^t \frac{1}{2} \langle \mu_s, \operatorname{div}(\sigma \sigma^\dagger \nabla f) \rangle ds \\ &\quad + \int_0^t \left\langle \mu_s, \nabla f \cdot \left\{ \int \Gamma(\cdot, \theta') \Psi(\cdot, x') \mu_s(d\theta', dx') + F \right\} \right\rangle ds \end{aligned} \quad (2.30)$$

The hypotheses on the kernel Ψ in [L3] are rather technical, let us only mention here the most important one: we suppose that there exists some $\iota \in (0, 1]$ such that

$$\delta\Psi(x, y) := \int_I |\Psi(x, z) - \Psi(y, z)| dz \leq C |x - y|^\iota, \quad x, y \in I. \quad (2.31)$$

In other words, although the kernel Ψ might be singular, the quantity

$$D_\Psi^{\text{in}}(x) := \int_I \Psi(x, z) dz, \quad x \in I \quad (2.32)$$

that one might interpret as the *macroscopic indegree* of x (i.e. the sum of all incoming contributions of other all nodes z within the macroscopic graph induced on I^2 by the kernel Ψ) is regular in x . The assumption (2.31) includes particularly the following examples:

Example 2.4.1 (*P*-nearest neighbor model). Take $I = [0, 1]^d$ and $\Psi(x, y) = \mathbf{1}_{|x-y| \leq R}$ for some $R \in (0, 1]$: this corresponds to the situation where one particle only interacts with a proportion R of closest particles [OMHS11]. Assumption (2.31) is satisfied for $\iota = 1$.

Example 2.4.2 (Singular polynomial interaction). Take $I = [0, 1]^d$ and

$$\Psi(x, y) = \frac{1}{|x - y|^\alpha}, \quad x \neq y \in I \quad (2.33)$$

for some $\alpha \in [0, d)$. This example is originally motivated by the XY model with long-range interaction [GPR12, GCR12]. Assumption (2.31) is satisfied for $\iota = (d - \alpha) \wedge 1$.

We see here from (2.30) and the above examples a technical difficulty that is intrinsic to the model: the kind of regularity we need on the test functions $f(\theta, x)$ in (2.30) is significantly different w.r.t. the θ variable (in this respect, some \mathcal{C}^2 regularity is sufficient) and w.r.t. the x -variable (dealing with possibly discontinuous functions, the only relevant regularity w.r.t. x is only L^p for some $p \geq 1$). This lack of regularity w.r.t. the spatial variable rises two main questions: first, concerning the well-posedness of (2.30) (and the regularity of the solution w.r.t. the x variable) and secondly, when it comes to the convergence of the empirical measure (2.29) towards (2.30), the question being which topology one should put on $\mathcal{P}(\mathcal{X} \times I)$, as the weak topology may not be sufficient. The first question of well-posedness of (2.30) is addressed in [L3, Prop. 2.19]:

Proposition 2.4.3 (Proposition 2.19 in [L3]). *Under the previous hypotheses and technical assumptions that include both Examples 2.4.1 and 2.4.2, for every initial condition $\mu_0(d\theta, dx)$, there exists a unique solution $\mu \in \mathcal{P}(\mathcal{C}([0, T], \mathcal{X}) \times I)$ to (2.30).*

Existence follows from ad-hoc extensions of Sznitman's fixed-point argument for the existence of a nonlinear process associated to (2.28), in the same spirit as for Proposition 2.1.1. Uniqueness follows from Theorem 2.4.5 below, based on the observation that *any solution* to (2.30) is necessarily the weak limit of the empirical measure (2.29).

Remark 2.4.4 (Spatial regularity of μ). *Writing the solution μ to (2.30) as $\mu(d\theta, dx) = \mu^x(d\theta)dx$, it is proven in [L10, Lemma A.3] that $x \mapsto \mu^x$ has the same regularity as $\delta\Psi$ in (2.31): there exists $C > 0$ such that for any Lipschitz function $\theta \mapsto f(\theta)$, $|\int f(\theta) \mu_t^x(d\theta) - \int f(\theta) \mu_t^y(d\theta)| \leq \|f\|_{Lip} C e^{Ct} \delta\Psi(x, y)$.*

The main point of [L3] is to address the second point raised above, that is the convergence of the empirical measure (2.29) towards (2.30). The strategy is to construct an ad-hoc distance on $\mathcal{P}(\mathcal{X} \times I)$ that is adapted to test functions with the kind of regularity of Examples 2.4.1 and 2.4.2. We specify here only the main lines of construction adapted to singular interactions of Example 2.4.2. For fixed $a \in I$, denote by \mathcal{C}_a (with corresponding semi-norm $\|\cdot\|_a$) as the set of test functions $(\theta, x) \mapsto f(\theta, x)$ such that, when renormalised by $|x - a|^\alpha$, are regular in both (θ, x) : require that $|x - a|^\alpha f(\theta, x)$ is bounded, Lipschitz in θ and with Hölder regularity in x . Then for fixed $K \geq 1$, if $\mathcal{D}_K := \left\{ \left(\frac{j_1}{K}, \dots, \frac{j_d}{K} \right), 0 \leq j_1 \leq K, \dots, 0 \leq j_d \leq K \right\}$ is the regular discretization of $[0, 1]^d$ of order K , define the sequence $(d_K)_{K \geq 1}$ of distances on $\mathcal{P}(\mathcal{X} \times I)$ given by $d_K(\lambda, \nu) := \sup_f \left(\mathbf{E}(|\langle f, \lambda - \nu \rangle|^p)^{1/p} \right)$ for some appropriate $p \geq 1$, where the supremum is taken over all $f \in \cup_{a \in \mathcal{D}_K, 1 \leq K' \leq K} \mathcal{C}_a, \|f\|_a \leq 1$. The final distance follows a standard Fréchet construction from the intermediate distances d_K : define $d_\infty(\lambda, \nu) := \sum_{K \geq 1} a_K d_K(\lambda, \nu)$, for some appropriate choice of summable family $(a_K)_{K \geq 1}$. We see here the point of d_∞ : it is to follow the singularity (uniformly as $n \rightarrow \infty$) of test functions of the form $g(\theta)\Psi(x, \frac{k}{n})$ that appear naturally in (2.28). Note that the distance d_∞ is somehow weaker than the usual bounded-Lipschitz distance d_{BL} in the sense that the supremum in test functions is outside the expectation, not inside (as we had for the extensions of Proposition 2.1.3 mentioned earlier). The main result of [L3] is then

Theorem 2.4.5 (Theorem 2.18 of [L3]). *Under the previous hypotheses, there exists some $C > 0$, such that for any $\epsilon > 0$ and any arbitrary solution μ to (2.30),*

$$\sup_{t \in [0, T]} d_\infty(\mu_t^n, \mu_t) \leq C \begin{cases} n^{-\{(\frac{d}{2} - \epsilon) \wedge 1\}} & \text{if } \alpha \in [0, \frac{d}{2}), \\ (\ln n) n^{-\{\frac{d}{2} \wedge 1\}} & \text{if } \alpha = \frac{d}{2}, \\ (\ln n) n^{-\{(d - \alpha) \wedge 1\}} & \text{if } \alpha \in (\frac{d}{2}, d). \end{cases} \quad (2.34)$$

The proof is based on a propagator argument similar to the one mentioned in the proof of Proposition 2.1.1: take here $d = 1$ for simplicity. Note at this point that we have not proven uniqueness of a solution to (2.30): let μ be *any of such solutions*. Using again the propagator $P_{s,t}$ defined in (2.10), a way to couple μ_t^n with μ_t is to consider the evolution of $P_{s,t}f(\theta^{k,n}(s))$ for any $k = 1, \dots, n$, where $\theta^{k,n}$ solves (2.28): Ito's formula gives (see [L3, Lem. 4.3]), for any regular test function $(\theta, x) \mapsto f(\theta, x)$,

$$\begin{aligned} \langle \mu_{n,t} - \mu_t, f \rangle &= \langle \mu_{n,0} - \mu_0, f \rangle + \frac{1}{n} \sum_{k=1}^n \int_0^t \partial_\theta (P_{s,t}f) (\theta_s^{k,n}, x_k) dB_{k,t} \\ &+ \frac{1}{n} \sum_{k=1}^n \int_0^t \partial_\theta (P_{s,t}f) (\theta_s^{k,n}, x_k) [\langle \mu_{n,s} - \mu_s, \Gamma(\theta_s^{k,n}, \cdot) \Psi(x_k, \cdot) \rangle] dt := (A) + (B) + (C) \end{aligned} \quad (2.35)$$

The key to Theorem 2.4.5 is to note that $\|P_{s,t}f\|_a \leq C \|f\|_a$, uniformly over regular test functions f (see [L3, Lem. 4.4]). Then, Theorem 2.4.5 boils to down to control each term in (2.35), once one has taken expectation and the supremum in f in (2.35): (B) is a standard Brownian term. The transition described in (2.34) comes from the scaling of the term (A), which captures both the dependence in the initial condition and in the space variable x . The distance d_∞ described above is precisely built up to deal with the coupling term (C), which gives rise to a Grönwall term. Uniqueness of a solution to (2.30) follows readily from (2.34), as the unique limit of the empirical measure μ_n . Note that we have left aside an important technical issue: the calculations leading to (2.35) require F to be Lipschitz, whereas we only require the one-sided Lipschitz condition (2.6). This can be circumvented by observing that all the previous calculations do not actually depend on the Lipschitz constant of F , only in the constant C_F in (2.6). Hence, if F is not Lipschitz, one can replace the dissipative F by its Yosida approximation [Cer01, App. A] $F_\lambda(\theta) := F(R_\lambda(\lambda\theta))$, where $R_\lambda(\theta) = (\lambda - F(\theta))^{-1}$: F_λ is now Lipschitz (with a constant that diverges with $\lambda \rightarrow \infty$) but one-sided Lipschitz with the same constant C_F , uniformly in λ . The point is then to apply the previous estimates with F_λ in place of F and then proceed with approximation arguments with $\lambda \rightarrow \infty$. We refer to [L3, § 7] for further details. The conclusion of Theorem 2.4.5 is a strong indication that the spatial singularities induces a phase transition in the scaling of the law

of large numbers: in the subcritical regime $\alpha < \frac{d}{2}$, the speed of convergence is arbitrarily close to the Gaussian scaling $n^{-\{\frac{d}{2} \wedge 1\}}$, whereas (forgetting the $\ln n$ term that is mostly technical) the speed of convergence is essentially given by $n^{-\{(d-\alpha) \wedge 1\}}$ (note here that the rates of convergence in (2.34) are non optimal, see the next paragraph below). Similar techniques apply to the case of Example 2.4.1 and one obtains the usual Gaussian scaling $n^{-\{\frac{d}{2} \wedge 1\}}$ for any $R \in (0, 1]$, see [L3, Th. 2.13].

2.4.2 Fluctuations

The natural question is to understand whether the rates given by (2.34) are optimal (they are not) and whether it is possible to obtain some exact CLT results concerning (2.29) around its mean-field limit (2.30). This is the purpose of [L6] where we consider Example 2.4.2 when $I = \mathbb{T}$ with $d = 1$ for simplicity. The above discussion suggests to consider the following fluctuation process

$$\eta_t^n := a_n (\mu_t^n - \mu_t) \quad (2.36)$$

where $a_n = n^{\frac{1}{2}}$ when $\alpha \in [0, 1/2)$ and $a_n := n^{1-\alpha}$ when $\alpha \in (1/2, 1)$. It is easy to see that these scalings are indeed the correct ones: the convergence of μ^n to μ comes from the competition between the convergence of the empirical distribution of both the initial condition $\theta_0^{i,n}$ and Brownian motions B_i (which scales typically as $n^{1/2}$) and the convergence with respect to the spatial variable x_i . Setting $F \equiv 0$ and $\Gamma \equiv 1$ in (2.28) to fix ideas, everything boils down to the approximation of the integral $\int_I \frac{1}{|x-x_i|^\alpha} dx$ by the Riemann sum $\frac{1}{n} \sum_{j=1}^n \frac{1}{|x_j-x_i|^\alpha}$, for which the rate of this last convergence is exactly $n^{1-\alpha}$. So the previous choice of a_n simply corresponds to the predominant scaling in both cases (see [FR01, eq. (27)] for heuristics on this derivation in a similar context). This intuition also suggests that when $\alpha < \frac{1}{2}$, the randomness prevails and one should obtain Gaussian fluctuations as $n \rightarrow \infty$, whereas when $\alpha > \frac{1}{2}$ the randomness disappears under the scaling $n^{1-\alpha}$ and one should obtain a deterministic limit for η^n . The main result of [L6] is precisely to make this intuition rigorous: we show that the fluctuation process η^n converges to the unique solution of a linear stochastic partial differential equation when $\alpha < \frac{1}{2}$ and that η^n has a deterministic limit in the supercritical case $\alpha > \frac{1}{2}$, see Theorem 2.4.6.

An auxiliary fluctuation process

We use here similar Hilbertian techniques [FM97] already introduced earlier (recall § 2.2.1): we consider η^n as a random distribution over test functions $(\theta, x) \mapsto f(\theta, x)$. There is however a major difficulty here: the embeddings techniques in [FM97] require a sufficient regularity (typically \mathcal{C}^k for sufficiently large k) on the test functions f . This is in apparent contradiction with our present case, since in order to study η^n in (2.36), one needs to consider test functions $(\theta, x) \mapsto f(\theta, x)$ that reproduce the singularities of the kernel Ψ w.r.t. the variable x , and hence are *not regular in the x variable*. We bypass this difficulty by introducing a key auxiliary process, that we call *two-particle fluctuation process*:

$$\mathcal{H}_t^n := a_n \left(\frac{1}{n^2} \sum_{i,j=1}^n \Psi(x_i, x_j) \delta_{(\theta_t^{i,n}, x_i, \theta_t^{j,n}, x_j)} - \mu_t^n \otimes \mu_t(\Psi \cdot) \right), \quad t \in [0, T], \quad (2.37)$$

One can easily see that \mathcal{H}^n captures the correct fluctuations induced by the space variables (especially in the supercritical case $\alpha > \frac{1}{2}$): applying \mathcal{H}^n to the constant test function $g \equiv 1$, one obtains that $\langle \mathcal{H}_t^n, 1 \rangle = \frac{1}{n} \sum_{i=1}^n \left\{ a_n \left(\frac{1}{n} \sum_{j=1}^n \Psi(x_i, x_j) - \int_I \Psi(x_i, \tilde{x}) d\tilde{x} \right) \right\}$, which is exactly of order 1 when $\alpha > \frac{1}{2}$.

From a purely technical point of view, what makes the use of \mathcal{H}^n critical is that it enables to separate the issue of the singularity of the spatial kernel Ψ from the issue of the regularity of the test functions: it is the weighted process \mathcal{H}^n itself that carries the singularity in (x, \tilde{x}) (through the weight Ψ), not the test functions. Hence, we are allowed to consider test functions as regular as required in *all* variables (θ, x) so that the Sobolev embeddings techniques of [FM97] remain applicable. Note that this two-level process is not only a convenient technical tool, but

a relevant part of the fluctuation result: in the supercritical case $\alpha > \frac{1}{2}$, the fluctuations of the system are necessarily described in terms of the joint limit (η, \mathcal{H}) of (η^n, \mathcal{H}^n) as $n \rightarrow \infty$ (see Theorem 2.4.6). The starting point of the analysis is to write a semimartingale decomposition for both processes η^n and \mathcal{H}^n . We see in particular that the part that involves the singular kernel Ψ in the semimartingale decomposition of η^n is completely expressed in terms of \mathcal{H}^n : for all $t \in [0, T]$, one has

$$\begin{cases} \eta_t^n &= \eta_0^n + \int_0^t \mathcal{L}_s^{1,*} \eta_s^n ds + \int_0^t \Phi^* \mathcal{H}_s^n ds + \mathcal{M}_t^{1,n}, \\ \mathcal{H}_t^n &= \mathcal{H}_0^n + \int_0^t \mathcal{L}_s^{2,*} \mathcal{H}_s^n ds + \int_0^t R_s^n ds + \mathcal{M}_t^{2,n}, \end{cases} \quad (2.38)$$

where \mathcal{L}_s^1 , \mathcal{L}_s^2 and Φ are explicit linear operators (acting on smooth functions), R^n is a remainder term and $\mathcal{M}_t^{l,n}$ for $l = 1, 2$ are martingales. Based on this joint semimartingale description, the main result is

Theorem 2.4.6 (Theorems 2.7 and 2.8 in [L6]). *Under regularity assumptions on F and Γ and moment conditions on the initial condition, there exist appropriate Sobolev spaces H_1 and H_2 with negative regularity, such that*

1. *if $\alpha \in [0, 1/2)$, the random process η^n converges in law in $\mathcal{C}([0, T], H_1)$ to the unique solution to the well-posed linear stochastic partial differential equation*

$$\eta_t = \eta_0 + \int_0^t \mathcal{L}_s^{1,*} \eta_s ds + \mathcal{M}_t^1 \quad (2.39)$$

where η_0 and \mathcal{M}^1 are explicit independent Gaussian processes.

2. *if $\alpha \in (1/2, 1)$, the random process (η^n, \mathcal{H}^n) converges in law in $\mathcal{C}([0, T], H_1 \otimes H_2)$ to the unique solution to the well-posed coupled system of linear partial differential equations*

$$\begin{cases} \eta_t &= \int_0^t \mathcal{L}_s^{1,*} \eta_s ds + \int_0^t \Phi^* \mathcal{H}_s ds, \\ \mathcal{H}_t &= \mathcal{H}_0 + \int_0^t \mathcal{L}_s^{2,*} \mathcal{H}_s ds, \end{cases} \quad (2.40)$$

where \mathcal{H}_0 is some nontrivial initial condition.

Some comments on Theorem 2.4.6: (2.39) reduces to the linear SPDE found in the homogeneous mean-field regime in [FM97] when $\alpha = 0$: in the subcritical case $\alpha < \frac{1}{2}$, the spatial damping on the interactions in (2.28) is not strong enough to have an effect on the behavior of the fluctuations of the system and the mean-field fluctuations remain universal and Gaussian. The main point is to realise that the two-level fluctuations \mathcal{H}^n vanishes as $n \rightarrow \infty$ in this case. To see this, the strategy is to show that the limit \mathcal{H} satisfies in any case the linear equation $\mathcal{H}_t = \mathcal{H}_0 + \int_0^t \mathcal{L}_s^{2,*} \mathcal{H}_s ds$ and that the initial condition $\mathcal{H}_0 \equiv 0$ in the subcritical case. Hence, by linearity, $\mathcal{H}_t \equiv 0$ for all $t \geq 0$ and the fluctuations are described in terms of the process η_t only. On the contrary, in the supercritical case, the spatial constraints prevail and the first correction to the fluctuations as $n \rightarrow \infty$ is deterministic. Note that in this case, the only point that makes this limit nondegenerate is the nontrivial initial condition \mathcal{H}_0 of the two-level fluctuation process. Having defined both processes (η^n, \mathcal{H}^n) , the strategy of proof follows the same procedure as in [FM97] (that is tightness and identification of the limit) with the notable technical difficulty that \mathcal{H}^n acts now on test functions on four variables $(\theta, x, \tilde{\theta}, \tilde{x})$. A key (and quite technical) argument at this point is to close the fluctuation decomposition by showing that the remainder term R^n in (2.38) effectively goes to 0 as $n \rightarrow \infty$. In comparison with the similar fluctuation result for random graphs in Section 3.2 below, dealing with this remainder term is actually conceptually easier than for Section 3.2, as there is no superposition here of an additional randomness: the control of R^n is essentially deterministic (although very technical as one has to deal with test functions of four variables $(\theta, x, \tilde{\theta}, \tilde{x})$ with singular dependence in x, \tilde{x}). The use of weighted empirical processes such as \mathcal{H}^n in the context of interacting particle systems is reminiscent of previous works (see in particular the articles of Kurtz and Xiong [KX99, KX04] on particle approximations for nonlinear SPDEs). The supercritical case may also be related to a class of models previously studied in the literature, that is *moderately interacting diffusions* [JM98, Oel87]. This class of models also exhibits deterministic fluctuations (see [JM98] p. 755), but one should point out that the precise scaling a_n is not explicitly known [JM98, Rem. 3.15].

2.5 Comments and perspectives

When $\alpha > \frac{1}{2}$, the leading term in the asymptotic expansion of the empirical measure μ^n around μ is deterministic, of order $n^{-(1-\alpha)}$. A natural question would be about the existence and the nature of the next term in this expansion. Concerning the dependence in the spatial variable, the next term in the expansion of $\frac{1}{n} \sum_{j=1}^n \frac{1}{|x_j|^\alpha}$ is of order n^{-1} , which is in any case smaller than the Gaussian scaling $n^{-1/2}$. Consequently, one should expect the next term in the expansion of μ^n to be Gaussian, of order $n^{-1/2}$. The precise form of this term remains unclear (see [Oel87] for a similar analysis concerning moderately interacting diffusions). Another natural question would be to ask what happens at the critical case $\alpha = \frac{1}{2}$ (that is when the spatial and Gaussian fluctuations are exactly of the same order). Although it is natural to think that the correct scaling is exactly \sqrt{n} , the present work only provides partial answers (all that is proven in [L6] is that the correct scaling is *at least* $a_n = \frac{\sqrt{n}}{\ln n}$). The behavior as $n \rightarrow \infty$ of systems similar to (2.28) in the case $\alpha > 1$ is also of interest (see [FR01] for heuristics). The analysis of this case goes beyond the framework of mean-field analysis and requires alternative techniques. As far the P -nearest neighbors model (Example 2.4.1) is concerned, the above results only cover the case where the proportion $R > 0$ is fixed (that would correspond to the dense case, with the formalism of Chapter 3). The diluted case where $R = R_n \rightarrow 0$ with $n \rightarrow \infty$ remains to be fully investigated, in connection with the framework of moderately interacting diffusions [JM98].

Chapter 3

Beyond the mean-field assumption: spatially-structured models and random graphs

Whereas the nature of the homogeneities considered in Chapter 2 was essentially local (each local disorder ω_i is attached to the particle $\theta^{i,n}$), we are interested here in the case where the disorder lies on the *interaction* between two particles $\theta^{i,n}$ and $\theta^{j,n}$. Recall the setting of (1.27): let $\mathcal{G}_n = (\mathcal{E}_n, \mathcal{V}_n)$ be a generic graph with sets of vertices $\mathcal{E}_n := \{1, \dots, n\}$ and edges $\mathcal{V}_n \subset \{1, \dots, n\}^2$. Encoding the presence (resp. absence) of the edge $i \rightarrow j$ in \mathcal{V}_n by $\xi_{i,j}^n := 1$ (resp. $\xi_{i,j}^n := 0$), define the particle system $(\theta^{1,n}, \dots, \theta^{n,n})$ interacting on the graph \mathcal{G}_n by

$$d\theta_t^{i,n} = F(\theta_t^{i,n})dt + \frac{\kappa_i^n}{n} \sum_{j=1}^n \xi_{i,j}^n \Gamma(\theta_t^{i,n}, \theta_t^{j,n}) dt + \sigma \cdot dB_t^i, \quad i = 1, \dots, n, \quad t \in [0, T], \quad (3.1)$$

Here, κ_i^n is a possibly inhomogeneous dilution parameter in order to compensate for the local sparsity of the graph around vertex i . At this point, the chosen graph \mathcal{G}_n could be random (build upon some probability \mathbb{P}_g) or simply deterministic.

Moving from the homogeneous (1.3) to (3.1) constitutes a considerable step further in the difficulty as one formally leaves the mean-field framework: the interaction in (3.1) is no longer a functional of the empirical measure $\mu_t^n = \frac{1}{n} \sum_{i=1}^n \delta_{\theta_t^{i,n}}$ but of the collection of local empirical measures

$$\mu_t^{i,n} := \frac{\kappa_i^n}{n} \sum_{j=1}^n \xi_{i,j}^n \delta_{\theta_t^{j,n}}, \quad i = 1, \dots, n. \quad (3.2)$$

In particular, there is no immediate argument that for two vertices $i \neq j$, the local empirical measure $\mu_t^{i,n}$ around i should be close to $\mu_t^{j,n}$, nor that both should be close to the global empirical measure μ_t^n . By Ito's formula to (3.1), the dynamics of these local empirical measures depend themselves on empirical measures of higher order in the graph so that the main question is to find how to properly close this hierarchy as $n \rightarrow \infty$. Having in mind Question 1.2.5, we want to understand under which condition on the graph \mathcal{G}_n the weighted measure

$$\hat{\mu}_t^{i,n} := \mu_t^{i,n} - \mu_t^n = \frac{1}{n} \sum_{j=1}^n (\kappa_i^n \xi_{i,j}^n - 1) \delta_{\theta_t^{j,n}} \quad (3.3)$$

effectively goes to 0 as $n \rightarrow \infty$. The key (and the major present difficulty) to this convergence is to find a way to decouple in (3.3) what is relevant to the proximity of the graph sequence w.r.t. the complete graph (that is $(\kappa_i^n \xi_{i,j}^n - 1)$ in (3.3)) from the dynamics $\delta_{\theta_t^{j,n}}$ itself. This is not a trivial task as $\theta_t^{j,n}$ intricately depends on the graph itself, so that it is unclear if any estimate one might have on graph observables such as $\frac{1}{n} \sum_{j=1}^n (\kappa_i^n \xi_{i,j}^n - 1)$ would propagate to $\hat{\mu}_t^{i,n}$.

3.1 Homogeneous graphs: Law of Large Numbers

3.1.1 Uniform propagation of chaos

We first address Question 1.2.5 at the level of the law of large numbers, that is the convergence of the empirical measure μ_t^n of (3.1) to μ_t solution to the nonlinear Fokker-Planck equation (1.4). The first elementary and natural approach of [L5] (but yet quite unsatisfactory, see commentaries below) is to tackle this question by proving global propagation of chaos for (3.1), that is to look at the following convergence

$$\sup_{i=1,\dots,n} \mathbf{E} \left[\sup_{s \in [0, T]} |\theta_s^{i,n} - \bar{\theta}_s^{i,n}|^2 \right] \xrightarrow{n \rightarrow \infty} 0 \quad (3.4)$$

where $(\bar{\theta}^{1,n}, \dots, \bar{\theta}^{n,n})$ is the usual sequence of i.i.d. nonlinear mean-field synchronous coupling (with the same initial conditions and Brownian motions as for (3.1), recall (1.5)). A very simple argument based on Ito's formula and the Lipschitz continuity of Γ shows that

$$\begin{aligned} \mathbf{E} \left[\sup_{0 \leq s \leq T} |\theta_s^{i,n} - \bar{\theta}_s^{i,n}|^2 \right] &\leq C \int_0^t \mathbf{E} \left[\sup_{0 \leq v \leq u} |\theta_v^{i,n} - \bar{\theta}_v^{i,n}|^2 \right] du \\ &+ \int_0^t \mathbf{E} \left[\left| \frac{\kappa_i^n}{n} \sum_{j=1}^n \xi_{i,j}^n \Gamma(\theta_u^{i,n}, \theta_u^{j,n}) - \int \Gamma(\bar{\theta}_u^i, \theta') \mu_u(d\theta') \right|^2 \right] du \end{aligned} \quad (3.5)$$

The idea of [L5] is very simple: whereas in the last term of (3.5), $\xi_{i,j}^n$ and $\Gamma(\theta_u^{i,n}, \theta_u^{j,n})$ are coupled in a nontrivial way, the situation gets much better if one considers instead $\frac{\kappa_i^n}{n} \sum_{j=1}^n \xi_{i,j}^n \Gamma(\bar{\theta}_u^{i,n}, \bar{\theta}_u^{j,n})$, as, by construction, the nonlinear process $\bar{\theta}_t^{i,n}$ no longer depends on the graph. Noting that the difference $\Gamma(\theta_u^{i,n}, \theta_u^{j,n}) - \Gamma(\bar{\theta}_u^{i,n}, \bar{\theta}_u^{j,n})$ only gives rise to Grönwall terms, we see from (3.5) that the natural condition to obtain the convergence (3.4) is to require that

$$b_n = b_n(\mathcal{G}_n) := \sup_{i=1,\dots,n} \left| \frac{1}{n} \sum_{j=1}^n \hat{\xi}_{i,j}^n \right| \xrightarrow{n \rightarrow \infty} 0, \text{ for } \hat{\xi}_{i,j}^n := \kappa_i^n \xi_{i,j}^n - 1. \quad (3.6)$$

Condition (3.6) is a way to measure the proximity of the graph \mathcal{G}_n with the complete graph K_n which is essentially a condition of *uniform homogeneity on the degrees*: suitably renormalized, each vertex in \mathcal{G}_n has its degree $d_n^i = \sum_{j=1}^n \xi_{i,j}^n$ which approaches n . Hence, we obtain

Theorem 3.1.1 (Theorem 1.1 in [L5]). *Suppose that the initial conditions in (3.1) are i.i.d. with law μ_0 . Suppose that there is some κ_n such that $\kappa_n^i = \kappa_n$ for all $i = 1, \dots, n$ and $\frac{\kappa_n}{n} \xrightarrow{n \rightarrow \infty} 0$. Under Lipschitz regularity of the coefficients F and Γ , assuming condition (3.6), there exists some constant $C = C_{F,\Gamma} > 0$ and n_0 such that for all $n \geq n_0$ and any $t \geq 0$,*

$$\sup_{i=1,\dots,n} \mathbf{E} \left[\sup_{s \in [0, t]} |\theta_s^{i,n} - \bar{\theta}_s^{i,n}|^2 \right] \leq C \left(\frac{\kappa_n}{n} + b_n^2 \right) \exp(Ct) \xrightarrow{n \rightarrow \infty} 0. \quad (3.7)$$

In particular, for any sequence t_n such that $\left(\frac{\kappa_n}{n} + b_n^2 \right) \exp(Ct_n) = o(1)$,

$$\mathbf{E} \left[\sup_{s \in [0, t_n]} d_{BL}(\mu_{n,s}, \mu_s) \right] \xrightarrow{n \rightarrow \infty} 0, \quad (3.8)$$

where $(\mu_t)_{t \in [0, T]}$ solves (1.4) with initial condition μ_0 .

Example 3.1.2 (Erdős-Rényi random graphs, $ER(n, p_n)$). *Let $p_n \in (0, 1]$ and let $\xi_{i,j}^n, i, j = 1, \dots, n$ be i.i.d. Bernoulli random variables with parameter p_n (that is $\mathbb{P}(\xi_{i,j}^n = 1) = p_n = 1 - \mathbb{P}(\xi_{i,j}^n = 0)$). One is interested in two situations here: the dense case, when $p_n \xrightarrow{n \rightarrow \infty} p \in (0, 1]$ (this corresponds to a situation where each vertex interacts with a proportion of size n of the population) and the diluted case (or vanishing degree case), when $p_n \xrightarrow{n \rightarrow \infty} 0$. In such a case,*

for $\kappa_n = \frac{1}{p_n}$, condition (3.6) is valid for almost every realisation of the graph (see [L5, Prop. 1.3]) provided the following dilution condition is met

$$\frac{np_n}{\ln n} \rightarrow \infty \text{ as } n \rightarrow \infty. \quad (3.9)$$

Example 3.1.3 (Random regular graphs). *One can build a graph in which each vertex has degree $d = d(n)$ provided $3 \leq d < n$ and dn even [Bol01] and condition (3.6) is met for $\kappa_n = \frac{n}{d(n)}$ with $d(n) \rightarrow \infty$.*

3.1.2 Discussion

On the optimality of the dilution condition

First note that the hypothesis $\frac{\kappa_n}{n} \xrightarrow[n \rightarrow \infty]{} 0$ in Theorem 3.1.1 is the minimal condition where to expect some mean-field behavior for (3.1). In case of Example 3.1.2, this corresponds to the condition (slightly better than (3.9))

$$np_n \rightarrow \infty \quad (3.10)$$

i.e., the expected degree of each vertex goes to ∞ as $n \rightarrow \infty$. As far as the convergence (3.8) of the empirical measure is concerned, condition (3.10) is certainly optimal: in the critical regime $np_n \xrightarrow[n \rightarrow \infty]{} \lambda$ for some $\lambda > 0$, the degree of each vertex remains a.s. finite (asymptotically Poisson(λ)) and it is well-known that $ER(n, p_n)$ converges (in the sense of local topology on graphs [Bor16]) to some Galton-Watson tree. The limit of (3.1) is then no longer mean-field but given as interacting diffusions on Galton-Watson trees [ORS20, LRW23].

In the annealed case (i.e. when (3.8) is also averaged w.r.t. the law of the graph), one easily obtains the optimal condition (3.10) [BCW20]. We stress here that there is an intrinsic difficulty in capturing the same optimal condition in the quenched regime. The main reason is that the dilution condition (3.9) (slightly improved into $\liminf_{n \rightarrow \infty} \frac{p_n n}{\log(n)} > 0$ in [CDG20, eq. (2.1)], see § 3.1.4 below for more details) is likely to be optimal for the uniform propagation of chaos (3.4). Indeed, considering Example 3.1.2, $p_n \sim \frac{\ln n}{n}$ is the threshold for connectivity for \mathcal{G}_n : when $p_n = a \frac{\ln n}{n}$ with $a < 1$, there is with high probability about n^{1-a} isolated vertices (see the discussion before [Dur07, Th. 2.8.1]). In particular, the dynamics of these particles are simply Brownian and their law not close to a McKean-Vlasov process. Hence, if one believes in the optimality of condition (3.10) for the convergence of the empirical measure (3.8), a uniform propagation result such as (3.4) is necessarily not an appropriate strategy in this intermediate regime $\frac{1}{n} \ll p_n \ll \frac{\ln n}{n}$. We see here a particularly striking illustration of the effect of non-exchangeability in (3.1): the convergence (3.4) of the law of one single particle is no longer equivalent to the convergence (3.8) of the empirical measure.

Going beyond logarithmic scale

There is further evidence that (3.4) (and its corresponding condition (3.6)), as appealing as it may seem at first look, may not be the correct approach for the convergence (3.8) of the empirical measure. This is discussed at length in [L5], we reproduce here the main arguments. The point is that condition (3.6), although uniform along the graph \mathcal{G}_n , is essentially of *local nature*, as it concerns only the behavior of the renormalized degree of each vertex. In particular, (3.6) does not even require \mathcal{G}_n to be connected. This leads to some phenomenological paradox that we present here.

Example 3.1.4. *Consider the non-disordered Kuramoto model (i.e. Example 1.2.2 in the case $\omega \equiv 0$). Condition (3.6) does not distinguish at all between*

1. *some fully connected $\mathcal{G}_n = ER(n, \frac{1}{2})$*
2. *two disjoint components $\mathcal{G}_n := K_n^{(1)} \sqcup K_n^{(2)}$ where $K_n^{(1)}$ (resp. $K_n^{(2)}$) is the complete graph with vertices $\{1, \dots, n/2\}$ (resp. $\{n/2 + 1, \dots, n\}$).*

Indeed, both cases satisfy (3.6) with $\kappa_n = 2$. Applying then Theorem 3.1.1, with i.i.d. initial conditions uniformly distributed on \mathbb{T} shows that in both cases, the empirical measure μ^n converges to the same μ , solution to the following non-disordered version of (1.15)

$$\partial_t \mu_t(\theta) = \frac{1}{2} \partial_\theta^2 \mu_t(\theta) - \partial_\theta [\mu_t(\theta) J * \mu_t(\theta)], \quad (3.11)$$

where $J(\cdot) = -K \sin(\cdot)$, with initial condition $\mu_0 = \frac{1}{2\pi}$.

Having in mind the dynamical properties of the Kuramoto model (Example 1.2.2) roughly sketched in Chapter 1, Example 3.1.4 contradicts somehow the dynamics for t large observed on simulations: when $K > 1$, the system would leave the unstable uniform stationary solution $\frac{1}{2\pi}$ and stabilize around some point $q_{r,\psi}(\cdot) = q_{r,0}(\cdot - \psi)$ of the stable circle of synchronized solutions $\{q_{r,\psi}(\theta) = q_{r,0}(\theta - \psi), \psi \in \mathbb{T}\}$, with $q_{r,0}$ given by (1.21). In the fully connected case 1. of Example 3.1.4, one observes a synchronized unimodal solution whereas the system converges to a bimodal profile (made of two independent unimodal distributions) in the disjoint case 2. Hence, there is evidently a breakdown between the information that carries the mean-field solution μ_t of (3.11) and the empirical measure μ_t^n .

The solution to this apparent paradox lies in the fact that the convergence (3.8) is only valid up to times t_n that cannot go beyond a time scale of order $\ln n$ (which is from a simulation perspective and most of real life systems, essentially a finite time). This impossibility of extending propagation of chaos beyond bounded times scales is something that is already relevant in the homogeneous mean-field case. In a word: there is no *uniform propagation of chaos in the Kuramoto model* (at least not uniformly w.r.t. the initial condition, see [DT21] for more refined results). Let us be more precise on this point: if we rewrite (3.11) in terms of the Fourier coefficients

$$c_k(t) := \int \mu_t(\theta) \cos(k\theta) d\theta \quad \text{and} \quad s_k(t) := \int \mu_t(\theta) \sin(k\theta) d\theta, \quad k \geq 1, \quad (3.12)$$

one verifies that at a linear level all modes decouple and solve the equation $\dot{x} = \lambda_k x$, with $\lambda_k = -k^2/2$ for every mode $k \geq 2$ and $\lambda_1 = (K-1)/2$ (hence these are the two unstable modes, the zero mode is conserved). At an informal level, for initial condition given by i.i.d. random variables uniformly distributed on the circle, a good approximation for the large n evolution of the empirical mean $c_{1,n}(t) := \frac{1}{n} \sum_{j=1}^n \cos(\theta_t^{j,n})$ is the linear SDE

$$dx(t) = \frac{(K-1)}{2} x(t) dt + \frac{1}{\sqrt{n}} dB_t, \quad (3.13)$$

with initial condition $\mathcal{N}(0, 1/n)$, independent of the (standard) Brownian motion B (a rigorous justification of this approximation is close in spirit to the analysis made in [DMPV86] in the context of reaction-diffusion particle systems). The same is true for the sine mode, call $s_{1,n}(t)$ the corresponding empirical mean, with new (independent) Brownian motion and initial condition. The solutions of these equations are centered Gaussian processes. Computing the *synchronization* degree $r_n(t) = \sqrt{(c_{1,n}(t))^2 + (s_{1,n}(t))^2}$ (using (3.13)), one obtains

$$\mathbf{E}[(r_n(t))^2] \approx \frac{1}{n} \exp(2\lambda_1 t) \left(1 + \frac{1}{2\lambda_1} (1 - \exp(-2\lambda_1 t)) \right), \quad (3.14)$$

for n large and as far as the linear approximation is reliable. In analogy with [DMPV86], the fluctuations lead to the escape from the flat state at a time $a \log n$, $a := (2\lambda_1)^{-1}$ in the sense that for $t \leq c \log n$, any $c \in (0, a)$, the empirical measure of the system converges for $n \rightarrow \infty$ to the uniform probability, but for $t = c \log n$, any $c > a$, it converges to one of the *synchronized* solutions. One can convince oneself easily that the previous argument is fairly general and valid for any mean-field process having both an unstable point and a stable structure (here a circle of stationary solutions): starting from this unstable point, the system would leave this point and reach the stable structure on logarithmic times scales. However, it is possible to obtain uniform propagation of chaos estimates, if one starts away from this unstable point, see [DT21] for more details.

3.1.3 Partial conclusion

We deduce from the previous discussion the following important points:

There is no uniform propagation of chaos for the Kuramoto model

This is something already true in the homogeneous mean-field case: the convergence (3.8) cannot be substantially improved beyond logarithmic time scales as there is a breakdown in the information that the mean-field limit μ_t carries. Note here that the previous argument for non uniformity in propagation of chaos is specific to the fact that the initial condition is the unstable point $\frac{1}{2\pi}$. The same argument no longer holds for any initial condition away from $\frac{1}{2\pi}$: in such a case, we indeed refer to the deep result of [DT21] where a weak form of uniform propagation of chaos is proven (in the mean-field case). In the homogeneous mean-field case, the cumulative influence of noise on time scales larger than $\ln n$ will be analysed in details in Chapter 5.

Condition (3.6) is not the correct one

Example 3.1.4 shows that Condition (3.6) is too weak and certainly not sufficient in itself to analyse the dynamics of the empirical measure of (3.1) beyond logarithmic time scales. One necessarily needs to complement this condition in order to push further the analysis of $\mu_{n,t}$. This is essentially due to the fact that (3.6) does not capture the right topology on graphs: choosing $t_j \equiv 1$, we have obviously $b_n \leq b_n^{\infty \rightarrow \infty}$ where

$$b_n^{\infty \rightarrow \infty} := \sup_{i=1, \dots, n} \sup_{t \in \{\pm 1\}^n} \left| \frac{1}{n} \sum_{j=1}^n \hat{\xi}_{i,j}^n t_j \right| \quad (3.15)$$

where the supremum is on every sequence of signs $t = (t_1, \dots, t_n) \in \{\pm 1\}^n$. Here the notation $\infty \rightarrow \infty$ (and $\infty \rightarrow 1$ below) borrows from the formalism of graphs limits introduced by Lovász and Szegedy [LS06] (see also [Jan13, DJ08] for further references): $b_n^{\infty \rightarrow \infty}$ is equal to $\|W^{\mathcal{G}_n} - \mathbf{1}\|_{\infty \rightarrow \infty}$ where $W^{\mathcal{G}_n}$ is the step-graphon associated to \mathcal{G}_n (see e.g. [BCL⁺08, § 3]) and

$$\|W\|_{\infty \rightarrow \infty} := \sup_{\|g\|_{\infty} \leq 1} \left\| \int_{[0,1]} W(\cdot, y) g(y) dy \right\|_{\infty} \quad (3.16)$$

Replacing the local condition (3.6) with the global and stricter $b_n^{\infty \rightarrow \infty} \xrightarrow{n \rightarrow \infty} 0$ is obviously too demanding: in the case of $ER(n, p_n)$ (Example 3.1.2), $b_n^{\infty \rightarrow \infty} = \sup_{i=1, \dots, n} \frac{1}{n} \sum_{j=1}^n |\hat{\xi}_{i,j}^n|$ is of order $2(1 - p_n)$, which only goes to 0 in the asymptotically mean-field case $p_n \rightarrow 1$. We see here the problem of considering (3.6): we are not looking at the right topology w.r.t. graph convergence. It would be preferable to consider the norm $\|W^{\mathcal{G}_n} - \mathbf{1}\|_{\infty \rightarrow 1}$ with

$$\|W\|_{\infty \rightarrow 1} := \sup_{\|g\|_{\infty} \leq 1} \int_{[0,1]} \left| \int_{[0,1]} W(x, y) g(y) dy \right| dx \quad (3.17)$$

This corresponds to requiring that

$$b_n^{\infty \rightarrow 1} := \|W^{\mathcal{G}_n} - \mathbf{1}\|_{\infty \rightarrow 1} = \sup_{s, t \in \{\pm 1\}^n} \frac{1}{n^2} \sum_{i, j=1}^n \hat{\xi}_{i,j}^n s_i t_j \xrightarrow{n \rightarrow \infty} 0. \quad (3.18)$$

The latter condition is well-known in the literature as it captures the convergence of the graph \mathcal{G}_n towards the complete graph in terms of the cut-norm $\|\cdot\|_{\square}$ (see e.g. [BCL⁺08])

$$\|W\|_{\square} := \sup_{S, T \subset [0,1]} \left| \int_{S \times T} W(x, y) dx dy \right| \quad (3.19)$$

Condition (3.18) has been considered by Coppini [Cop22] for the longtime dynamics of the Kuramoto model on ER graphs and is valid under the optimal condition (3.10) only (see [Cop22, Lem. A.3]). Another indication that condition (3.18) may be much more relevant than (3.6) for the convergence of the empirical measure is that (3.18) carries much more information on the structure of \mathcal{G}_n : it implies in particular the existence of a unique giant component in \mathcal{G}_n [Cop22, Lem. A.2] so that Example 3.1.4 is no longer significant.

There is absolute necessity of considering initial conditions for (3.1) that are not exchangeable

The convergence result (3.7) requires crucially to have independent initial conditions $\theta_0^{1,n}, \dots, \theta_0^{n,n}$ (this is inherent to the synchronous coupling technique, as already mentioned in Remark 1.1.1). In view of the discussion of the first paragraph of § 3.1.2, improving the condition (3.9) into (3.10) for the convergence (3.8) necessarily requires to discard this assumption of independence. There is also a crucial need to consider initial conditions with laws that are non necessarily identical. Here, Example 3.1.4 is somehow misleading, as its conclusions rely heavily on the choice of uniform distribution on \mathbb{T} of the initial condition $(\theta_0^{1,n}, \dots, \theta_0^{n,n})$: in Case 2. of Example 3.1.4, choosing different laws μ_0^1 and μ_0^2 for the initial configuration of the two components leads to the convergence of $\mu_{n,t}$ to $\mu_t^{\text{Case 2}} := \frac{1}{2}(\mu_t^1 + \mu_t^2)$ where both μ^k solves (3.11) with initial condition μ_0^k , $k = 1, 2$. The same initial condition for Case 1. leads to the convergence to $\mu^{\text{Case 1}}$ solving (3.11) with initial condition $\mu_0^{\text{Case 1}} := \frac{1}{2}(\mu_0^1 + \mu_0^2)$. But (3.11) is nonlinear: $\mu^{\text{Case 1}} \neq \mu^{\text{Case 2}}$ unless $\mu_0^1 = \mu_0^2$. This observation is elementary, but it reveals the fundamental difficulty in going beyond the results of Theorem 3.1.1: the state of the system depends on the graph at any positive times even though it might not be the case at $t = 0$.

3.1.4 Extensions and semimartingale approach

A first generalisation of Theorem 3.1.1 was provided by Coppini, Dietert and Giacomin in [CDG20]: the strategy is still to proceed with some coupling, but the point is now to couple the particle system $\theta^{i,n}$ (3.1) not to its nonlinear limit $\bar{\theta}^{i,n}$ but rather to its mean-field companion $\tilde{\theta}^{i,n}$ (that is the same system as (3.1) but that lives on the complete graph, with the same initial condition and Brownian motion). The proof is then two-fold: first, to prove that $\frac{1}{n} \sum_{i=1}^n \mathbf{E} \left[\sup_{s \in [0, T]} \left| \theta_s^{i,n} - \tilde{\theta}_s^{i,n} \right|^2 \right] \xrightarrow{n \rightarrow \infty} 0$ (deterministically in the graph and noise) and second, to take advantage of the fact that $\tilde{\mu}^n$ (the empirical measure for the mean-field system $\tilde{\theta}$) converges to (1.4) whatever the initial condition, provided only $\tilde{\mu}_0^n$ converges to μ_0 at time $t = 0$ [Gär88, CDFM20]. The result of [CDG20] hence extends the convergence (3.8) to $ER(n, p_n)$ graphs with non-necessarily i.i.d. initial conditions (one only requires that they do not depend on the graph) under the single assumption that μ_0^n converges to μ_0 , under the dilution condition (slightly weaker than (3.9)) $\liminf_{n \rightarrow \infty} \frac{p_n n}{\log(n)} > 0$. Note also that some Large Deviation Principle for the empirical measure (with the same rate function as in the mean-field case) is also proven in [CDG20].

General convergence result

These results being stated, there is however room for improvement in at least two directions: first to consider initial conditions that possibly depend on the graph itself, and second to prove the convergence (3.8) under the optimal condition (3.10) only. This point is precisely addressed in a previous work by Coppini [Cop22, Th. 2.1], for the particular dynamics of Kuramoto oscillators. The convergence in [Cop22, Th. 2.1] is somehow weaker than Theorem 3.1.5 below, as it is stated in probability, whereas (3.21) is in expectation. Note also that Theorem 3.1.5 also addresses the convergence of the local empirical measures $\mu^{i,n}$ defined in (3.2).

Theorem 3.1.5 (Theorem 2.4 in [L14]). *Suppose that the state space is \mathbb{T} with regularity conditions on Γ and F . Suppose that \mathcal{G}_n is the realisation of an $ER(n, p_n)$ graph. Suppose that the initial conditions $(\theta_0^{1,n}, \dots, \theta_0^{n,n})$ are chosen independently on the Brownian motions (B^1, \dots, B^n) (but not necessarily i.i.d. and they may depend on the graph), such that their empirical measure μ_0^n converges weakly to some μ_0 in the following way:*

$$d_{BL}(\mu_0^n, \mu_0) \xrightarrow{n \rightarrow \infty} 0, \text{ a.s.}, \quad (3.20)$$

where d_{BL} is the bounded Lipschitz distance. Under (3.20) and the dilution condition (3.10),

the global empirical measure μ^n verifies, for almost every realisation of the graph

$$\mathbf{E} \left[\sup_{s \leq T} d_{BL}(\mu_s^n, \mu_s) \right] \xrightarrow{n \rightarrow \infty} 0. \quad (3.21)$$

Supposing moreover that the initial condition is independent of the graph, if $np_n^3 \rightarrow \infty$ as $n \rightarrow \infty$, then for any fixed $l \geq 1$, the local empirical measure $\mu^{n,l}$ defined in (3.2) verifies, for almost every realisation of the graph

$$\mathbf{E} \left[\sup_{s \leq T} d_{BL}(\mu_s^{n,l}, \mu_s) \right] \xrightarrow{n \rightarrow \infty} 0. \quad (3.22)$$

Remark 3.1.6. Although stated in [L14] for simplicity for Erdős-Rényi graphs, the convergence (3.21) is really more general: the convergence (3.21) is valid for any (possibly deterministic) graph \mathcal{G}_n satisfying (3.18) i.e. $\|W^{\mathcal{G}_n} - \mathbf{1}\|_{\infty \rightarrow 1} \xrightarrow{n \rightarrow \infty} 0$. In case of Erdős-Rényi graphs, this latter condition is valid under (3.10). This indicates that the cut-norm is effectively the right topology on graphs to ensure convergence of mean-field models.

Contrary to the previous strategies based on couplings (for which we have discussed at length above on their potential limitations) the idea is here to proceed via a semimartingale argument. Introducing as for the proof of Proposition 2.1.1 the propagator $P_{s,t}f(\theta) := \mathbf{E}_B[f(\Phi_s^t(\theta))]$, where $t \mapsto \Phi_s^t(\theta)$ is the solution to the mean-field equation $d\theta_t = \int \Gamma(\theta_t, \theta) \mu_t(d\theta) dt + dB_t$ with $\Phi_s^s(\theta) = \theta$, one obtains with straightforward calculations, for all f regular

$$\begin{aligned} \mathbf{E} |\langle \mu_T^n - \mu_T, f \rangle| &\leq \mathbf{E} |\langle \mu_0^n - \mu_0, P_{0,T}f \rangle| + \mathbf{E} \left| \frac{1}{n} \sum_{k=1}^n \int_0^T \partial_\theta P_{t,T}f(\theta_t^{k,n}) dB_t^k \right| \\ &\quad + \int_0^T \mathbf{E} \left| \frac{1}{n^2} \sum_{i,j=1}^n \hat{\xi}_{i,j}^n \partial_\theta P_{t,T}f(\theta_t^{i,n}) \Gamma(\theta_t^{i,n}, \theta_t^{j,n}) \right| dt \\ &+ \int_0^T \mathbf{E} \left| \frac{1}{n} \sum_{i=1}^n \partial_\theta P_{t,T}f(\theta_t^{i,n}) \langle \Gamma(\theta_t^{i,n}, \cdot), \mu_t^n - \mu_t \rangle \right| dt := (A) + (B) + (C) + (D) \end{aligned} \quad (3.23)$$

Recall that $\hat{\xi}_{i,j}^n = \frac{\xi_{i,j}^n}{p_n} - 1$. Term (A) is controlled by the hypothesis (3.20) on the initial condition, (B) is a standard noise term, scaling as $n^{-1/2}$ and (D) is controlled by $d_{BL}(\mu^n, \mu)$ and will give rise to a Grönwall term.

Key idea: Grothendieck inequality

The main difficulty is to deal with the term (C) in (3.23): in view of the structure of (C), one would be inclined to use concentration arguments on the weighted sum of i.i.d. centered variables $\frac{1}{n^2} \sum_{i,j=1}^n \hat{\xi}_{i,j}^n u_{i,j}$ for any fixed bounded sequence of weights $(u_{i,j})$. This would work very well in the case $u_{i,j}$ would be *independent on the graph*, as one would be able to apply standard concentration inequalities. The problem is that $u_{i,j} := \partial_\theta P_{t,T}f(\theta_t^{i,n}) \Gamma(\theta_t^{i,n}, \theta_t^{j,n})$ depends in a highly nontrivial way on the graph sequence $\xi_{i,j}^n$ itself. The key argument is to make use of a powerful algebraic tool, the Grothendieck inequality (see e.g. [AN06, Ver18]) that can be stated as follows: consider an infinite dimensional Euclidian space with coordinates indexed by a space A , $l^2(A) = \{x = (x_\alpha)_{\alpha \in A} \in \mathbb{C}^A : \sum_{\alpha \in A} |x_\alpha|^2 < \infty\}$, endowed with the usual scalar product $\langle x, y \rangle_{l^2(A)} = \sum_{\alpha \in A} x_\alpha \bar{y}_\alpha$ and the associated norm $\|\cdot\|_{l^2(A)}$. Then there exists a universal constant \mathcal{K} such that for any finite scalar array (a_{jk}) ,

$$\begin{aligned} \sup \left\{ \left| \sum_{j,k} a_{jk} \langle x_j, y_k \rangle_{l^2(A)} \right| : x_j, y_k \in l^2(A), \|x_j\|_{l^2(A)} \leq 1, \|y_k\|_{l^2(A)} \leq 1 \right\} \\ \leq \mathcal{K} \sup \left\{ \left| \sum_{j,k} a_{jk} s_j t_k \right| : s_j = \pm 1, t_k = \pm 1 \right\}. \end{aligned} \quad (3.24)$$

In a word, one can replace $\langle x_j, y_k \rangle$ (with possible nontrivial dependence with the coefficients $a_{i,j}$ by arbitrary signs s_j, t_k . This is precisely our point here: we deduce^[1] that the term (C) is exactly controlled by $b_n^{\infty \rightarrow 1}$ defined in (3.18). Hence, the convergence holds for any graph \mathcal{G}_n satisfying $b_n^{\infty \rightarrow 1} \rightarrow 0$. In the present Erdős-Rényi case, all that remains is to show that $b_n^{\infty \rightarrow 1}$ is exactly of order $\frac{1}{\sqrt{np_n}} \rightarrow 0$, under assumption (3.10): indeed, by Bernstein's inequality and on a union bound, we obtain $\mathbb{P}\left(\sup_{s,t \in \{\pm 1\}^n} \left| \frac{1}{n^2} \sum_{i,j=1}^n \hat{\xi}_{i,j}^n s_i t_j \right| > t\right) \leq 2 \cdot 4^n \exp\left(-\frac{1}{2} \frac{t^2 n^2 p_n}{2 + \frac{t}{3}}\right)$. Thus, the choice $t = \frac{c}{\sqrt{np_n}}$ leads to

$$\mathbb{P}\left(b_n^{\infty \rightarrow 1} > \frac{c}{\sqrt{np_n}}\right) \leq 2 \cdot 4^n \exp\left(-\frac{1}{2} \frac{c^2 n}{2 + \frac{c}{3\sqrt{np_n}}}\right), \quad (3.25)$$

which is summable for $c = 3$ under the hypothesis $np_n \xrightarrow[n \rightarrow \infty]{} \infty$. The last difficulty is that one cannot put the supremum in $f \in BL$ inside the expectation in the lefthand side of (3.23). This is circumvented via a compactness argument.

Remark 3.1.7. For fixed signs s_i, t_j , the standard deviation of $\frac{1}{n^2} \sum_{i,j=1}^n \hat{\xi}_{i,j}^n s_i t_j$ is of order $\frac{1}{np_n^{1/2}}$: one easily sees with the same Bernstein/Borel-Cantelli argument that, for fixed s_i, t_j , for arbitrary $\varepsilon \in (0, 1/2)$, a.s., $n^{1-\varepsilon} p_n^{1/2-\varepsilon} \left| \frac{1}{n^2} \sum_{i,j=1}^n \hat{\xi}_{i,j}^n s_i t_j \right| \rightarrow 0$. We see here the tradeoff for the use of Grothendieck inequality: dealing with the supremum w.r.t. all 4^n possible signs requires to keep a factor n in the exponential bound in order to maintain summability. Hence, we loose some speed of convergence w.r.t. the standard deviations (what we need is really moderate deviations rather than Gaussian deviations from the CLT). In this sense, let us appreciate the present miracle: the Grothendieck inequality gives that $b_n^{\infty \rightarrow 1} \rightarrow 0$ under the optimal (3.10) only. The fact that one needs to assume the (certainly non optimal) condition $np_n^3 \rightarrow \infty$ for the local empirical measure (3.22) comes from the use of Grothendieck inequalities of higher order, for which the quality of speed of convergence given by Bernstein inequality degrades as $p_n \rightarrow 0$.

3.2 Homogeneous graphs: Central Limit Theorem

At this point, the reader may be convinced that we have addressed Question 1.2.5 as much as possible (at least in the case of homogeneous graphs) at the level of the law of large numbers concerning the empirical measure: the same universal macroscopic limit holds as long as the optimal condition (3.10) is met. We now address Question 1.2.5 at the level of fluctuations. This is the main purpose of [L14]. We are not only interested in the *global fluctuation process*

$$\eta^n := \sqrt{n}(\mu^n - \mu) \quad (3.26)$$

but also in the mutual convergence of *local fluctuation processes* around two vertices 1 and 2 (recall (3.2)), that is

$$\zeta_t^n := (\zeta_t^{n,1}, \zeta_t^{n,2}) := (\sqrt{np_n}(\mu_t^{n,1} - \mu_t), \sqrt{np_n}(\mu_t^{n,2} - \mu_t)) \quad (3.27)$$

One point is to understand how the limits of both η^n and ζ^n may or may not differ from their mean-field counterparts and whether or not their convergence depend on a specific realisation of the graph or on the graph structure itself (in particular the fact that the graph may be dense or diluted). As far as CLT on random graphs are concerned, we are only aware of [BBW19] where the global fluctuation of systems like (3.1) are analysed in an annealed framework (that is the behavior of (3.26) under the joint law $\mathbf{P} \otimes \mathbb{P}$), in the dense case $\liminf_{n \rightarrow \infty} p_n > 0$, with i.i.d. initial condition. The point of [L14] is precisely to address these three points: we work as much as possible in a *quenched case*, in a possibly *diluted regime* $p_n \rightarrow 0$ and we will drop the exchangeability assumption on the initial condition. Note also that [BBW19] consider the convergence of (3.26) in terms of finite dimensional marginals, whereas we address the full convergence of (3.26) as a process.

[1]. This requires to interpret (C) in terms of a scalar product: this is done in [L14] using Fourier series.

On the importance of the initial conditions

We place ourselves again for simplicity on the state space $\mathbb{T}^{[2]}$, with appropriate regularity and boundedness assumptions on both Γ and F . An important motivation of [L14] concerns the influence of the choice of the initial condition $(\theta_0^{1,n}, \dots, \theta_0^{n,n})$ in (3.1) on the nature of the fluctuations: the original CLT result in [FM97] relied heavily on the fact that the initial datum consisted of i.i.d. random variables. One significant improvement of [L14] is that we only suppose here that the initial conditions are chosen independently on the Brownian motions (B^1, \dots, B^n) (but not necessarily i.i.d. and they may depend on the graph), requiring only the convergence of the empirical measure at time $t = 0$ (3.20) as well as initial convergence and control on the fluctuation process η_0^n , i.e. $\sup_n \mathbb{E}_0(\|\eta_0^n\|_{-r}^{1+\alpha}) < +\infty$, \mathbb{P}_g -a.s. for some $\alpha \in (0, 1)$ and the convergence of η_0^n to some η_0 in some suitable Sobolev space $H^{-r}(\mathbb{T})$ for some $r > 0$. These assumptions not only include the case of i.i.d. initial conditions but also sequences satisfying suitable mixing conditions (including geometrically ergodic Markov processes, see [L14] for further details).

Semimartingale decompositions of the fluctuation processes

As before, start with the semimartingale decomposition for η^n : for any f regular,

$$\langle \eta_t^n, f \rangle = \langle \eta_0^n, f \rangle + \int_0^t \langle \eta_s^n, \mathcal{L}_s^{(1),n} f \rangle ds + H_t^{n,1}(f) + \langle W_t^n, f \rangle, \quad (3.28)$$

where $\mathcal{L}_s^{(1),n}$ is an explicit linear operator and W^n is a martingale. Decomposition (3.28) differs from the mean-field semimartingale decomposition only by the presence of the remaining term

$$H_t^{n,1}(f) := \int_0^t \frac{1}{n^{3/2}} \sum_{i,j=1}^n \hat{\xi}_{i,j}^n \Gamma(\theta_s^{i,n}, \theta_s^{j,n}) \partial_{\theta} f(\theta_s^{i,n}) ds \quad (3.29)$$

Recall that $\hat{\xi}_{i,j}^n := \frac{\xi_{i,j}^n}{p_n} - 1$, so that $H_t^{n,1}$ exactly captures the proximity of the graph \mathcal{G}_n w.r.t. K_n . This decomposition is not closed in η^n : $H^{n,1}$ is a functional of the second order fluctuation process

$$\hat{\eta}_n := \frac{1}{n^{3/2}} \sum_{i,j=1}^n \hat{\xi}_{i,j}^n \delta_{(\theta^{i,n}, \theta^{j,n})} \quad (3.30)$$

One could rightfully hope that $H^{n,1}$ would go to 0 as $n \rightarrow \infty$: applying again Grothendieck inequality (3.24), one obtains that $|H_t^{n,1}(f)| \leq K_{\Gamma, f} \sup_{s_i, t_j \in \{\pm 1\}} \frac{1}{n^{3/2}} \sum_{i,j=1}^n \hat{\xi}_{i,j}^n s_i t_j$. Bad news: concentration estimates (recall (3.25)) only tell us that this bound is of order $\frac{c}{\sqrt{p_n}}$: this at best tells us that H^n is bounded (and only in the dense case $p_n \rightarrow p > 0$!) but in any case, gives no information whatsoever on the fact that $H^{n,1}$ goes to 0. The idea is then to proceed further in the asymptotic expansion of μ^n and write also the semimartingale decomposition of $\hat{\eta}^n$: one has again

$$\hat{\eta}_t^n = \hat{\eta}_0^n + \int_0^t \mathcal{L}_s^{(2),n,*} \hat{\eta}_s^n ds + \sqrt{n} H_t^{n,2} + \hat{W}_t^n \quad (3.31)$$

in a suitable Sobolev space. Here, we are more optimistic w.r.t. the remainder term

$$\sqrt{n} H^{n,2}(g) := \sqrt{n} H^{n,\zeta}(g) + \sqrt{n} H^{n,\leftrightarrow}(g) \quad (3.32)$$

with

$$H^{n,\zeta}(g) = \int_0^t \frac{1}{n^3} \sum_{i,j,k=1}^n \hat{\xi}_{ij}^n \hat{\xi}_{ik}^n \partial_{\theta_1} g(\theta_s^{i,n}, \theta_s^{j,n}) \Gamma(\theta_s^{i,n}, \theta_s^{k,n}) ds \quad (3.33)$$

$$H^{n,\leftrightarrow}(g) = \int_0^t \frac{1}{n^3} \sum_{i,j,k=1}^n \hat{\xi}_{ij}^n \hat{\xi}_{jk}^n \partial_{\theta_2} g(\theta_s^{i,n}, \theta_s^{j,n}) \Gamma(\theta_s^{j,n}, \theta_s^{k,n}) ds, \quad (3.34)$$

as it concerns higher order functionals within the graph, so that one might expect them to be of smaller order in n . The main problem is that the standard Grothendieck inequality (3.24) is no longer applicable, as it concerns now trilinear functionals, not only bilinear scalar products.

[2]. Dealing with diffusions in \mathbb{R}^d would only entail technical complications, through especially the use of weighted Sobolev norms but the result would certainly be unchanged.

Generalised Grothendieck inequalities

Extensions of Grothendieck inequalities are known to fail in general when the scalar product in (3.24) is replaced by a bounded trilinear functional, see e.g. [Pis12]. However, in [Ble11], R. Blei describes a family of multilinear functionals for which this inequality remains valid. Consider a positive integer m and a sequence $\mathcal{U} = (S_1, \dots, S_N)$ of non empty sets that satisfy $\cup_{i=1}^N S_i = \{1, 2, \dots, m\}$. For $\alpha = (\alpha_j)_{1 \leq j \leq m} \in A^m$ define the projections $\pi_{S_i}(\alpha) = (\alpha_j)_{j \in S_i}$. Consider the functional $\nu_{\mathcal{U}}: l^2(A^{|S_1|}) \times \dots \times l^2(A^{|S_N|}) \rightarrow \mathbb{C}$ defined as $\nu_{\mathcal{U}}(x_1, \dots, x_N) = \sum_{\alpha \in A^m} x_1(\pi_{S_1}(\alpha)) \cdots x_N(\pi_{S_N}(\alpha))$. Denote, for $1 \leq j \leq m$, by k_j the incidence of j in the covering sequence \mathcal{U} , that is $k_j(\mathcal{U}) = |\{i : j \in S_i\}|$ and by $\mathcal{I}_{\mathcal{U}}$ the minimal incidence $\mathcal{I}_{\mathcal{U}} = \min \{k_j(\mathcal{U}) : j \in \{1, \dots, m\}\}$.

Theorem 3.2.1 (Theorem 11.11, Section 12.4 in [Ble11]). *Suppose that $\mathcal{I}_{\mathcal{U}} \geq 2$. Then there exists a positive constant $\mathcal{K}_{\mathcal{U}}$, depending only on the covering \mathcal{U} , such that for any finitely supported scalar n -array $a_{j_1 \dots j_N}$,*

$$\sup \left\{ \left| \sum_{j_1, \dots, j_N} a_{j_1 \dots j_N} \nu_{\mathcal{U}}(x_1, \dots, x_N) : \|x_1\|_{l^2(A^{|S_1|})} \leq 1, \dots, \|x_N\|_{l^2(A^{|S_N|})} \leq 1 \right| \right\} \leq \mathcal{K}_{\mathcal{U}} \sup \left\{ \left| \sum_{j_1, \dots, j_N} a_{j_1 \dots j_N} s_{1, j_1} \cdots s_{N, j_N} : s_{1, j_1} = \pm 1, \dots, s_{N, j_N} = \pm 1 \right| \right\}. \quad (3.35)$$

An additional technical difficulty concerning the estimation of (3.32) is that uniform estimates (in the test functions g) are required, as we consider $H^{n,2}$ in distribution spaces. Secondly, using (3.35) requires to identify (3.32) as proper multilinear functionals on appropriate $l^2(A)$. These two issues are treated in [L14] using Fourier series and Hilbertian techniques. Let us make a representative calculation here: let $(\Xi_{ijk}^{(n)})_{i,j,k=1, \dots, n}$ a field that depends on the graph \mathcal{G}_n (e.g. in (3.33), $\Xi_{ijk}^{(n)} = n^{-3} \hat{\xi}_{ij}^n \hat{\xi}_{jk}^n$) and consider the weighted statistics

$$S_n(f) := \int_0^T \sum_{i,j,k=1}^n \Xi_{ijk}^{(n)} \partial_{\theta_1} f(\theta_s^{i,n}, \theta_s^{j,n}) \Gamma(\theta_s^{i,n}, \theta_s^{k,n}) ds. \quad (3.36)$$

The point is to establish a bound on $S_n(f)$ that only depends on T , some norm of Γ and some Sobolev norm of f . Expand $\partial_{\theta_1} f$ and Γ into Fourier series in $L^2(\mathbb{T}^2)$ ($(e_a)_{a \in \mathbb{Z}}$ is the canonical basis of $L^2(\mathbb{T})$):

$$\partial_{\theta_1} f(\theta_1, \theta_2) = \sum_{a \in \mathbb{Z}} e_a(\theta_2) \int_{\mathbb{T}} \partial_{\theta_1} f(\theta_1, \theta) \bar{e}_a(\theta) d\theta \quad \text{and} \quad \Gamma(\theta_1, \theta_2) = \sum_{b \in \mathbb{Z}} e_b(\theta_2) \int_{\mathbb{T}} \Gamma(\theta_1, \theta) \bar{e}_b(\theta) d\theta.$$

For some $\delta > 0$ define the following sequences, indexed by $a, b \in \mathbb{Z}$: $x_{1,i}(a, b) := C_{\delta}^2 ((1+a^2)(1+b^2))^{1/4+\delta} \left(\int_{\mathbb{T}} \partial_{\theta_1} f(\theta_s^{i,n}, \theta) \bar{e}_a(\theta) d\theta \right) \left(\int_{\mathbb{T}} \Gamma(\theta_s^{i,n}, \theta) \bar{e}_b(\theta) d\theta \right)$, $x_{2,j}(a) := C_{\delta}^{-1} (1+a^2)^{-1/4-\delta} e_a(\theta_s^{j,n})$ and $x_{3,k}(b) := C_{\delta}^{-1} (1+b^2)^{-1/4-\delta} e_b(\theta_s^{k,n})$, where $C_{\delta} = (\sum_{a \in \mathbb{Z}} (1+a^2)^{-1/2-2\delta})^{1/2}$ for some $\delta > 0$. By construction

$$S_n(f) = \int_0^t \sum_{i,j,k=1}^n \Xi_{ijk}^{(n)} \sum_{a,b \in \mathbb{Z}} x_{1,i}(a, b) x_{2,j}(a) x_{3,k}(b) ds,$$

with $\|x_{2,j}\|_{\ell^2(\mathbb{Z})} = \|x_{3,k}\|_{\ell^2(\mathbb{Z})} = 1$ and a direct calculation show that

$$\|x_{1,i}\|_{\ell^2(\mathbb{T}^2)}^2 \leq C_{\delta}^4 C_{\Gamma, \delta} \sum_{a \in \mathbb{Z}} (1+a^2)^{1/2+2\delta} \left| \int_{\mathbb{T}} \partial_{\theta_1} f(\theta_s^{i,n}, \theta) \bar{e}_a(\theta) d\theta \right|^2 = C_{\delta}^4 C_{\Gamma, \delta} \left\| \partial_{\theta_1} f(\theta_s^{i,n}, \cdot) \right\|_{H^{1/2+2\delta}(\mathbb{d}\theta_2)}^2,$$

where the constant $C_{\Gamma, \delta}$ only depends on Γ and δ . Here, we use the standard fractional H^s -norm ($0 < s < 1$) on \mathbb{T} and $\left\| \partial_{\theta_1} f(\theta_s^{i,n}, \cdot) \right\|_{H^{1/2+2\delta}(\mathbb{d}\theta_2)}^2 = \left\| \partial_{\theta_2}^{1/2+2\delta} \partial_{\theta_1} f(\theta_s^{i,n}, \cdot) \right\|_{L^2(\mathbb{d}\theta_2)}^2 \leq \sup_{\theta \in \mathbb{T}} \left\| \partial_{\theta_2}^{1/2+2\delta} \partial_{\theta_1} f(\theta, \cdot) \right\|_{L^2(\mathbb{d}\theta_2)}^2$.

The last term is a mixed Sobolev-norm on test functions $(\theta_1, \theta_2) \mapsto f(\theta_1, \theta_2)$, that is $W^{1, \infty}$ in θ_1 and $H^{1/2+2\delta}$ in θ_2 . Choosing $\delta = 1/4$, one obtains finally, by Sobolev embeddings (see [L14,

App. A] for details), $\|x_{1,i}\|_{\ell^2(\mathbb{T}^2)}^2 \leq C_\Gamma \|f\|_{H^3(d\theta_1, d\theta_2)}^2 < \infty$. Hence, we can apply Theorem 3.2.1 with $A = \mathbb{Z}$, $m = 2$, $N = 3$, $\mathcal{U} = \{\{1, 2\}, \{1\}, \{2\}\}$ to obtain $|S_n(f)| \leq C \|f\|_{H^3(d\theta_1, d\theta_2)} \sup_{r,s,t \in \{\pm 1\}^n} \left| \sum_{i,j,k=1}^n \Xi_{ijk}^n r_i s_j t_k \right|$ so that taking the supremum in $t \in [0, T]$ and in $f \in H^3(\mathbb{T}^2)$, we finally obtain

$$\|S_n\|_{\mathcal{C}([0,T], H^{-3}(\mathbb{T}^2))} \leq C \sup_{r,s,t \in \{\pm 1\}^n} \left| \sum_{i,j,k=1}^n \Xi_{ijk}^n r_i s_j t_k \right| \quad (3.37)$$

Applying this calculation to (3.33) (the same holds for (3.34)) and using standard concentration results [Tao12] on the righthand side of (3.37) gives that $\left\| \sqrt{n} H^n \mathring{\zeta} \right\|_{-3}$ is of order $\frac{1}{\sqrt{np_n^2}}$ which goes to 0 provided that

$$np_n^4 \xrightarrow{n \rightarrow \infty} 0. \quad (3.38)$$

The main conclusion is that one can close the asymptotic expansion of the fluctuation process at the level of (3.30): the global fluctuations of the system are described only in terms of $(\eta^n, \hat{\eta}^n)$.

Remark 3.2.2 (On extensions to \mathbb{R}^d). In [L14], the state space is assumed for simplicity to be \mathbb{T} . This has two simplifying advantages: (i) straightforward boundedness assumptions on the coefficients of (3.1) and (ii) the use of Fourier series as above. Extending the present results to \mathbb{R}^d is however possible: the point (i) is easily circumvented by the use of weighted Sobolev spaces as in [Mit85, FM97]. Concerning point (ii), a possible approach in \mathbb{R}^d to a similar Fourier expansion as above would be to assume that the interaction kernel $(\theta, \tilde{\theta}) \mapsto \Gamma(\theta, \tilde{\theta})$ is in L^1 Fourier class (see e.g. [OR19, Ass. 4] for more details), that is to assume the existence of a complex measure m over $(\mathbb{R}^d)^2$ such that $\Gamma(\theta, \tilde{\theta}) = \int_{(\mathbb{R}^d)^2} \exp(2\pi i \langle (\theta, \tilde{\theta}), (u, v) \rangle) m(du, dv)$. Developpements similar as before follow easily from this identity.

The global fluctuation result

This use of generalised Grothendieck inequalities coupled with more standard Hilbertian techniques (similar to the ones used for Theorem 2.4.6) lead to the first main result of [L14]:

Theorem 3.2.3 (General global fluctuations, Theorem 2.7 in [L14]). *Under the previous hypotheses and under the dilution condition (3.38), for some appropriate Sobolev spaces of distribution H_1 and H_2 , for almost every realisation of the graph \mathcal{G}_n , $(\eta^n, \hat{\eta}^n)$ converges in law in $\mathcal{C}([0, T], H_1 \otimes H_2)$ to $(\eta, \hat{\eta})$ to the well-posed coupled system*

$$\begin{cases} \eta_t = \eta_0 + \int_0^t \mathcal{L}_s^{(1),*} \eta_s ds + \int_0^t \Theta^* \hat{\eta}_s ds + W_t, \\ \hat{\eta}_t = \hat{\eta}_0 + \int_0^t \mathcal{L}_s^{(2),*} \hat{\eta}_s ds \end{cases} \quad (3.39)$$

where $\mathcal{L}_s^{(1)}, \mathcal{L}_s^{(2)}, \Theta$ are explicit linear operators and $(W_t)_{t \in [0, T]}$ is a Gaussian process with explicit covariance, independent of $(\eta_0, \hat{\eta}_0)$.

We stress again that we consider much more general initial conditions than i.i.d.: we only impose the convergence (3.20) and that $(\eta_0^n, \hat{\eta}_0^n) \rightarrow (\eta_0, \hat{\eta}_0)$. The proof here diverges significantly from [FM97]: a crucial step in the tightness estimate (see [FM97, Prop. 3.5]) is to derive uniform bounds on moments of the fluctuation process of the form $\sup_n \sup_{t \leq T} \mathbf{E}(\|\eta_t^n\|_H^2) < \infty$, for an appropriate Hilbert space H . The proof of [FM97, Prop. 3.5] relies heavily on the exchangeability of the particle system when the graph of interaction is complete, which is not a path one can follow here. We circumvent this difficulty by rewriting (3.28) and (3.31) into equivalent weak-mild formulations: denoting by $S_t := e^{t\Delta}$ the semigroup of the Laplacian on \mathbb{T} , we have, for all test functions h

$$\begin{cases} \langle \eta_t^n, h \rangle = \langle \eta_0^n, S_t h \rangle + \int_0^t \langle \eta_s^n, (\Gamma * \mu_s^n) \partial_\theta S_{t-s} h \rangle ds + \int_0^t \langle \mu_s, (\Gamma * \eta_s^n) \partial_\theta S_{t-s} h \rangle ds + w_t^n(h) \\ \quad + \int_0^t \frac{1}{n^{3/2}} \sum_{i,j=1}^n \hat{\xi}_{ij}^{(n)} \langle \delta_{\theta_s^{i,n}}, (\Gamma * \delta_{\theta_s^{j,n}}) \partial_\theta S_{t-s} h \rangle ds, \\ \langle \hat{\eta}_t^n, h \rangle = \langle \hat{\eta}_0^n, S_t h \rangle + \int_0^t \langle \hat{\eta}_s^n, \Lambda_s^n \cdot [\nabla S_{t-s} h] \rangle ds + \hat{w}_t^n(h) \\ \quad + \int_0^t \frac{1}{n^{3/2}} \sum_{i,j=1}^n \hat{\xi}_{ij}^{(n)} \langle \delta_{\theta_s^{i,n}} \otimes \delta_{\theta_s^{j,n}}, \hat{\Lambda}_{s,ij}^n \cdot \nabla S_{t-s} h \rangle ds, \end{cases} \quad (3.40)$$

where

$$\Lambda_s^n(\theta_1, \theta_2) := (\langle \mu_s^n(d\theta'), \Gamma(\theta_1, \theta') \rangle, \langle \mu_s^n(d\theta'), \Gamma(\theta_2, \theta') \rangle) \in \mathbb{R}^2, \quad (3.41)$$

$$\hat{\Lambda}_{s,ij}^n(\theta_1, \theta_2) := (\langle \hat{\mu}_s^{n,i}(d\theta'), \Gamma(\theta_1, \theta') \rangle, \langle \hat{\mu}_s^{n,j}(d\theta'), \Gamma(\theta_2, \theta') \rangle) \in \mathbb{R}^2, \quad (3.42)$$

and w_t^n and \hat{w}_t^n are explicit martingale terms. In (3.40), the bracket $\langle \cdot, \cdot \rangle$ stands for the duality H^{-r}, H^r for some appropriate index r . One now can take advantage of the regularising properties of the heat kernel S_t : one has that for $h \in H^r(\mathbb{T}^d)$, $\|S_{t-s}h\|_k \leq C \left(1 + \frac{1}{(t-s)^{\frac{k-r}{2}}}\right) \|h\|_r$. Plugging this kind of estimate into the first line of (3.40) gives for $\alpha \in [0, 1)$,

$$\mathbf{E}(\|\eta_t^n\|_{-r}^{1+\alpha}) \leq \|\eta_0^n\|_{-r}^{1+\alpha} + C \int_0^t \left(1 + \frac{1}{(t-s)^{\frac{1+\alpha}{2}}}\right) \mathbf{E}(\|\eta_s^n\|_{-r}^{1+\alpha}) ds + \mathbf{E}(\|w_t^n\|_{-r}^{1+\alpha}) + C \sup_{t \in [0, T]} \mathbf{E}(\|\hat{\eta}_t^n\|_{-r}^{1+\alpha}) \quad (3.43)$$

Two conclusions from (3.43): firstly, η^n is controlled in terms of the auxiliary process $\hat{\eta}^n$. Applying again the same techniques to the second line of (3.40) (and using controls on the remaining terms similar to (3.37)) gives a uniform control on this auxiliary process of the form $\sup_n \sup_{t \leq T} \mathbf{E}(\|\hat{\eta}_t^n\|_{-r}^{1+\alpha}) < \infty$ for some appropriate r . In return, this uniform control on $\hat{\eta}^n$ propagates back to η^n via (3.43). The second and main observation from (3.43) is that the singularity of the term $\frac{1}{(t-s)^{\frac{1+\alpha}{2}}}$ requires to have $\alpha < 1$ (and one then concludes from a Grönwall-type argument a similar control on η^n). In particular, one cannot reach $\alpha = 1$, as it was originally the case in [FM97] (but this $1 + \alpha < 2$ is actually sufficient for the remaining of the tightness argument). Dealing with this singular kernel is the price to pay to use the regularity of the heat kernel (and hence to bypass the loss of exchangeability of the system and being able to consider very general initial conditions, not necessarily i.i.d.).

Universality of mean-field fluctuations

A crucial observation from (3.39) is that one retrieves the universal mean-field fluctuation SPDE for η_t (obtained in [FM97] in the pure mean-field case $p_n \equiv 1$ under i.i.d. initial conditions)

$$\eta_t = \eta_0 + \int_0^t \mathcal{L}_s^{(1),*} \eta_s ds + W_t, \quad (3.44)$$

if and only if the auxiliary term $\int_0^t \Theta^* \hat{\eta}_s ds$ vanishes (and it is the case if $\hat{\eta} \equiv 0$). Since the dynamics of $\hat{\eta}$ in (3.39) is linear and deterministic, we have that $\hat{\eta}_t \equiv 0$ if $\hat{\eta}_0 \equiv 0$. An important point of [L14] (see Proposition 2.11) is to note that this is indeed the case under the following condition

Condition 3.2.4. *Suppose that the initial condition $(\theta_0^{1,n}, \dots, \theta_0^{n,n})$ is chosen independently of the graph (but not necessarily i.i.d.).*

Based on the previous observation, we have the important corollary:

Theorem 3.2.5 (Universality of mean-field fluctuations, Th. 2.8, Prop. 2.11 in [L14]). *Under the assumptions of Theorem 3.2.3, suppose in addition Condition 3.2.4. Then, (η^n) converges in law as $n \rightarrow \infty$ to η , unique solution to the mean-field fluctuation SPDE (3.44) with η_0 independent of W .*

Theorem 3.2.5 is a first positive answer to Question 1.2.5 at the level of the CLT: for a generic Erdős-Rényi graphs of parameter p_n (satisfying the dilution condition (3.38)), the system (3.1) conserves the same global fluctuations as in the mean-field case, provided one chooses the initial condition independently from the graph. Note that we do not expect the dilution condition (3.38) to be sharp: the bound $p_n \gg \frac{1}{n^{1/4}}$ is required by the use of concentration inequalities for Grothendieck estimates that may not be optimal in the diluted regime $p_n \rightarrow 0$ (this is due to the lack of sub-Gaussianity of $\hat{\xi}_{i,j}^n = \frac{\xi_{i,j}^n}{p_n} - 1$ as $p_n \rightarrow 0$). Any improvement in these bounds would necessarily improve condition (3.38). Note nonetheless that there is good confidence that these

universal fluctuations should hold beyond (3.38): an indication in this direction concerns the large deviations principle considered by Coppini, Dietert and Giacomin: it is proven in [CDG20, Th. 2.2] that under the condition $\liminf_{n \rightarrow \infty} \frac{pn^n}{\log(n)} > 0$, the empirical measure of (3.1) satisfies the same LDP than in the mean-field framework. But in case the rate function has sufficient convexity properties around its unique zero, it is possible to derive a weak form of CLT (in the sense of convergence of finite dimensional marginals) from the LDP estimate. We refer to [Bol86] and [DdH96] where this procedure has been carried out for Kuramoto-type oscillators. Hence, the same LDP indicates the possibility of the same CLT beyond (3.38).

An example of non-universal fluctuations

An interesting question is what happens when one relaxes Condition 3.2.4, that is when one considers initial conditions that depend on the graph. We describe in [L14] an example of such initial conditions, such that $\hat{\eta}_0^n$ has a non zero limit, so that the limit fluctuations are non-universal, i.e. completely described by (3.39) and no longer by the mean field fluctuations (3.44) (so that, in a sense, Condition 3.2.4 is optimal w.r.t. mean-field universality):

Proposition 3.2.6 (Proposition 2.14 in [L14]). *Take $\Gamma(\theta, \theta') = -\sin(\theta - \theta')$. For any graph sequence $(\xi_{i,j}^n)$, there exists a choice of initial conditions $(\theta_0^{1,n}, \dots, \theta_0^{n,n})$ such that $(\eta_0^n, \hat{\eta}_0^n)$ converges in law (w.r.t. the annealed law $\mathbb{P} \otimes \mathbf{P}$) to $(\eta_0, \hat{\eta}_0)$ with $\eta_0 = Z_1 \delta_0 + Z_2 \delta_{\frac{\pi}{2}}$ (where $(Z_1, Z_2) \sim \mathcal{N}(0, K)$ with $K = \begin{pmatrix} \frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{1}{4} \end{pmatrix}$) and $\hat{\eta}_0 = \frac{1}{6\sqrt{\pi}} \left(-\delta_{(0,0)} + 2\delta_{(\frac{\pi}{2}, 0)} - \delta_{(\frac{\pi}{2}, \frac{\pi}{2})} \right)$. In particular, $\Gamma * \hat{\eta}_0^n$ converges to $-\frac{1}{3\sqrt{2\pi}} \delta_{\frac{\pi}{2}} \neq 0$, so that the limiting process η_t is governed by (3.39) and not by the universal mean-field SPDE (3.44).*

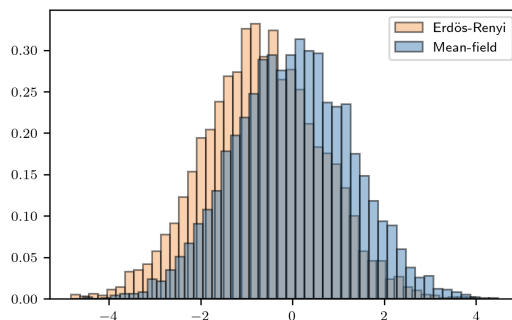


Figure 3.1 – Histograms representing 5000 realizations of $\sqrt{n}(\psi_1^n - \psi_1)$, where $r_t^n e^{i\psi_t^n} = \langle \mu_t^n(d\theta), e^{i\theta} \rangle$ and $r_t e^{i\psi_t} = \langle \mu_t(d\theta), e^{i\theta} \rangle$, for the choice of interaction kernel $\Gamma(\theta, \theta') = -K \sin(\theta - \theta')$ with $K = 2$. For the blue histogram the interaction is of mean-field type with i.i.d. initial condition of distribution $\frac{1}{2}\delta_0 + \frac{1}{2}\delta_{\frac{\pi}{2}}$, while for the brown one it is of symmetric Erdős-Renyi type with $p = 0.5$ and initial condition as described by Proposition 3.2.6. We observe a dephasing at time $t = 1$ at the level of fluctuations, induced by the graph-dependent initial condition.

Local fluctuation results

The same techniques can be applied to the local random field (3.27). As global fluctuations compete with local fluctuations, the main result concerns the joint convergence of the joint fluctuation process $(\zeta^{n,1}, \zeta^{n,2}, \eta^n)$. For simplicity, we place ourselves in the case of i.i.d. initial conditions, independent on the graph.

Theorem 3.2.7 (Local fluctuations, Theorem 2.9 in [L14]). *Suppose that $(\theta_0^{1,n}, \dots, \theta_0^{n,n})$ are i.i.d. random variables with law μ_0 , independent from the graph. Suppose that $\liminf_n np_n^5 = \infty$ and denote by $p := \lim_{n \rightarrow \infty} p_n \in [0, 1]$. Then, there exists some Sobolev space of distribution H_1*

such that for almost every realisation of the graph \mathcal{G}_n , the joint fluctuation process $(\zeta^{n,1}, \zeta^{n,2}, \eta^n)$ converges as $n \rightarrow \infty$ in $\mathcal{C}([0, T], (H_1)^3)$ to (ζ^1, ζ^2, η) solution to the well-posed system

$$\begin{cases} \zeta_t^l = \zeta_0^l + \int_0^t \mathcal{U}_s^* \zeta_s^l ds + \sqrt{p} \int_0^t \mathcal{V}_s^* \eta_s ds + W_t^l, & l = 1, 2, \\ \eta_t = \eta_0 + \int_0^t \mathcal{L}_s^{(1),*} \eta_s ds + W_t. \end{cases} \quad (3.45)$$

where $\mathcal{U}_s, \mathcal{V}_s$ are explicit differential operators, $(\zeta_0^1, \zeta_0^2, \eta_0)$ and (W_t^1, W_t^2, W_t) Gaussian processes with explicit covariance, the initial condition $(\zeta_0^1, \zeta_0^2, \eta_0)$ and the noise (W^1, W^2, W) being independent.

We refer to [L14] for an explicit description of the covariance structure of both the noise and initial condition in (3.45). What we only need to retain here is that there is a significant breakdown of the mean-field universality at the level of local fluctuations: in the *dense case* $p_n \rightarrow p > 0$, ζ^1 and ζ^2 are correlated through a nontrivial correlation of their noise and initial conditions, and through the coupling of the global fluctuation process η (this is in accordance with the fully coupled case $p \equiv 1$ where $\eta = \zeta^1 = \zeta^2$). However, in the *diluted case* $p_n \rightarrow p = 0$, the processes (ζ^1, ζ^2, η) become mutually independent.

3.3 Inhomogeneous graphs

So far, we have addressed the case of homogeneous random graphs \mathcal{G}_n . One crucial observation is that most of the techniques used above in the homogeneous case do not actually require the $\xi_{i,j}^n$ to be identically distributed. This leads to the following generalisation of the law of large numbers result to inhomogeneous graphs, carried out in [L10]. This analysis of mean-field particle systems interacting on inhomogeneous random graphs originates from a series of paper from Medvedev and coauthors [Med14a, KVM17, KVM18, CM19, Med14b] on macroscopic limits for Kuramoto-type models, in the deterministic case $\sigma = 0$. In addition to the fact that we consider here more general dynamics, the crucial point is the presence of noise in (3.1) that changes considerably the analysis (in particular, the fixed-point argument [Neu84] used in [KVM18, CM19] for the convergence of the empirical measure in the deterministic case does not seem to generalize easily to the case $\sigma \neq 0$). The work that is closest to the present analysis is the recent work [OR19] where annealed large deviations estimates are given in the case of bounded graphons. Some of the results of [L10] have been partially and independently reproduced in a subsequent [BCW20], in an annealed regime. Among results that are closest to [L10], one should cite [OR19], for an annealed LLN and LDP in a case of bounded kernels W , [BCN20] for an annealed LLN in an exchangeable framework with possibly random graphons (still in a bounded case). Contrary to [OR19], we do not address large deviation estimates here. Let us finally mention the recent works [JPS21, KX22] where similar law of large numbers is established for sparse graphs represented by measure-valued graphons.

3.3.1 A generic law of large numbers for the empirical measure

An inhomogeneous graph of interaction

Consider again the system (3.1) (with standard hypotheses on F and Γ) when now the $\xi_{i,j}^n$ are sampled in a inhomogeneous way: take for simplicity the compact $I := [0, 1]$ and associate to each vertex $i = 1, \dots, n$ a position variable $x_i^n := \frac{i}{n} \in I$ encoding some local inhomogeneity for the vertex i in the graph \mathcal{G}_n [3]. Then, introduce a kernel $W_n : I^2 \rightarrow [0, 1]$ such that $W_n(x_i^n, x_j^n) \in [0, 1]$ represents the probability of the presence of the edge $i \rightarrow j$ in the graph \mathcal{G}_n : on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$, $(\xi_{i,j}^n)_{i,j=1, \dots, n; n \geq 1}$ is a family of independent Bernoulli random variables with parameter $W_n(x_i^n, x_j^n)$. Contrary to the homogeneous case, the dilution

[3]. the case of a general set I with generic, possibly random, positions x_i^n can be also considered, provided one supposes the convergence of the empirical measure of the positions, see [L10] for details.

parameter $\kappa_i^n > 0$ in (3.1) (which compensates for the possible local sparsity of the graph \mathcal{G}_n around vertex i) is intrinsically inhomogeneous: κ_i^n may actually depend on the whole sequence of positions in the graph \mathcal{G}_n : $\kappa_i^n = \kappa_i^n(x_1^n, \dots, x_n^n)$. We assume for simplicity that the initial condition $(\theta_0^{1,n}, \dots, \theta_0^{n,n})$ in (3.1) are independent with law $\theta_0^{i,n} \sim \nu_0^{x_i^n}$, with appropriate moment conditions. Note here that we allow a priori the initial law of the particles to depend on their positions $x_k = x_k^n$. An intermediate particle system that naturally approximates (3.1) is

$$d\theta_t^{i,n} = F(\theta_t^{i,n})dt + \frac{\kappa_i^n}{n} \sum_{j=1}^n W_n(x_i^n, x_j^n) \Gamma(\theta_t^{i,n}, \theta_t^{j,n}) dt + \sigma \cdot dB_t^i, \quad i = 1, \dots, n, \quad t \in [0, T]. \quad (3.46)$$

Here, the sequence $(\kappa_i^n W_n(x_i^n, x_j^n))_{i,j=1, \dots, n}$ encodes for a graph $\bar{\mathcal{G}}_n$ which is the *annealed* (as $\xi_{i,j}^n$ is replaced by its expectation $W_n(x_i, x_j)$) and *weighted* (as each edge $i \rightarrow j$ carries the weight $\beta_{i,j} = \kappa_i^n$) version of the original graph \mathcal{G}_n .

Assumptions on the graph $\bar{\mathcal{G}}_n$ and its limit W

The second main ingredient is a notion of proximity of the graph $\bar{\mathcal{G}}_n$ to some macroscopic kernel $W : I^2 \rightarrow [0, +\infty)$ (in the Erdős-Rényi case, this is simply $W \equiv 1$). Suppose the existence of $\kappa_n \geq 1$ and $w_n \in (0, 1]$ such that $\kappa_\infty^n(\underline{x}) := \max_{i=1, \dots, n} (\kappa_i^n(\underline{x})) \leq \kappa_n$ and $\max_{i,j=1, \dots, n} (W_n(x_i, x_j)) \leq w_n$ satisfying, as $n \rightarrow \infty$,

$$\kappa_n^2 w_n = o\left(\frac{n}{\ln n}\right), \quad \text{as } n \rightarrow \infty. \quad (3.47)$$

Assumption (3.47) is the exact counterpart of the dilution condition (3.9) in our inhomogeneous framework (which one recovers in Example 3.1.2, taking $w_n = 1$ and $\kappa_i^n \equiv \kappa_n = \frac{1}{p_n}$ for all $i = 1, \dots, n$). The main assumption we ask is the existence of some kernel $W : [0, 1]^2 \rightarrow [0, +\infty)$ verifying (recall the definition of $\|\cdot\|_{\infty \rightarrow \infty}$ in (3.16))

$$\left\| W^{\bar{\mathcal{G}}_n} - W \right\|_{\infty \rightarrow \infty} \xrightarrow{n \rightarrow \infty} 0. \quad (3.48)$$

Here, for any weighted graph \mathcal{G} (with weights $\beta_{i,j}$ on each edge $i, j = 1, \dots, n$), $W^{\mathcal{G}}$ is the step function $W^{\mathcal{G}}(x, y) = \sum_{i,j=1}^n \beta_{i,j} \mathbf{1}_{(x,y) \in I_i \times I_j}$ for $I_k = \left[\frac{k-1}{n}, \frac{k}{n}\right[$. We refer to § 3.3.3 for further comments about (3.48), regarding especially the observations made in § 3.1.3 in the homogeneous case $W \equiv 1$.

We need also some regularity estimates on W . We leave aside all the (technical but natural) assumptions on W and only mention here that these hypotheses do not directly concern W itself but rather the degree functionals $x \mapsto \int W(x, y)^r dy$ for some $r \geq 1$. In particular, one crucially needs the following uniform control

$$\sup_{x \in I} D^{\text{in}}(x) := \sup_{x \in I} \int W(x, y) dy < \infty. \quad (3.49)$$

i.e. that in the limit $n \rightarrow \infty$, the indegree of each node $x \in I$ (i.e. the global incoming contribution onto x of all other vertices y in the macroscopic graph W) remains uniformly bounded [DDS20]^[4]. A sufficient condition for (3.49) is of course that W is bounded itself (assumption that is commonly met in the literature [LS06]), but there are interesting examples of unbounded kernels W satisfying (3.49). Note that we do not suppose any symmetry of the kernels W and W_n , nor that W and W_n are functions of the distance $x - y$ (this is a natural hypothesis if one thinks of applications in neuroscience, as the mutual influence between neuron i on neuron j need not be symmetric). Interestingly, this set-up unifies the framework of diffusions interacting on a random graph with spatially extended McKean-Vlasov diffusions considered in § 2.4: the natural coupling of the original particle system (3.1) is then described by independent copies $(\theta^{x_1}, \dots, \theta^{x_n})$ (with same initial conditions and noise) of the nonlinear process

$$d\bar{\theta}_t^x = F(\bar{\theta}_t^x)dt + \int W(x, y) \Gamma(\bar{\theta}_t^x, \bar{\theta}) \mu_t^y(d\bar{\theta}) dy dt + \sigma dB_t, \quad 0 \leq t \leq T. \quad (3.50)$$

[4]. actually one requires a further bound on the moment of order 2, i.e. $\|W_2\|_\infty < \infty$, but this is mostly technical.

where, for fixed (x, t) , $\mu_t^x(d\theta)$ is the law of $\bar{\theta}_t^x$. The joint law $\mu(d\theta, dx) = \mu^x(d\theta)dx$ of $(\bar{\theta}^x, x)$ solves the nonlinear Fokker-Planck equation ($\varphi(\theta, x)$ being a regular test function)

$$\begin{aligned} \langle \mu_t, \varphi \rangle &= \langle \mu_0, \varphi \rangle + \int_0^t \left\langle \mu_s, \frac{1}{2} \nabla_\theta (\sigma \sigma^\dagger \nabla_\theta \varphi) + \nabla_\theta \varphi \cdot F \right\rangle ds \\ &\quad + \int_0^t \left\langle \mu_s(d\theta, dx), \nabla_\theta \varphi(\theta, x) \cdot \int W(x, y) \Gamma(\theta, \tilde{\theta}) \mu_s(d\tilde{\theta}, dy) \right\rangle ds, \end{aligned} \quad (3.51)$$

that is (writing formally $\mu_t(d\theta, dx) = p_t(\theta, x)d\theta dx$) the weak formulation of

$$\partial_t p_t = \frac{1}{2} \nabla_\theta (\sigma \sigma^\dagger \nabla_\theta p_t) - \nabla_\theta \left(p_t \left(F + \int \Gamma(\cdot, \theta') W(\cdot, y) p_t(\theta', y) d\theta' dy \right) \right). \quad (3.52)$$

The main convergence result

The main result of [L10] is

Theorem 3.3.1 (Theorems 2.13 and 2.15 in [L10]). *Fix $T \geq 0$. Under the previous hypotheses on $F, \Gamma, \bar{\mathcal{G}}_n$ and W , we have that*

$$\sup_{i=1, \dots, n} \mathbf{E} \left[\sup_{s \leq T} |\theta_s^{i,n} - \bar{\theta}_s^{x_i}|^2 \right] \rightarrow 0 \text{ as } n \rightarrow \infty, \quad (3.53)$$

for almost every realization of the graph \mathcal{G}_n and one has the convergence of the empirical measure towards the solution to (3.51)

$$\mathbf{E} \left[\sup_{s \leq T} d_{BL}(\mu_s^n, \mu_s) \right] \rightarrow 0 \text{ as } n \rightarrow \infty. \quad (3.54)$$

Proof of Theorem 3.3.1 relies on a similar coupling as for Theorem 3.1.1, with the substantial nontrivial technical complexity due to the presence of spatial inhomogeneities in (3.46). The bound on $\sup_{i=1, \dots, n} \mathbf{E} \left[\sup_{s \leq T} |\theta_s^{i,n} - \bar{\theta}_s^{x_i}|^2 \right]$ is now expressed in terms of

$$b_n(\mathcal{G}_n, \bar{\mathcal{G}}_n) = \sup_{i=1, \dots, n} \left| \frac{1}{n} \sum_{j=1}^n (\kappa_i^n \xi_{i,j}^n - W_n(x_i, x_j)) \right|$$

(term that can be seen as an inhomogeneous version of b_n in (3.6) and is treated by concentration arguments) as well as terms involving $\|W^{\bar{\mathcal{G}}_n} - W\|_{\infty \rightarrow \infty}$. Note that the convergence to 0 of the previous term captures in particular a notion a regularity of the kernel W (this consideration is obviously totally absent in the Erdős-Rényi case as $W \equiv 1$). An important point of the proof is to obtain precise estimates on the regularity of $x \mapsto \mu^x$ in (3.51). For this, one needs to understand the kind of regularity that is required on the kernel W itself. This strongly depends on the nature of the spatial sequence (x_1, \dots, x_n) one considers: taking regular positions $x_i = \frac{i}{n}$ basically requires Hölder regularity on the degree $x \mapsto \int W(x, y) dy$ whereas choosing random i.i.d. positions only requires L^p estimates on W (this point is detailed at length in [L10, § 3]).

3.3.2 Examples of graphs satisfying Theorem 3.3.1.

The present examples are directly inspired by the formalism of W -random graphs, introduced in [LS06, BCL⁺08, BCL⁺12, BCCZ19, BCCZ18] and used in [Med14a, KVM17, KVM18, CM19, Med14b] (and references therein). The probability field $W_n(x, y)$ is directly constructed upon a predetermined nonnegative deterministic kernel, possibly unbounded, $\mathcal{P}(x, y)$.

Definition 3.3.2 (Generic random graph with graphon \mathcal{P}). *For fixed $n \geq 1$ and a given positive measurable kernel $(x, y) \mapsto \mathcal{P}(x, y)$ on I^2 , we define*

$$W_n(x, y) := \rho_n \min \left(\frac{1}{\rho_n}, \mathcal{P}(x, y) \right), \quad (3.55)$$

where $\rho_n \in [0, 1]$.

One important aim of [LS06, BCL⁺08, BCL⁺12, BCCZ19, BCCZ18] (and references therein) is precisely to prove that \mathcal{G}_n given by Definition 3.3.2 converges to \mathcal{P} for the cut-norm (3.19). In this context, one generally distinguishes between *bounded graphons* \mathcal{P} [LS06, BCL⁺08, BCL⁺12] and *unbounded graphons* (a typical hypothesis being that $\mathcal{P} \in L^p(I^2)$ for some $p \geq 1$ see [BCCZ19, BCCZ18] and references therein). Note that when the graphon \mathcal{P} is bounded (and up to the change $\rho_n \leftrightarrow \frac{\rho_n}{\|\mathcal{P}\|_\infty}$, one can always suppose that $\|\mathcal{P}\|_\infty = 1$), (3.55) boils down to

$$W_n(x, y) = \rho_n \mathcal{P}(x, y), \quad x, y \in I. \quad (3.56)$$

When $\rho_n = 1$, we are dealing with dense graphs, whereas in the case $\rho_n \xrightarrow{n \rightarrow \infty} 0$, we consider diluted graphs. (3.56) is an inhomogeneous version of the Erdős-Rényi case (Example 3.1.2 is $\mathcal{P}(x, y) \equiv 1$). When \mathcal{P} is not bounded, one usually assumes in (3.55) that $\rho_n \rightarrow 0$ and $n\rho_n \rightarrow \infty$, as $n \rightarrow \infty$. Here the distinction is not really on the boundedness of \mathcal{P} in (3.55), but rather a control on the degree of \mathcal{P} , that is either uniformly bounded $\sup_x \int \mathcal{P}(x, y) dy < +\infty$ or not: $\sup_x \int \mathcal{P}(x, y) dy = +\infty$.

The case of uniformly bounded degrees

Assumption 3.3.3. *Suppose here that*

$$\sup_{x \in I} \int \mathcal{P}(x, y) dy < +\infty. \quad (3.57)$$

Then the result is the following

Proposition 3.3.4 (§ 3.3 of [L10]). *Assume here that W_n is given by (3.55) for some \mathcal{P} verifying Assumption 3.3.3. Choose a renormalization that is uniform on the nodes $i = 1, \dots, n$, i.e.*

$$\kappa_i^n = \kappa_n := \frac{1}{\rho_n}, \quad i = 1, \dots, n \text{ and } w_n = \rho_n, \quad (3.58)$$

satisfying $\kappa_n = \frac{1}{\rho_n} = o\left(\frac{n}{\ln n}\right)$, as $n \rightarrow \infty$. Then, under appropriate regularity hypotheses on \mathcal{P} and for an appropriate choice of ρ_n , the limit for $\bar{\mathcal{G}}_n$ in the sense of (3.48) is simply given by $W := \mathcal{P}$ itself (see Proposition 3.6 in [L10]). In particular Theorem 3.3.1 is true.

Example 3.3.5. *Interesting choices include $\mathcal{P}(x, y) = 1 - \max(x, y)$ or $\mathcal{P}(x, y) = f(x)g(y)$ (Expected Degree Distribution [CL02] with $f \geq 0$ bounded and g integrable, see [BCCZ18, BCL⁺11] for many other interesting examples). Our hypotheses are sufficiently general to capture some interesting cases where \mathcal{P} is not continuous, such as the P -nearest neighbor model (Example 2.4.1) or singular polynomial interaction $\mathcal{P}(x, y) := \frac{1}{|x-y|^\alpha}$ on $I = [0, 1]$ (for $0 < \alpha < \frac{1}{2}$) as in Example 2.4.2. They also include the case of kernels with values in $\{0, 1\}$ (with $\rho_n = 1$) (this case corresponds to deterministic graphs, see [Med14a], § 4).*

Let us mention another example of unbounded \mathcal{P} such that (3.57) is nonetheless verified:

Example 3.3.6 ([KVM17]). *Let W_n be given by (3.55) where*

$$\mathcal{P}(x, y) := (1 - \alpha)y^{-\alpha}, \quad \alpha \in \left[0, \frac{1}{2}\right), \quad x, y \in I = [0, 1], \quad (3.59)$$

and κ_i^n is given by (3.58) and $\rho_n = n^{-\delta}$ for some $\alpha < \delta < \frac{1}{2}$. Then, $\bar{\mathcal{G}}_n$ converges to $W = \mathcal{P}$ (see [L10, Ex. 3.7]).

The case of diverging degree We discuss also in [L10, § 3.4] the consequence of having \mathcal{P} satisfying

$$\sup_{x \in I} \int \mathcal{P}(x, y) dy = +\infty. \quad (3.60)$$

Here, the uniform renormalization (3.58) is no longer adapted: we consider instead

$$\kappa_i^n = \frac{n}{\rho_n \sum_{j=1}^n \min\left(\frac{1}{\rho_n}, \mathcal{P}(x_i, x_j)\right)}, \quad i = 1, \dots, n. \quad (3.61)$$

This corresponds to renormalising the interaction in (3.1) by the averaged degree $\sum_j W_n(x_i, x_j)$ of each vertex. Here, the correct choice for the macroscopic interaction kernel W is

$$W(x, y) := \frac{\mathcal{P}(x, y)}{\int \mathcal{P}(x, z) dz}, \quad x, y \in I. \quad (3.62)$$

Note here that despite that the indegree of \mathcal{P} is unbounded (3.60), the indegree of W is uniformly 1 so that (3.49) remains valid. We see here the effect of the inhomogeneous renormalisation in the interaction, see Remark 3.3.8 and § 3.3.3 for further comments on this point. Then again, under suitable regularity conditions on \mathcal{P} , Theorem 3.3.1 is true (see [L10, § 3.4] for details). An example is the following:

Example 3.3.7 ([KVM17], Ex. 2.1). *Consider here (3.55) and (3.61) where*

$$\mathcal{P}(x, y) := (1 - \alpha)^2 x^{-\alpha} y^{-\alpha}, \quad \alpha \in \left[0, \frac{1}{2}\right), \quad x, y \in I = [0, 1]. \quad (3.63)$$

In this case, for an appropriate choice of ρ_n , $\bar{\mathcal{G}}_n$ converges to $W(x, y) = (1 - \alpha)y^{-\alpha}$, see [L10, Prop. 3.12].

Remark 3.3.8. *The macroscopic limits in Example 3.3.6 and Example 3.3.7 are the same, although the underlying graphs \mathcal{G}_n have really different structures. In Example 3.3.6, \mathcal{G}_n is more or less homogeneous whereas Example 3.3.7 is much more hub-like: nodes with positions close to 0 are connected to the whole population with probability close to 1. We see here the effect of the renormalization (3.61): it compensates for the hubs in the graph \mathcal{G}_n so that, even though the graphs \mathcal{G}_n might be different, the renormalized graphs $\bar{\mathcal{G}}^n$ are actually quite similar.*

3.3.3 Some remarks on graph convergence

On the importance of graph distances

It is worth comparing the assumption (3.48) with the hypotheses made in the homogeneous case § 3.1. Note in particular that we do not ask that $\|W^{\mathcal{G}_n} - W\|_{\infty \rightarrow \infty} \xrightarrow{n \rightarrow \infty} 0$ (which is too demanding, as noticed already in § 3.1.3). The convergence is really about the annealed graph $\bar{\mathcal{G}}_n$ towards W , not about \mathcal{G}_n towards W . In the case of Example 3.1.2, assumption (3.48) is empty, as $\|W^{\bar{\mathcal{G}}_n} - \mathbf{1}\|_{\infty \rightarrow \infty}$ is equally 0 for all n : the difficulty in handling the norm $\|\cdot\|_{\infty \rightarrow \infty}$ is really w.r.t. the randomness of the graph, not from its deterministic annealed structure. Here, the proximity between $\bar{\mathcal{G}}_n$ and \mathcal{G}_n is handled directly via concentration estimates, in the same way as in [L5]. We provide below examples of graphs where (3.48) is satisfied. The choice of the norm $\|\cdot\|_{\infty \rightarrow \infty}$ is intrinsically linked with the fact that we want to have a global propagation of chaos result, i.e. to estimate the supremum over all particles $i = 1, \dots, n$ in (3.53). The following weaker convergence (which also implies (3.54)) is possible

$$\frac{1}{n} \sum_{i=1}^n \mathbf{E} \left[\sup_{s \leq T} |\theta_s^{i,n} - \bar{\theta}_s^{x_i}|^2 \right] \rightarrow 0 \text{ as } n \rightarrow \infty. \quad (3.64)$$

Although this is not written in [L10], (3.64) is then valid under the weaker assumption (recall (3.17))

$$\|W^{\bar{\mathcal{G}}_n} - W\|_{\infty \rightarrow 1} \xrightarrow{n \rightarrow \infty} 0, \quad (3.65)$$

provided the following additional hypothesis is met

$$\sup_{y \in I} D^{\text{out}}(y) := \sup_{y \in I} \int W(x, y) dx < \infty. \quad (3.66)$$

One may find the weaker convergence assumption (3.65) more natural than (3.48) (since (3.65) is equivalent to the convergence of $\bar{\mathcal{G}}_n$ to W w.r.t. the cut-norm (3.19) [LS06]). However, it comes to a price of requiring a supplementary uniform control on the outdegrees (3.66) (the contribution of edges $z \rightarrow x$ leaving z) in the macroscopic graph induced by W , whereas (3.49) only requires the same uniform bound on the *indegree* of vertex x (i.e. the contribution of all *incoming* edges $y \rightarrow x$). This additional cost is of course transparent when W is bounded or symmetric (and this is in particular why e.g. [BCW20, BCN20] only require (3.65), as they restrict to bounded graphons). However, hypothesis (3.66) may be not satisfied in many interesting case, such as Example 3.3.6. Proving (3.64) where (3.66) is discarded remains an open question. In case of bounded W , (3.65) is indeed the correct hypothesis (this is a minor extension of Remark 3.1.6 in the Erdős-Rényi case).

On the importance of renormalisation

An important point concerning the convergence of Theorem 3.3.1 is that a lot of the structure of the microscopic graph \mathcal{G}_n is lost in the limit $n \rightarrow \infty$: the macroscopic limit (3.50) essentially captures a dynamics that lives on the *renormalized* graph $\bar{\mathcal{G}}^n$, which may be significantly different to \mathcal{G}_n , due to the presence of the renormalizing coefficients $(\kappa_i^n)_{i=1,\dots,n}$ in (3.1): when \mathcal{G}_n has vertices with diverging degree as $n \rightarrow \infty$, the limit \mathcal{P} of \mathcal{G}_n is in general different from the limit W of $\bar{\mathcal{G}}^n$ given by (3.48): Examples 3.3.6, 3.3.7 give two different \mathcal{G}_n , converging to different \mathcal{P} , such that their renormalized graphs $\bar{\mathcal{G}}^n$ converge to the same W . In particular, even though the graph of interaction \mathcal{G}_n might be of power-law type, the renormalized graph $\bar{\mathcal{G}}^n$ and its macroscopic counterpart W are *never* of power-law type: the degree of each macroscopic node of W remains uniformly of order 1, recall (3.49).

3.3.4 The nonlinear spatial profile

Let us mention briefly an auxiliary result in [L10]. In [Med14a, Med14b, KVM17, KVM18], a different approach to the large population behavior of (3.1) is considered. The point of view here is to consider the deterministic macroscopic spatial profile $(\psi(\cdot, t))_{t \in [0, T]}$ that, in our context, solves the following nonlinear integro-differential equation

$$\partial_t \psi(x, t) = F(\psi(x, t)) + \int_I \Gamma(\psi(x, t), \psi(y, t)) W(x, y) dy. \quad (3.67)$$

In the context of [Med14b], (3.67) is referred to as the nonlinear heat equation on the graph W . For FitzHugh-Nagumo dynamics with linear interaction, (3.67) corresponds to the reaction-diffusion equation addressed in the recent work [CFF19]. We consider here weak solutions to (3.67) in the sense of the following definition: if $\mathcal{C}([0, T], L^k(I))$ is the set of continuous functions with values in $L^k(I)$ for some $k \geq 2$, we say that $\psi(\cdot, t)_{t \in [0, T]} \in \mathcal{C}([0, T], L^k(I))$ is a weak solution to (3.67) if for all regular test functions $J : I \rightarrow \mathbb{R}^d$, for all $t \in [0, T]$, we have

$$\begin{aligned} \int_I \langle \psi(x, t), J(x) \rangle dx &= \int_I \langle \psi(x, 0), J(x) \rangle dx + \int_0^t \int_I \langle F(\psi(x, s)), J(x) \rangle dx ds \\ &+ \int_0^t \int_{I^2} \langle \Gamma(\psi(x, s), \psi(y, s)), J(x) \rangle W(x, y) dy dx ds. \end{aligned} \quad (3.68)$$

Following the approach of [Med14b], it is possible to consider the spatial field:

$$\theta_n(x, t) := \theta_t^{[nx+1], n} = \sum_{i=1}^n \theta_t^{i, n} \mathbf{1}_{[x_{i-1}^n, x_i^n)}(x), \quad x \in I, \quad t \geq 0. \quad (3.69)$$

The convergence is the following:

Theorem 3.3.9. *Under the previous assumptions, for almost every realization of the graph \mathcal{G}_n , the spatial field (θ_n) given in (3.69) converges weakly in $\mathcal{C}([0, T], L^k)$ to $\psi(\cdot, t)_{t \in [0, T]}$, unique solution in $\mathcal{C}([0, T], L^k)$ to (3.68) with initial condition*

$$\psi_0(x) := \int_{\mathbb{R}^d} \theta \nu_0^x(d\theta). \quad (3.70)$$

The present result can be seen as a generalization of [Med14a, Med14b, KVM17, KVM18], where the case of Kuramoto-type interaction (namely Γ of the form $\Gamma(\theta - \tilde{\theta})$ with $\Gamma(\cdot)$ and $F(\cdot)$ Lipschitz and bounded) in absence of noise ($\sigma = 0$) is considered.

3.4 Perspectives

3.4.1 Spatial extensions of phase oscillators

In the context of phase oscillators (recall Definition 1.2.1), (3.1) and its mean-field counterpart (3.51) give rise to a family of Kuramoto models with spatial extension already considered in the literature (P -nearest neighbor model [OMHS11], long-range interactions [GPR12]). In the simple case without disorder, (3.52) reads

$$\partial_t p_t(\theta, x) = \frac{\sigma^2}{2} \partial_\theta^2 p_t(\theta, x) + K \partial_\theta \left(p_t(\theta, x) \left\{ \int_I W(x, y) \int_{\mathbb{T}} \sin(\theta - \theta') p_t(\theta', y) dy \right\} \right), \quad \theta \in \mathbb{T}, \quad x \in I \quad (3.71)$$

Properly characterizing the synchronized states of (3.71), as for the original homogeneous Kuramoto model, is still an ongoing question (see [AS06, GCR14] and references therein). The fixed-point procedure recalled in § 1.2.2 in the homogeneous case gets considerably more complex as the self-consistency relation (1.18) involves now the space variable x (see [GCR14]). Although solving this relation in full generality of W might be complicated, the interesting cases where $W(x, y) = W(x - y)$ is translation invariant (in such a case the spatial interaction in (3.71) is a convolution) or when $W(x, y) = f(x)g(y)$ for two probability densities f and g on I (i.e. the so-called *Expected Degree Distribution Model*, *EDD* [CL02]) seem to be more manageable. The characterization of the dynamics of (3.1) on such graphs is a natural perspective, the important question being: to what extent the inhomogeneity of the graph may or may not induce an inhomogeneity in the dynamics? To illustrate this, we give in Figure 3.2 an explicit representation of a typical stationary measure for (3.71) in the particular case of $f(x) = g(x) = 2(1 - x)$.

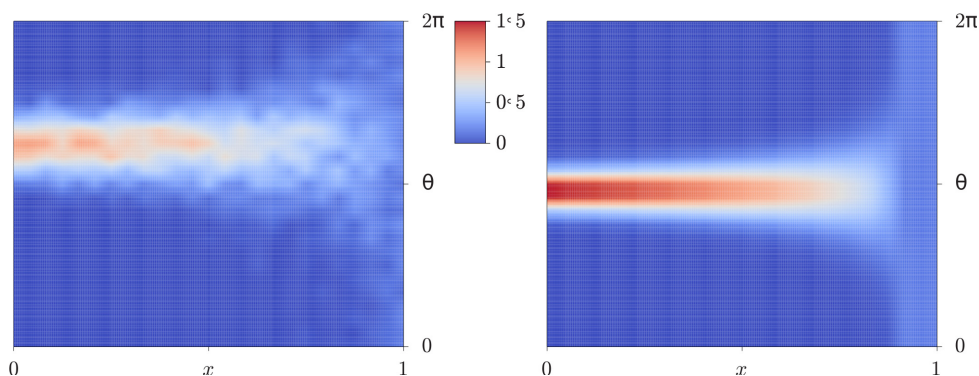


Figure 3.2 – The Kuramoto model interacting on a W -random graph given by $W(x, y) = 4(1 - x)(1 - y)$. Left: representation of the empirical distribution $\mu_{n,t} = \frac{1}{n} \sum_{i=1}^n \delta_{(\theta_i, t, x_i)}$ of the system (3.46) for regular positions $x_i = \frac{i}{n}$ on $[0, 1]$, with $n = 7000$, $K = 10$ and $t = 50$ and i.i.d. initial condition uniformly distributed on the circle \mathbb{T} . Right: a typical stationary solution to (3.71), centered in $\theta = \pi$ (any other stationary solution to (3.71) being a translation of the previous one). We see here the influence of the inhomogeneity of the graph on the dynamics: vertices with positions x close to 1 have fewer connections so that the corresponding oscillators tend to incoherence.

3.4.2 Discarding the bounded indegree assumption (3.49)

We have discussed at length in Section 3.3.3 on the importance of the hypothesis that the indegree of each node in the macroscopic graph induced by W remains uniformly bounded (recall (3.49)). We claim that this uniform control on degrees crucially depends on the choice we make

on the dilution coefficients (κ_i^n) . To illustrate this, consider the graph \mathcal{G}_n with diverging degrees defined in Example 3.3.7 where, instead of (3.61), we choose now

$$\kappa_i^n = \frac{1}{\rho_n}, \quad i = 1, \dots, n, \quad (3.72)$$

that is, the same uniform dilution as for bounded kernels (3.58) (example already considered in [KVM17], § 6.2). The graph $\bar{\mathcal{G}}_n$ remains convergent ([L10, p. 6794]) in the sense of (3.48) to

$$W(x, y) := \mathcal{P}(x, y) = (1 - \alpha)^2 x^{-\alpha} y^{-\alpha}. \quad (3.73)$$

But now, the uniform renormalization (3.72) (well adapted to vertices with low degree, with position away from 0) is no longer sufficient to compensate for vertices with high degree with position close to 0 and the boundedness assumption (3.49) is no longer satisfied for (3.73): macroscopic nodes $x \in [0, 1]$ have diverging degrees as $x \rightarrow 0$. At the level of generality considered in this work (but even for Kuramoto-type interaction), it is unclear if the present convergence results remain true when assumption (3.49) is discarded. Interestingly, the previous analysis of stationary solutions to (3.71) in the case (3.73) reveals some *explosive synchronisation*: stationary solutions to (3.71) approach formally to some Dirac mass as $x \rightarrow 0$. The rigorous analysis of this phenomenon remains to be done.

3.4.3 Going beyond the first order

The previous analysis reveals a fundamental property of (3.1): how inhomogeneous the renormalisation by κ_i^n may be, one intrinsically renormalises the sum in (3.1) by the (expected) degree of vertex i . Doing so, the leading order in the sum (and the only part that survives in the $n \rightarrow \infty$ limit) only concerns the *dense part* of the neighbors of vertex i : the diluted connections are not conserved as $n \rightarrow \infty$. To illustrate this, let us consider one simple example (which is an instance of the Stochastic Block model with only two communities): let n an even number and divide the population into two clusters $C_1^n = \{1, \dots, \frac{n}{2}\}$ and $C_2^n = \{\frac{n}{2} + 1, \dots, n\}$, and suppose that the $\xi_{i,j}^n$ are independent with Bernoulli law with parameter $p_{i,j} = p$ if i, j belong to the same cluster and $p_{i,j} = q$ if i, j belong to different clusters. To fix ideas, suppose that $p = 1$ but that $q = q_n \rightarrow 0$ as $n \rightarrow \infty$. Then the mean degree of each node is $\frac{n(1+q_n)}{2} \sim \frac{n}{2}$ so that one renormalizes with $\kappa_i^n = 2$. Obviously, the macroscopic limit as $n \rightarrow \infty$ of the empirical measure is then given by (3.51) for $W = \mathbf{1}_{[0, \frac{1}{2}]^2} + \mathbf{1}_{[\frac{1}{2}, 1]^2}$: at first order, the mean-field limit only captures the dynamics of two disjoint mean-field communities! But obviously this does not represent the actual dynamics of the particle system for n large but finite as the two communities are connected: there is synchronization (see Figure 3.3). Open questions at this point are: (i) can we derive correctly a second-order expansion of the empirical measure around (3.51) capturing the connectivity between clusters? (ii) is it at the level of the LLN or at the level of fluctuations? (iii) on which time scale (possibly depending on q_n) do we see synchronisation? (iv) and what about a general graph with a dense *skeleton* with a diluted neighborhood?

3.4.4 Fluctuations and long-time behavior

Concerning fluctuations, there is every indication that the CLT analysis done in [L14] should not be specific to the homogeneous Erdős-Rényi case and should readily apply to the inhomogeneous case of W -random graphs (at least for regular and bounded graphons W). The exact form of the fluctuations processes is not absolutely clear, though, but one can expect straightforward extensions. Theorem 3.3.1 comes with all the comments and restrictions raised in [L5]: the convergence results are only valid on bounded time intervals $[0, T]$, where T is independent of n (although that, with a little more work, it would certainly be possible to extend the result up to times T which grows logarithmically in n , as in [L5], Corollary 1.2). The question of the long-time analysis of the empirical measure μ^n on larger time scales (for which we will discuss at length in Chapter 5 in the mean-field case) remains open even in the homogeneous Erdős-Rényi graph. Applying the formalism of Grothendieck inequalities on this matter is an obvious natural issue.

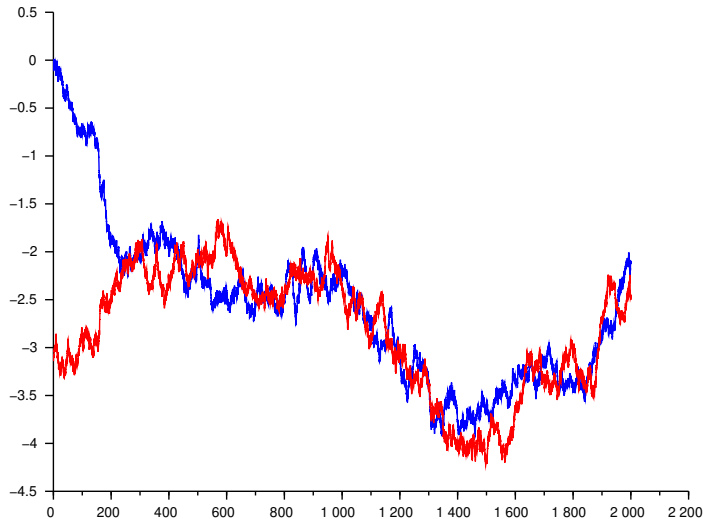


Figure 3.3 – Trajectories of the center of synchronization of two mean-field communities of Kuramoto oscillators, with additional random connections between the two communities with probability $p_n = \frac{\lambda}{n}$, for $n = 500$, $K = 5$, $\lambda = 4$, up to $T = 2000$. Although the asymptotic graphon consists of two disjoint communities $W = \mathbf{1}_{[0, \frac{1}{2}]^2} + \mathbf{1}_{[\frac{1}{2}, 1]^2}$, the finite size clusters synchronize.

3.4.5 Application to neuroscience: neural field equation and traveling waves

A case of interest for (3.51) concerns the case of FitzHugh-Nagumo oscillators (recall Example 1.3.2). Theorem 3.3.1 gives a new interpretation of such spatially-extended PDEs in terms of the mean-field limit of diffusions on random graphs. Several existing works [CFF19, Sta13, ABR04] have considered the same PDE, especially when it comes to its stability of synchronized patterns in presence of macroscopic noise. The question of the links between these macroscopic features and the microscopic system (3.1) remain unclear. Another interesting case is when one considers (3.67) when $d = 1$, $F(\theta) = -\alpha\theta$ (for some $\alpha > 0$) and $\Gamma(\theta, \tilde{\theta}) = f(\tilde{\theta})$ (typically f is a sigmoid function):

$$\partial_t \psi(x, t) = -\alpha \psi(x, t) + \int_I f(\psi(y, t)) W(x, y) dy. \quad (3.74)$$

Equation (3.74) is nothing else than the neural field equation, introduced from a phenomenological point of view by Wilson and Cowan [WC72] and Amari [Ama77] in order to describe the macroscopic activity of a population of neurons with spatial extension (we will come back to it in Chapter 6). Eq. (3.74) has been the subject of an extensive literature (see [Bre14b, Bre12] and references therein; see in particular the recent works [CDLO19, AN22b] showing that (3.74) is a proper limit for spatially-extended Hawkes processes). An important issue here is the existence and stability of traveling waves [FK18, RAB⁺14]. In this context, it is reasonable to expect that the effect of thermal noise in (3.1) will persist on larger time intervals, resulting in stochastic neural field equations [FI15, IM16, KS14, Lan16].

3.4.6 Interaction on balanced networks

In (3.51), the kernel W is supposed to be nonnegative. In a neuroscience context, this corresponds to an excitatory network. It is certainly possible to accommodate for inhibitory connections in (3.1): in order to keep it simple, think e.g. to the following homogeneous case (but the reader may convince himself that the extension to W -random graphs would follow easily): simply choose $\xi_{i,j}^n = 1$ with probability p and $\xi_{i,j}^n = -1$ with probability $1 - p$. The mean-field limit would follow easily with kernel $W \equiv (1 - 2p)$. An interesting case would correspond to a balanced network when $p = \frac{1}{2}$. Then, we fall in a CLT regime where the interaction typically

scales as $\frac{1}{\sqrt{n}}$, not $\frac{1}{n}$. This situation is reminiscent to spin glass systems [BAG97, CT18b] with only difference that the interaction is discrete, not Gaussian. This situation has already been considered in [PRS22, ELL22] in a context of Hawkes processes, but the diffusion case, to the best of my knowledge, does not seem to have been considered (especially when it comes to general inhomogeneous random graphs).

3.4.7 Other topologies and time dependent graphs

Going beyond the optimal dilution condition $np_n \rightarrow \infty$ (3.10) requires to look at the case where $np_n \rightarrow \lambda$. As already mentioned, we leave here the framework of mean-field analysis, as the correct limit is given in terms of diffusions interacting on Galton-Watson trees [OR19, LRW23]. The dynamics of such diffusions on trees (and their extensions e.g. to deterministic regular trees) is far from being understood. A possibility would be to look first at stationary versions of these processes (e.g. the XY model), see [DM10a, DM10b] for related works in the Ising case.

This chapter only concerns the case of diffusions interacting on graphs to do not depend on time. There is from a modelling perspective a crucial need to understand the corresponding dynamics where the graph of interaction itself evolves with time (to mention only one application, the need to understand synaptic plasticity in neuronal networks [RV21b, RV21a]). So far, existing results mainly tackle the easier case where the evolution of the graph is independent from the particles [BBW19], but the interesting point is obviously when the graphs depends itself on the particle dynamics. Understanding the role of the dynamics on the graphs (possibly on toy models derived from the formalism of W -random graphs of the present chapter), and the different time scales for both dynamics is a crucial issue, on a longer perspective.

Chapter 4

Periodicity for nonlinear Fokker-Planck equations

The present chapter is concerned with the difficult question of the dynamics of the solution μ_t to the Fokker-Planck equation (1.4) when $t \rightarrow \infty$. As discussed in § 1.3, this issue is model-dependent and has been mostly considered in reversible situations. We are mostly interested in situations where μ_t exhibits periodic behaviors.

4.1 About Stable Normally Hyperbolic Manifolds (SNHM)

4.1.1 A general framework

In order to analyse (1.4) as $t \rightarrow \infty$ we need structure: the key notion of this chapter concerns Stable Normally Hyperbolic Manifolds (SNHM). Let us briefly give an informal definition (see [BLZ98]): let X be a functional (Banach) space with norm $|\cdot|$. Let $T^t : X \rightarrow X$, $t \geq 0$ a semiflow that gives rise to an evolution $\{u_t\}_{t \geq 0}$, with $u_0 = u \in X$ given by $u_t := T^t u$. One says that a compact regular manifold M is a Stable Normally Hyperbolic Manifold (SNHM) if for all $m \in M$, there exists a decomposition

$$X = X_m^c \oplus X_m^u \oplus X_m^s \quad (4.1)$$

of closed subspaces in X with X_m^c the tangent space to M at m (here, c : *center*, u : *unstable* and s : *stable*) such that the following holds

1. For each $m \in M$, $t \geq 0$, for $m_t = T^t(m)$, the differential of the semiflow $DT^t(m)|_{X_m^\alpha} : X_m^\alpha \rightarrow X_{m_t}^\alpha$ for $\alpha = c, u, s$ and $DT^t(m)|_{X_m^u}$ is an isomorphism from X_m^u onto $X_{m_t}^u$.
2. There exists $t_0 \geq 0$ and $\lambda < 1$ such that for all $t \geq t_0$

$$\lambda \inf \{ |DT^t(m)x^u|, x^u \in X_m^u, |x^u| = 1 \} > \max(1, \|DT^t(m)|_{X_m^c}\|), \quad (4.2)$$

$$\lambda \min \{ 1, \inf \{ |DT^t(m)x^c|, x^c \in X_m^c, |x^c| = 1 \} \} > \|DT^t(m)|_{X_m^s}\| \quad (4.3)$$

Condition (4.2) essentially says that the semiflow T^t is, locally around m , expansive in the unstable direction X_m^u at a rate that is strictly greater than on X_m^c , while (4.3) states that T^t is locally contractive along the stable direction X_m^s at a rate greater than on X_m^c . This kind of dynamical structure originally goes back to Fenichel [Fen72, Fen79] for flows in finite dimension and then generalized in [HPS77] in the case of Riemannian manifolds and in [BLZ98, SY13] in the infinite dimensional setting. A crucial point of the previous references (and an essential argument for us) is to note that these structures are stable under perturbations [BLZ98, SY13].

The main goal of this chapter is to address the existence of SNHM for the nonlinear Fokker-Planck equation (1.4) and to analyse the reduced dynamics on these manifolds (in particular limit cycles). Note that the solutions μ to (1.4) take values in the set of probability measures: we are intrinsically working in an infinite dimensional set-up, a typical choice of X being some $L^2_\rho(\mathbb{R}^d)$ or some Sobolev space $H_\rho^{-k}(\mathbb{R}^d)$ for some index $k \geq 0$ and possibly some weight ρ , (or affine spaces constructed upon such spaces in order to keep track of the fact that we are dealing with probability measures). In this infinite dimensional setting, the fact that we may

have to deal with e.g. unbounded coefficients F in (1.4) (think of the FitzHugh-Nagumo case, Example 1.3.2) is a major difficulty: applying the previously mentioned theoretical stability results of [BLZ98, SY13] will not be necessarily possible. A serious issue will be to build ad-hoc procedures and to find suitable functional spaces where to obtain similar existence and stability results.

4.1.2 An alternative definition

We are here mostly interested in situations with oscillations, that is where M is a one-dimensional circle or more generally some stable limit cycle. In such a case, the unstable component X_m^u in (4.1) is trivial and the definition of a SNHM can be expressed in terms of the following equivalent formulation: associate to the dynamics $u_t = T^t u$, with $u_0 = u$, the linear evolution semigroup $\{\Phi(u, t)\}_{t \geq 0}$ in X , satisfying $\partial_t \Phi(u, t)v = L(t)\Phi(u, t)v$ and $\Phi(u, 0)v = v$, where $L(t)$ is the operator obtained by linearizing the evolution around u_t . Then, a 1-dimensional curve, $M \subseteq X$ is a SNHM of characteristics λ_1, λ_2 ($0 \leq \lambda_1 < \lambda_2$) and $C > 0$ if M is a C^1 compact connected manifold which is invariant under the dynamics and for every $u \in M$ there exists a projection $P^o(u)$ on the tangent space of M at u , that is $\mathcal{R}(P^o(u)) =: T_u M$, which, for $v \in L_0^2$, satisfies the following properties:

1. for every $t \geq 0$ we have

$$\Phi(u, t)P^o(u_0)v = P^o(u_t)\Phi(u, t)v, \quad (4.4)$$

2. we have

$$|\Phi(u, t)P^o(u_0)v| \leq C \exp(\lambda_1 t)|v|, \quad (4.5)$$

and, for $P^s := 1 - P^o$, we have

$$|\Phi(u, t)P^s(u_0)v| \leq C \exp(-\lambda_2 t)|v|, \quad (4.6)$$

for every $t \geq 0$;

3. there exists a negative continuation of the dynamics $\{u_t\}_{t \leq 0}$ and of the linearized semigroup $\{\Phi(u, t)P^o(u_0)v\}_{t \leq 0}$ and for any such continuation, we have, for $t \leq 0$,

$$|\Phi(u, t)P^o(u_0)v| \leq C \exp(-\lambda_1 t)|v|. \quad (4.7)$$

Once again, the condition $0 \leq \lambda_1 < \lambda_2$ expresses the fact that the rate of contraction of the linear dynamics along the orthogonal component is strictly greater than the rate of expansion along the tangential component.

The problem of existence of such manifolds left apart, another crucial issue concerns the regularity of such structures (e.g. for limit cycles, expressed in terms of some regular isochron map associated to the dynamics). Whereas this regularity is essentially straightforward in the finite-dimensional case $X = \mathbb{R}^d$ (see [Fen79]), the same issue in the infinite-dimensional setting is no longer clear (see [BLZ98] for results in this direction). Some effort will be spend here to obtain regularity estimates that will be crucial for us in the next chapter, where the microscopic companion problem will be addressed.

4.1.3 A seminal work: perturbation analysis for phase oscillators

The first work to address the emergence of nontrivial structured collective behaviors for nonlinear Fokker-Planck equations goes back to the paper of Giacomin, Pakdaman, Pellegrin and Poquet [GPPP12] in the case of the Active rotators model (Example 1.2.3). In order to motivate the next results (and in particular to motivate why we *cannot* directly apply the same techniques in the case of FitzHugh-Nagumo oscillators), we briefly review here the results of [GPPP12]: [GPPP12] slightly generalizes (1.16) into

$$\partial_t p_t(\theta) = \frac{1}{2} \partial_\theta^2 p_t(\theta) - \partial_\theta (p_t(\theta)(J * p_t(\theta))) + \delta G[p_t](\theta), \quad \theta \in \mathbb{T}. \quad (4.8)$$

where G is a smooth and bounded functional from $L^2(\mathbb{T})$ to $H^{-1}(\mathbb{T})$ (see [GPPP12], § 2.2 for more examples). This is indeed the case for $G[p] := \partial_\theta(p\partial_\theta V)$ in (1.16), as V is supposed to be smooth and bounded. The main result of [GPPP12] (Th. 2.1) states the existence of a smooth SNHM M^δ for the dynamics (4.8) for $\delta > 0$ small enough. Asymptotics in δ of the dynamics of the phase along this manifold are also provided, in the general case [GPPP12, Th. 2.2] and in the case of Active rotators [GPPP12, § 3].

As it will be at the core of the arguments given in the following, let us briefly comment on the main structure of the proof here (which relies on arguments given in [SY13] with special attention given in [GPPP12] to the relationships between the various small parameters). The starting point is to note that when $\delta = 0$ in (4.8), we retrieve the reversible equation (1.19). Recalling Section 1.2.2, when $K > 1$, the circle M_0 of stationary solutions to (1.19) defined in (1.21) is linearly stable (recall Theorem 1.2.4). Hence, M_0 is a simple instance of a SNHM in the sense of Section 4.1.2, for the choice of $\lambda_1 = 0$ and $\lambda_2 = \lambda_K$ (where we recall the spectral gap λ_K in Theorem 1.2.4). Note here that $X = L^2(\mathbb{T})$ and for $q_\psi \in M_0$, the projection $P_{q_\psi}^o$ on the tangent space is explicitly written as $P_{q_\psi}^o u = \frac{\langle u, \partial_\theta q_\psi \rangle_{-1, q_\psi}}{\langle \partial_\theta q_\psi, \partial_\theta q_\psi \rangle_{-1, q_\psi}} \partial_\theta q_\psi$, where $\langle \cdot, \cdot \rangle_{-1, q}$ is the scalar product associated to the norm (1.25).

Now proceed with perturbation arguments for small $\delta > 0$. At least in a small neighborhood N of M_0 , one can construct a suitable projection on M_0 that is adapted to the underlying local Hilbertian structure: for any $p \in N$, there exists a unique $q := v(p) \in M_0$ such that $p - v(p)$ is orthogonal to M_0 at point $v(p)$: $\langle p - q, \partial_\theta q \rangle_{-1, q} = 0$. Define then the space $\mathcal{F}(\varepsilon, l)$ of continuous functions $f : M_0 \rightarrow L^2$ such that $\|f\|_\infty \leq \varepsilon$ and $\|f\|_{Lip} \leq l$ for some $\varepsilon > 0$ and $l \in [0, 1]$, as well as $\langle f(q), \partial_\theta q \rangle_{-1, q} = 0$ (think as $f(q)$ as a notion of a *distance* around q , so that for any f , $M' := (Id + f)(M_0)$ is a perturbation of size ε of the original manifold M_0). The main point is to find a suitable f so that the new manifold M' remains invariant for the dynamics of (4.8): in order to do so, construct a set of functionals $\{X_t\}_{t \geq 0} : \mathcal{F}(\varepsilon, 1) \rightarrow \mathcal{C}(M_0, L^2(\mathbb{T}))$ that is adapted to the dynamics (4.8) such that, for some T sufficiently large, $(X_t)_{t \in [T, 2T]}$ stabilises $\mathcal{F}(\varepsilon, 1)$ and is contracting on such space:

$$\|X_T(f_1) - X_T(f_2)\|_\infty \leq \frac{1}{2} \|f_1 - f_2\|_\infty, \quad f_1, f_2 \in \mathcal{F}(\varepsilon, 1). \quad (4.9)$$

More precisely, one defines X_t as the unique mapping satisfying for any $t > 0$,

$$X_t(f)(v(p(t; p_0))) = p(t; p_0) - v(p(t; p_0)), \quad p_0 = q + f(q), \quad q \in M_0, \quad (4.10)$$

where $p(t; p_0)$ is the solution to (4.8) at time t with initial condition p_0 . In words, starting at a distance $f(q)$ from $q \in M_0$, for any t , $X_t(f)$ associates to the phase $v(p(t; p_0)) \in M_0$ of the solution $p(t; p_0)$ at time t , its actual *height* above M_0 , i.e. $p(t; p_0) - v(p(t; p_0))$. Hence, if X_t possesses a unique fixed-point f_0 , i.e. $X_t(f_0) = f_0$, one obtains readily from the previous identity that $M_\delta := (Id + f_0)(M_0)$ is indeed an invariant manifold for the dynamics that has all the requirements for a SHNM, see Figure 4.1 and [GPPP12] for further details. Proving both the stability $X_T(\mathcal{F}(\varepsilon, 1)) \subset \mathcal{F}(\varepsilon, 1)$ and the contraction (4.9) requires to estimate the proximity of a solution p_t to (4.8) to some generic $q \in M$ (as well as the dependence of such quantities when the initial condition varies). The key to such estimates is based on the following mild formulation: for $q \in M_0$ and $w_t = p_t - q$, w_t verifies

$$w_t = e^{-tL_q^0} w_0 + \int_0^t e^{-(t-s)L_q^0} (\partial_\theta(w_s J * w_s) + \delta G[q + w_s]) ds. \quad (4.11)$$

All of the arguments developed above rely on a careful control of two antagonistic terms in (4.11): first, the contraction of the semigroup $e^{-tL_q^0}$ (Theorem 1.2.4) so that if w_0 is of order ε , $e^{-tL_q^0} w_0$ will be of order, say, $\frac{\varepsilon}{4}$ if t is sufficiently large and secondly, the remaining perturbations terms in (4.11): firstly the term δG , that however remains of order δ , by boundedness of G and secondly, the quadratic term $\partial_\theta(w_s J * w_s)$. Dealing with this last term is not an essential problem: proximity arguments around M_0 require w to be small, so that $\|w\|^2$ is even smaller and standard extensions of Grönwall lemmas apply. Putting everything together roughly says that

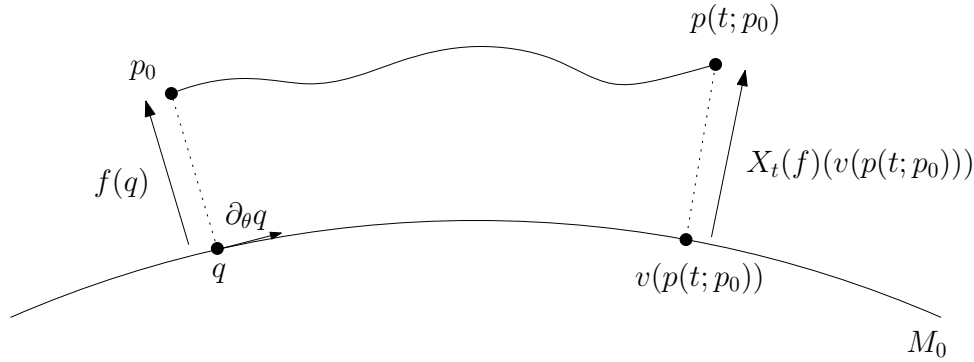


Figure 4.1 – Construction of the functional X_t : starting from p_0 at a distance $f(q)$ from $q \in M_0$, one follows the dynamics of (4.8) until time t where one projects back onto M_0 . $X_t(f)$ is the distance between $p(t; p_0)$ and its projection $v(p(t; p_0))$ onto M_0 . Hence, if f is the unique fixed-point to X_t and one defines $M' := (Id + f)(M_0)$, we have $v(p(t; p_0)) + f(v(p(t; p_0))) = p(t; p_0)$ so that $p(t; p_0) \in M'$: M' is invariant for the dynamics.

w_t will be of order $\frac{\varepsilon}{4} + O(\delta)$ which can be rendered smaller than $\frac{\varepsilon}{2}$ if δ is taken sufficiently small: up to technical details, we have essentially obtained our contraction result (4.9). Note here that a key argument of the previous proof is to require that G is bounded and smooth. An important difficulty we will be facing in § 4.3 is when we will have to move from the compact state space \mathbb{T} to \mathbb{R}^d : we will have to deal with *unbounded* functionals G (a prominent example being the polynomial FitzHugh-Nagumo case in Example 1.3.2), so that the control of the counterpart of (4.11) will be less clear.

4.2 Traveling waves in the asymmetric Kuramoto model

Before dealing with these difficulties, let us mention briefly the results of [L2] where the techniques developed in [GPPP12] are applied to the disordered Kuramoto model (Example 1.2.2). As the proof relies on the same perturbation techniques, we leave the technical details to [L2]. We reproduce here (1.15) for convenience

$$\partial_t p_t(\theta, \omega) = \frac{1}{2} \partial_\theta^2 p_t(\theta, \omega) - \partial_\theta \left(p_t(\theta, \omega) \left\{ \int J * p_t(\cdot, \omega') \nu(d\omega') + \delta U(\theta, \omega) \right\} \right), \quad \theta \in \mathbb{T}, \quad \omega \in \text{Supp}(\nu). \quad (4.12)$$

for the choice of $U(\theta, \omega) := \omega$. Hence, the dependence in the disorder taken apart, (4.12) is formally equivalent to (4.8), for the bounded perturbation $G[p](\theta, \omega) = -\partial_\theta(p(\theta, \omega)\omega)$. Note that this approach generalises easily to generic perturbations U , namely the case of Active rotators $U(\theta, \omega) = b + \omega + a \sin(\theta)$, $a, b \in \mathbb{R}$. We refer to [SK86b, GPPP12] for more details. The aim of [L2] is to analyse the influence of the law ν of the disorder on the long-time dynamics of (4.12). Note first that one can always assume ν to be centered (otherwise, the system is mapped into the same system rotating at constant speed $\int \omega \nu(d\omega)$). Two natural questions arise: 1) in case of symmetric ν , what are the stability properties of the manifold of synchronised profiles $q_{r,\psi}(\theta, \omega)$ parameterized by $r > 0$ solving the fixed-point relation (1.23) and 2) what happens if the distribution ν is not symmetric?

4.2.1 Emergence of rotating waves for asymmetric disorder

The main result of [L2] is actually independent on any symmetry assumption of ν . Making $\delta = 0$ in (4.12) does not here exactly give the non-disordered equation (1.19) but rather the similar

$$\partial_t p_t(\theta, \omega) = \frac{1}{2} \partial_\theta^2 p_t(\theta, \omega) - \partial_\theta \left(p_t(\theta, \omega) \int J * p_t(\cdot, \omega') \nu(d\omega') \right) \quad (4.13)$$

where the dependence in ω simply reduces to ineffective labels, but its evolution turns out to be identical to (1.19). In the statement below, $\|\cdot\|_{2,\nu}$ is the L^2 -norm w.r.t. the measure $d\theta \otimes \nu$

and $q \in M_0$ is the element of the synchronized manifold such that for all ω , $q(\cdot, \omega) = q_0(\cdot)$. The main result of [L2] is then

Theorem 4.2.1 (Theorem 2.2 in [L2]). *Suppose that $\text{Supp}(\nu) \subset [-1, 1]$. For every $K > 1$, there exists $\delta_0 = \delta_0(K) > 0$ such that for $|\delta| \leq \delta_0$, there exists $\tilde{q}_\delta \in L^2(d\theta \otimes \nu)$, satisfying $\|\tilde{q}_\delta - q\|_{2,\nu} = O(\delta)$ and a value $c_\mu(\delta) \in \mathbb{R}$ such that if we set*

$$q_t^{(\psi)}(\theta, \omega) := \tilde{q}_\delta(\theta - c_\mu(\delta)t - \psi), \quad (4.14)$$

then $q_t^{(0)}$ solves (4.12). Moreover, the family of solutions $\{q_t^{(\psi)}\}_\psi$ is stable: there exist $\beta = \beta(K) > 0$ and $C = C(K) > 0$ such that if p_0^δ is close enough to M_0 and $\int_{\mathbb{T}} p_0^\delta(\theta, \omega) d\theta = 0$ $\nu(d\omega)$ -a.s., then there exists $\psi_0 \in \mathbb{T}$ such that for all $t \geq 0$

$$\|q_t^{(\psi_0)} - p_t^\delta\|_{2,\nu} \leq 2C \exp(-\beta t). \quad (4.15)$$

In other words, a nontrivial disorder ν induces a macroscopic rotating wave for (4.12) with speed $c_\nu(\delta)$, on a bounded time scale (that is on $[0, T]$ for any fixed $T > 0$). Theorem 4.2.1 is complemented in [L2] with asymptotic development as $\delta \rightarrow 0$ of the speed of rotation:

$$c_\mu(\delta) = c_3\delta^3 + c_5\delta^5 + \dots \quad (4.16)$$

where the coefficients c_k are expressed in terms of the linearized evolution operator of (4.13) around $q_0 \in M_0$, $Au(\theta, \omega) := \frac{1}{2}\partial_\theta^2 u(\theta, \omega) - \partial_\theta \left(q_0(\theta) \int J * u(\theta) d\nu + u(\theta, \omega) J * q_0(\theta) \right)$. We refer to [L2, § 2.2] for more details on this expansion.

4.2.2 Stability of synchronization in the symmetric case

When ν is symmetric, the development (4.16) yields zero terms to all orders: this only reflects the fact that the manifold $M_\delta = \{\tilde{q}_\delta(\cdot - \psi), \psi \in \mathbb{T}\}$ (unique in a neighborhood of M_0 of size δ for δ small enough, which converges to M_0 as $\delta \searrow 0$, in C^j , for every j) consists of stationary points. In this case, Theorem 4.2.1 provides only a stability statement for M_δ . This result can be sharpened, with precise results on the linearisation around each element $q \in M_\delta$: define the linear operator

$$L_q u(\theta, \omega) := \frac{1}{2}\partial_\theta^2 u(\theta, \omega) - \partial_\theta \left(u(\theta, \omega) (\langle J * q \rangle_\nu(\theta) + \delta\omega) + q(\theta, \omega) \langle J * u \rangle_\nu(\theta) \right). \quad (4.17)$$

A summarised statement of Th. 2.5 in [L2] is then

Theorem 4.2.2 (Th. 2.5 in [L2]). *Assume $\text{Supp}(\nu) \subset [-1, 1]$. Then for all $K > 1$, for $\delta > 0$ sufficiently small, L_q is the infinitesimal generator of an analytic semi-group, the spectrum of L_q lies in a cone $C \subseteq \{z \in \mathbb{C}; \Re(z) \leq 0\}$, 0 is a simple eigenvalue, with eigenspace spanned by $(\theta, \omega) \mapsto \partial_\theta q(\theta, \omega)$ and the distance between 0 and the rest of the spectrum is strictly positive, perturbation of the spectral gap λ_K of the Kuramoto model without disorder.*

Although written in a qualitative way, what we really prove are quantitative explicit estimates. Proof of Theorem 4.2.2 relies on standard techniques concerning relatively bounded perturbations of linear operators (see e.g. [Paz83, Kat95]). We refer to [L2] for more details. The present analysis shows that, in the symmetric case, the disorder has essentially no influence on the *macroscopic* system (4.12). This in particular does not explain the non-self-averaging effect mentioned in Section 2.2 concerning the *microscopic* system (1.14). Pushing further the analysis from the macroscopic PDE (4.12) to the microscopic empirical measure (2.2) of the system requires some further work that will be discussed later in Chapter 5.

4.3 Excitable systems and the FitzHugh-Nagumo model

We now turn to the main point of the present chapter, that is the analysis of the more general case of emergence of oscillations for interacting excitable systems in \mathbb{R}^d in the presence of noise and linear interaction, that we write as follows

$$\partial_t \mu_t(x) = \nabla \cdot (\sigma^2 \nabla \mu_t)(x) + \nabla \cdot \left(K \mu_t(x) \left(x - \int_{\mathbb{R}^d} z \mu_t(dz) \right) \right) - \delta \nabla \cdot (\mu_t(x) F(x)), t \geq 0, \quad (4.18)$$

where $F : \mathbb{R}^d \rightarrow \mathbb{R}^d$. Solution $(\mu_t)_{t \geq 0}$ to (4.18) is a probability measure-valued process on \mathbb{R}^d , describing the law of the nonlinear process

$$dX_t = \left(\delta F(X_t) - K(X_t - \mathbf{E}[X_t]) \right) dt + \sqrt{2\sigma} dB_t, t \geq 0. \quad (4.19)$$

The microscopic counterpart of (4.19) is of course

$$dX_{i,t} = \left(\delta F(X_{i,t}) - K \left(X_{i,t} - \frac{1}{n} \sum_{j=1}^n X_{j,t} \right) \right) dt + \sqrt{2\sigma} dB_{i,t}, i = 1, \dots, n, t \geq 0, \quad (4.20)$$

where B_1, \dots, B_n is a collection of independent standard Brownian motions in \mathbb{R}^d . Two parameters are relevant for (4.18): $K = \text{diag}(k_1, \dots, k_d)$ and $\sigma = \text{diag}(\sigma_1, \dots, \sigma_d)$ (with $k_i \geq 0, \sigma_i \geq 0$), two matrices (supposed to be diagonal for simplicity) modelling respectively the intensity of interaction and noise within the population.

4.3.1 Persistence and emergence of synchronized structures under noise and interaction

Some general questions

Our main (but not exclusive) interest concerns situations where the isolated system (i.e. $K = \sigma = 0$ and $\delta = 1$ in (4.19))

$$dX_t = F(X_t) dt, \quad (4.21)$$

possesses excitable features, as described in Section 1.3.1.

Question 4.3.1. *Two main issues are particularly relevant here:*

1. Persistence of periodicity: *in case the isolated system (4.21) has a periodic behavior, does this periodicity persist with nontrivial interaction and noise in (4.19)?*
2. Emergence of periodicity: *is there a possibility that noise and interaction may induce for (4.20) a structured dynamics (e.g. synchronization, collective periodic behavior, oscillations) that is not originally observed for the isolated system (4.21)?*

Generically, in order to maintain and observe macroscopic oscillations in (4.18), some balance has to be found in the intensity of noise and interaction that one needs to put in the system. Hence, addressing Question 4.3.1 requires to obtain quantitative estimates on the relative size of the interaction K and the noise σ in order to maintain these macroscopic structures. The aim of this section is to provide a general framework in order to answer rigorously to Question 4.3.1 and to give concrete illustrations of applications of this framework.

The particular case of FitzHugh-Nagumo dynamics

Although several examples are discussed at length in [L9] (see e.g. Stuart-Landau oscillators or the Cucker-Smale alignment model [BCnCD16]), for simplicity, the only example that we discuss in this manuscript is the FitzHugh-Nagumo model (Example 1.3.2). In this context, (4.18) corresponds to FitzHugh-Nagumo oscillators with electrical synapses [BFFT12]. We are interested in two instances of the model (4.18): the *elliptic case*, where both interaction and noise are nontrivial on all coordinates: $k_i > 0, \sigma_i > 0$ for all $i = 1, \dots, d$ and the *kinetic case* (mathematically more challenging but biologically more adequate), where interaction and noise

in only present on the voltage variable. These two cases will be addressed in this section. Addressing Question 4.3.1 in the whole generality of Example 1.3.2 is certainly a complicated task, as the bifurcation diagram for the isolated system (4.21) is known to be complex [RsGGt00]. One may highlight however two main scenarios for (4.21) (both of interest in biological applications [LGONSG04, TPVE20]): the *excitable case*, where the nullcline of w in (1.29) intersects only once the nullcline of v and the *bistable case* where three distinct equilibrium points (one unstable, the other two stable) coexist for the dynamics (4.21). Again, for simplicity of exposition, we will be only concerned here with the excitable case and we refer to [L9] where the bistable case is addressed.

The model (4.18) with FHN dynamics has been the subject of several previous mathematical papers: [MQT16] addresses well-posedness results for (4.18) and proves the existence of equilibria for the limit mean-field dynamics in the weak interaction regime ($K \rightarrow 0$), in the kinetic case. [nT20] addresses the same model in the case of strong interaction ($K \rightarrow \infty$) and proves the convergence of solutions to (4.18) to singular Dirac solutions (clamping) in this case. Note that neither $K \rightarrow 0$ nor $K \rightarrow \infty$ are proper frameworks for the emergence of collective oscillations: intuitively, a reasonable balance between noise and interaction is needed, so that K should not be too small or too large. Our analysis lies precisely within this range of parameters:

Hypothesis 4.3.2. *We suppose that the parameters in (4.18) are such $\delta \rightarrow 0$ and K and σ remain of order 1.*

The persistence and emergence of periodic behaviors under noise and interaction in the FitzHugh-Nagumo model has been observed in several previous papers based on numerical simulations (see e.g. [BFFT12, TPVE20]), but a rigorous proof was lacking. To our knowledge, the results of the present section provide the first mathematical proof of these phenomena.

Why addressing Question 4.3.1 is difficult

Our analysis will rely again on SNHM structures as in § 4.1.1. Note however that the nature of the dynamics F poses serious major difficulties that we want to emphasize here. Firstly, in whole generality, contrary to phase oscillators, there is no rotational invariance for (4.18), so that one cannot build upon the hypothetic existence of one limit cycle for (4.18) when $\delta = 0$: as we will see below, (4.18) when $\delta = 0$ simply reduces to some stationary Ornstein-Uhlenbeck process with Gaussian invariant measures. Hence, one difficulty will be to *prove* the existence of limit cycles, whereas it was more or less obvious *by construction* in the case of phase oscillators. Secondly, a major simplifying technical assumption in case of phase oscillators is that \mathbb{T} is compact. In case of generic particles in \mathbb{R}^d , one will have to be careful about the control one has on the dynamics of (4.19) at infinity: as we will see below, a proper functional setting for the analysis of (4.18) is to work with L^2 -spaces L^2_ρ with Gaussian weights (in which the Ornstein-Uhlenbeck operator is self-adjoint). In such spaces, the major difficulty is that the perturbation $G[\mu] := \nabla \cdot (\mu(x)F(x))$ is *not bounded*, typical examples being F with sub-exponential (or polynomial) growth, as in Example 1.3.2. Hence, we cannot apply directly the existing perturbations results as in [GPPP12, BLZ98]. Control of such perturbation will rely on a careful interplay between such L^2_ρ weighted spaces with different weights ρ .

4.3.2 The elliptic case

Let us first review the work [L9], where existence of limit cycles for (4.18) is proven in the *elliptic case*, that is $k_i > 0$, $\sigma_i > 0$ for all $i = 1, \dots, d$. Note here that the result of [L9] is general and does not restrict to the FHN model.

Slow-fast dynamics and invariant manifold

Our approach is based on the fact that, since the two first terms of the right hand side of (4.18) leave the mean

$$m_t := \Pi_m(\mu_t) := \int_{\mathbb{R}^d} x \mu_t(dx) \in \mathbb{R}^d \quad (4.22)$$

invariant, this PDE is in fact equivalent to the system

$$\begin{cases} \partial_t p_t(x) &= \nabla \cdot (\sigma^2 \nabla p_t(x)) + \nabla \cdot (K p_t(x) x) + \nabla \cdot (p_t(x) (\dot{m}_t - \delta F(x + m_t))) \\ \dot{m}_t &= \delta \int_{\mathbb{R}^d} F(x + m_t) p_t(dx) \end{cases}, \quad (4.23)$$

where $p_t := \Pi_0(\mu_t)$ is the centered version of μ_t , i.e. satisfies for all test function φ

$$\int_{\mathbb{R}^d} \varphi(x) p_t(dx) := \int_{\mathbb{R}^d} \varphi(x) \Pi_0(\mu_t)(dx) := \int_{\mathbb{R}^d} \varphi(x - m_t) \mu_t(dx). \quad (4.24)$$

The mapping $\mu \mapsto (\Pi_0(\mu), \Pi_m(\mu))$ is an obvious one-to-one correspondence between $\mathcal{P}_1(\mathbb{R}^d)$, the set of probability measures on \mathbb{R}^d with first finite moment and $\mathcal{P}_0(\mathbb{R}^d) \times \mathbb{R}^d$, where $\mathcal{P}_0(\mathbb{R}^d) := \{p \in \mathcal{P}_1(\mathbb{R}^d), \int_{\mathbb{R}^d} xp(dx) = 0\}$, whose inverse given by $(p, m) \mapsto \mu$ defined by $\int \varphi(x) \mu(dx) = \int_{\mathbb{R}^d} \varphi(x + m) p(dx)$.

Taking δ small, the system (4.23) defines a slow-fast dynamics. Making $\delta = 0$ in (4.23) gives $\partial_t p_t(x) = \nabla \cdot (\sigma^2 \nabla p_t(x)) + \nabla \cdot (K p_t(x) x)$ that is nothing else than an Ornstein-Uhlenbeck dynamics, for which it is standard to prove exponential convergence to the Gaussian invariant measure with density $q_{0, \sigma^2 K^{-1}}$, where

$$q_{m, \Gamma}(x) = \frac{1}{((2\pi)^d \det(\Gamma))^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x - m) \cdot \Gamma^{-1}(x - m)\right), \quad x \in \mathbb{R}^d. \quad (4.25)$$

When $\delta > 0$ is small, the intuition would be to replace p_t by this limit in the evolution of m_t in (4.23) obtaining the approximation

$$\dot{m}_t \approx \delta \int_{\mathbb{R}^d} F(x + m_t) q_{0, \sigma^2 K^{-1}}(x) dx = \delta \int_{\mathbb{R}^d} F(x) q_{m_t, \sigma^2 K^{-1}}(x) dx, \quad (4.26)$$

which simply corresponds to replacing the right-hand side of (4.21) with its average with respect to a Gaussian measure centered in m_t , and slowing down the dynamics by a factor δ . The purpose of [L9] is make this approximation rigorous, and thus reducing drastically the dimension of the problem: one can look for structured dynamics for (4.19) at the level of the d -dimensional problem (4.26). In our case the manifold

$$\mathcal{M}^0 = \{(q_{0, \sigma^2 K^{-1}}, m) : m \in \mathbb{R}^d\} \quad (4.27)$$

is a stable manifold of stationary solutions for (4.23) with $\delta = 0$, and our aim is to prove that it persists in an invariant manifold \mathcal{M}^δ for $\delta > 0$ small, and that the phase dynamics on \mathcal{M}^δ can be approximated by (4.26) (more precisely we only prove the existence of positively invariant manifolds, see Theorem 4.3.3).

Set-up and main hypotheses

We work in essentially two functional settings, i.e. weighted L^2 and H^1 norms defined as $\|u\|_{L^2(w)} = \left(\int_{\mathbb{R}^d} |u(x)|^2 w(x) dx\right)^{\frac{1}{2}}$ and $\|u\|_{H^1(w)} = \left(\|u\|_{L^2(w)}^2 + \sum_{i=1}^d \|\partial_{x_i} u\|_{L^2(w)}^2\right)^{\frac{1}{2}}$, for weights

$$w_\theta(x) := \exp\left(\frac{\theta}{2} x \cdot K \sigma^{-2} x\right), \quad x \in \mathbb{R}^d \quad (4.28)$$

for different well chosen values of $\theta \in \mathbb{R}$. We give here only the main assumptions on F , under a simplified setting (see [L9] for more details): we suppose exponential moments on the initial condition of (4.19) and that there exist $c_F, C_F, r > 0$ and some (explicitly) sufficiently small $\varepsilon > 0$ such that the following holds

1. *One-sided Lipschitz continuity:*

$$(F(x) - F(y)) \cdot (x - y) \leq C_F |x - y|^2, \quad x, y \in \mathbb{R}^d. \quad (4.29)$$

2. *Confinement outside some ball:*

$$F(x) \cdot x \leq C_F \mathbf{1}_{\{|x| \leq r\}} - c_F |x|^2, \quad x \in \mathbb{R}^d \quad (4.30)$$

3. *The confinement controls the derivative:* the following limits holds:

$$\lim_{|x| \rightarrow \infty} \frac{|\partial_{x_k} F(x)|}{F(x) \cdot K \sigma^{-2} x} = 0 \text{ for } k = 1, \dots, d. \quad (4.31)$$

4. *Exponential control of the dynamics:*

$$\max \left(|F(x)|, \sup_{k=1, \dots, d} |\partial_{x_k} F(x)|, \sup_{k, l=1, \dots, d} \left| \partial_{x_k, x_l}^2 F^{(l)}(x) \right| \right) \leq C_F w_\varepsilon(x), \quad x \in \mathbb{R}^d. \quad (4.32)$$

Main result

The main result of [L9] is the following:

Theorem 4.3.3 (Th.2.3 and 2.5 in [L9]). *Under the previous assumptions, there exist a bounded open set $\mathcal{V} \subset \mathbb{R}^d$, $\alpha \in (0, 1)$ and $C > 0$ such that the following is true: for all $\delta \geq 0$ small enough, there exists a positively invariant manifold $\mathcal{M}^\delta = \{(p_m^\delta, m) : m \in \mathcal{V}\}$ for (4.18), where p_m^δ is a probability measure on \mathbb{R}^d for all $m \in \mathcal{V}$, and \mathcal{M}^δ is a perturbation of size δ of the manifold \mathcal{M}^0 in the following sense:*

$$\sup_{m \in \mathcal{V}} \|p_m^\delta - q_{0, \sigma^2 K^{-1}}\|_{L^2(w_\alpha)} \leq C \delta. \quad (4.33)$$

Furthermore, \mathcal{M}^δ is stable: there exist $\beta \in (0, \alpha)$, $\lambda, c, c', c'', C' > 0$ such that the following is true: if (p_0, m_0) satisfies $m_0 \in \mathcal{V}$, $\int_{\mathbb{R}^d} x p_0(dx) = 0$, $\int_{\mathbb{R}^d} w_\alpha(x + m_0) p_0(dx) \leq c$, $\|p_0 - q_{0, \sigma^2 K^{-1}}\|_{L^2(w_\alpha)} \leq c' \delta$ and $\|p_0 - p_{m_0}^\delta\|_{L^2(w_\beta)} \leq c'' \delta$, then for (p_t, m_t) the solution of (4.23) with initial condition (p_0, m_0) we have for all $t > 0$:

$$m_t \in \mathcal{V}, \quad \text{and} \quad \|p_t - p_{m_t}^\delta\|_{L^2(w_\beta)} \leq C' e^{-\lambda t} \|p_0 - p_{m_0}^\delta\|_{L^2(w_\beta)}. \quad (4.34)$$

Finally, the phase dynamics $t \mapsto m_t^\delta$ of the solution of (4.23) starting from $(p_{m_0}^\delta, m_0) \in \mathcal{M}^\delta$ has the following expansion in $\delta \rightarrow 0$

$$\dot{m}_t^\delta = \delta \int_{\mathbb{R}^d} F(u) q_{m_t^\delta, \sigma^2 K^{-1}}(u) du + \delta^2 g^\delta(m_t^\delta), \quad (4.35)$$

with $\|g^\delta\|_{C^1} \leq C$.

Consequences

The main conclusion of Theorem 4.3.3 is that, in the regime $\delta \rightarrow 0$, one can reduce (up to a time change of order δ) the analysis of the (infinite dimensional) mean-field system (4.18) to the d -dimensional synchronised system

$$\dot{x}_t = \int_{\mathbb{R}^d} F(u) q_{x_t, \sigma^2 K^{-1}}(u) du. \quad (4.36)$$

In (4.36), the presence of noise and interaction lies in the parameters $\varpi_1 = \sigma_1^2/k_1, \dots, \varpi_d = \sigma_d^2/k_d$ of the diagonal matrix $\sigma^2 K^{-1}$. Addressing Question 4.3.1 simply boils down to comparing the isolated system (4.21) (without noise and interaction) with the reduced dynamics (4.36) where noise and interaction is present.

Persistence of oscillations

Answering the question of *persistence* is easy: supposing here that (4.21) possesses a dynamical structure persistent under \mathcal{C}^1 -perturbation that is included in a bounded open set \mathcal{V}_0 with smooth boundaries (think of hyperbolic fixed-points, limit cycles, and more generally normally hyperbolic invariant manifolds [Fen72, Wig13], but also chaotic structures, as given by Lorenz-like flows [GW79, APV10]) classical convolution arguments show that the dynamics given by (4.36) is a \mathcal{C}^1 -perturbation (with perturbation of order $\max_i \varpi_i$) of (4.21), so that (4.36) admits

a similar persistent structure for $\max_i \varpi_i$ small enough. This observation holds for various examples that are extensively detailed in [L9, § 3.2] (e.g. the Stuart-Landau model [L9, eq. (3.3)]). Concerning the FHN case, starting from (1.29), the key point is to note that the convolution kernel driving the synchronised system (4.36) is simply

$$\int_{\mathbb{R}^2} F(u)q_{(x,y),\sigma^2K^{-1}}(u)du = \left((1 - \varpi_1)x - \frac{x^3}{3} - y, \frac{1}{c}(x + a - by) \right) \quad (4.37)$$

and so (4.37) defines again a FitzHugh-Nagumo system, with a modified prefactor $1 - \varpi_1$ in front of the x variable. Thus, if the parameters (a, b, c) in (1.29) are chosen so that the isolated system (4.21) is away from a bifurcation point (e.g. in the case of a limit cycle) and if $\varpi_1 \in [0, 1)$ is small enough, (4.36) will conserve the same type of dynamics: we have persistence of periodicity for the synchronised system (4.18), at least when the noise is small w.r.t. the interaction (as it had already been observed in [ABR04]). Similar persistence of periodic dynamics had already been obtained in [Sch86] for the mean-field Brusselator model, without any results on local stability of the solutions, using different techniques.

Emergence of structured dynamics

The question of *emergence* of oscillations is also intuitive: suppose that (4.21) exhibits a bifurcation and that a careful choice of parameters in the functional F brings (4.21) close to the bifurcation point. It may be that the introduction of the parameters $\varpi_1, \dots, \varpi_d$ in (4.36) makes the system cross this bifurcation point: we would then be precisely in a situation where the addition of noise and interaction induce a structured dynamical behavior that is not initially present in the IDS (4.21) (see [L9, § 3.3] for detailed applications to Stuart-Landau oscillators, [L9, Th. 3.1] and the Cucker-Smale model for collective dynamics [BCnCD16], see [L9, Th. 3.4]). For FHN oscillators, take (1.29) with parameters $a = \frac{1}{3}$, $b = 1$, $c = 10$. This is an excitable case when (4.21) has a unique stationary point $(x, y) = (-1, -\frac{2}{3})$. The transitions of (4.36) driven by

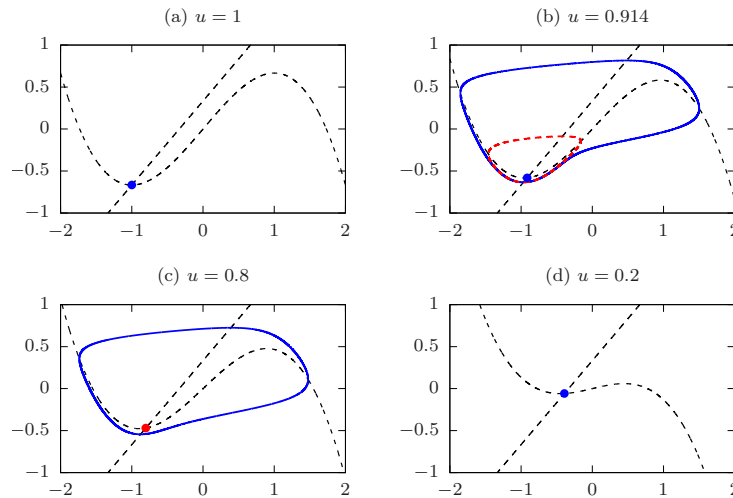


Figure 4.2 – Phase diagrams for the dynamics driven by (4.36) for parameters $a = \frac{1}{3}$, $b = 1$, $c = 10$ and different choices of $u := 1 - \varpi_1$. Stable (resp. unstable) points and limit cycles are represented in blue (resp. red). The nullclines $y = ux - \frac{x^3}{3}$ and $y = \frac{x+a}{b}$ are represented in black dashed lines. Starting from the fixed-point dynamics of $u = 1$ ($\varpi_1 = 0$, no noise, case (a)), a saddle-node bifurcation of cycles then occurs (numerically estimated at $u \approx 0.91435$) after which a stable point and a stable cycle coexist, separated by an unstable cycle (case (b)). Then, at $u \approx 0.88604$ the stable point and the unstable cycle collide in a subcritical Andronov-Hopf bifurcation. The dynamics is then given by a limit cycle surrounding an unstable point (case (c)), until the supercritical Andronov-Hopf bifurcation at $u \approx 0.28383$, after which the dynamics is again given by a fixed-point (case (d)).

(4.37) from $u := 1 - \varpi_1 = 1$ ($\varpi_1 = 0$, no noise) to larger values of the ratio ϖ_1 are described in Figure 4.2. While the whole description of Figure 4.2 (especially the saddle node bifurcation of

cycles (a)→(b)) is based on numerical simulations, it is possible to give a rigorous proof of the Andronov-Hopf bifurcations ((b)→(c) and (c)→(d) in Figure 4.2) in the following Theorem 4.3.4:

Theorem 4.3.4 (Emergence of oscillations for the FHN model, see Th. 3.3 in [L9]). *Suppose that $d = 2$ and that F is given by (1.29), for the choice of $a = \frac{1}{3}, b = 1, c = 10$. Then Theorem 4.3.3 is valid and there exist $0 < \varpi'_1 < \varpi''_1$ at which the averaged dynamics (4.36) undergoes an Andronov-Hopf bifurcation (that is subcritical in ϖ'_1 and supercritical in ϖ''_1). In particular, there exist values of ϖ_1 such that, for δ small enough, the PDE (4.18) admits a periodic solution whereas the isolated system (4.21) converges to its unique stable fixed-point $(-1, -\frac{2}{3})$.*

Remark 4.3.5. *Theorem 4.3.4 states the existence of a limit cycle for (4.18) for δ small. Let us be more precise on the dependence of this limit cycle in δ : for the prescribed values of the parameters a, b, c and the interaction/noise ratio ϖ_1 (note that the value of ϖ_1 is independent of δ), the system (4.36) (independent of δ) has a limit cycle, say Γ^0 . Then*

$$\dot{y}_t^\delta = \int_{\mathbb{R}^d} F(u) q_{y_t^\delta, \sigma^2 K^{-1}}(u) du + \delta g^\delta(y_t^\delta) \quad (4.38)$$

is a \mathcal{C}^1 -perturbation of (4.36): there exists $\delta^* > 0$ such that for $\delta \in [0, \delta^*)$, (4.38) admits a limit cycle Γ^δ (perturbation of Γ^0). But then the slow-downed version of (4.38) $m_t^\delta := y_{\delta t}^\delta$ solves (4.35) and admits the same limit cycle Γ^δ . In other words, the limit cycle Γ^δ we observe for (4.35) (and hence for (4.18)) for $\delta > 0$ is a perturbation of order δ (covered at small speed δ) of a pre-existing limit cycle Γ^0 , see Figure 4.3. The fact that one tends to a trivial (constant) dynamics for (4.18) as $\delta \rightarrow 0$ does not come from the fact that this limit cycle Γ^δ disappears as $\delta \rightarrow 0$, this is only due to the fact that the speed of (4.35) along $\Gamma^\delta \approx \Gamma^0$ vanishes as $\delta \approx 0$.

The existence of oscillations for (4.36) translates into similar oscillations for the system (4.18), see Figure 4.4.

Main lines of proof of Theorem 4.3.3

The main arguments for the proof of Theorem 4.3.3 rely on the same normally hyperbolic structure as developed in Section 4.1.3. The main point is to construct the invariant manifold \mathcal{M}^δ for (4.18) as $\mathcal{M}^\delta = (Id + f)(\mathcal{M}^0)$ where f is some distance functional satisfying an appropriate fixed-point relation in a similar fashion as in § 4.1.3. First mention a technical point that differs from the case of phase oscillators which concerns the control of the dynamics at infinity: one needs to make sure that, starting from a neighborhood of \mathcal{M}^0 , the dynamics of (4.18) sufficiently discriminates between trajectories with means m_t that remain in a sufficiently large ball from the ones with mean that eventually escape from this ball. This is done by artificially modifying the dynamics (4.18) outside a sufficiently large ball in \mathbb{R}^d , an operation that is transparent to the existence of the manifold \mathcal{M}^δ due to the confining properties of F , see [L9, Lemma 5.9] for a precise statement.

Let us now deal with our main problem, that is the control of the unbounded terms involving F in (4.18). The first remark is that under the dissipativity assumption of F (4.29), we have existence of exponential moments for (4.19) [L9, Lem. 4.1]: for any $T > 0, \alpha \in (0, 1)$, supposing that $\mathbf{E}(w_\alpha(X_0))$ (recall (4.28)), then there is some $\kappa_0 > 0$ such that $\sup_{t \in [0, T]} \mathbf{E}(w_\alpha(X_t)) \leq \max(\kappa_0, \mathbf{E}(w_\alpha(X_0)))$. This estimate is crucial in order to properly control the behavior of F in infinity (recall (4.32)). As for § 4.1.3, the main point is to control both the proximity of a solution p_t to (4.23) to the Gaussian $q_{0, \sigma^2 K^{-1}}$ (recall (4.33)) as well as the dependence of (4.23) w.r.t. its initial condition. In order to illustrate the difficulty of working with unbounded F on \mathbb{R}^d , let us make a typical calculation: compare two solutions, say p_t^1, p_t^2 , to (4.23) with close initial conditions p_0^1 and p_0^2 . The main starting point is similar to (4.11): for $F_t^i = F(\cdot + m_t^i)$ and $F_t^{12} := F_t^1 - F_t^2$, recalling the definition of the weight w in (4.28), the difference $\pi_t := p_t^1 - p_t^2$ verifies

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|\pi_t\|_{L^2(w_\beta)}^2 &= \langle \mathcal{L}\pi_t, \pi_t \rangle_{L^2(w_\beta)} - \delta \langle \nabla \cdot (\pi_t (F_t^1 - \langle \mu_t^1, F \rangle)), \pi_t \rangle_{L^2(w_\beta)} \\ &\quad - \delta \langle \nabla \cdot (p_t^2 (F_t^{12} - \langle \mu_t^1 - \mu_t^2, F \rangle)), \pi_t \rangle_{L^2(w_\beta)}. \end{aligned} \quad (4.39)$$

Here, $\mathcal{L}u := \mathcal{L}_\theta u$ for $\theta = 1$ where

$$\mathcal{L}u = \nabla \cdot (\sigma^2 \nabla u) + \theta \nabla \cdot (Kxu) \quad (4.40)$$

is the standard Ornstein-Uhlenbeck operator. It is well-known that such an operator \mathcal{L} is dissipative in $L^2(w_\theta)$ for the choice $\theta = 1$ (recall the definition of w_θ in (4.28)). As it will be clear below, an essential point of the proof is to obtain the dissipativity of \mathcal{L} in $L^2(w_\theta)$ for a large choice of θ , not only for $\theta = 1$: we have (see [L9, Lem. 4.8]) that for any $u \in L^2(w_\theta)$ such that $\int u(x) dx = 0$, for $\theta \in (0, 1)$,

$$\langle \mathcal{L}u, u \rangle_{L^2(w_\theta)} \leq -\frac{k_\theta}{\theta(\text{Tr}(K) + \underline{k})} \|\sigma \nabla u\|_{L^2(w_\theta)}^2 \leq -k_\theta \|u\|_{L^2(w_\theta)}^2, \quad (4.41)$$

for $k_\theta := \theta \underline{k} - \frac{1-\theta}{2} \text{Tr} K$, with $\underline{k} := \min_i k_i$. Inequality (4.41) is the generic dissipativity estimate to deal with the first term in the right-hand side of (4.39): one has $\langle \mathcal{L}\pi_t, \pi_t \rangle_{L^2(w_\beta)} \leq -\frac{k_\beta}{\beta(\text{Tr}(K) + \underline{k})} \|\sigma \nabla \pi_t\|_{L^2(w_\beta)}^2$. The second term in the right-hand side of (4.39) can be dealt directly using the hypotheses made on F : by integration by parts, $-\delta \langle \nabla \cdot (\pi_t(F_t^1 - \langle \mu_t^1, F \rangle)) , \pi_t \rangle_{L^2(w_\beta)} = \frac{\delta}{2} \int |\pi_t|^2 (\beta F_t^1 \cdot K \sigma^{-2} x - \nabla \cdot F_t^1 - \beta \langle \mu_t^1, F \rangle \cdot K \sigma^{-2} x) w_\beta \leq C\delta \|\nabla \pi_t\|_{L^2(w_\beta)}^2$, using (4.30) and (4.31). The necessity of working with general parameters β and θ in both (4.39) and (4.41) comes from the difficulty in handling the remaining term in (4.39). The technical details left apart, calculations give an upper bound in terms of $C\delta \|\nabla \pi_t\|_{L^2(w_\beta)} \|G\|_{L^2(w_\beta)}$, where G is generically written as $G = H(p_t^1, p_t^2)F$ for some explicit smooth functional H . The issue is that the only bound one has on F is exponential $|F| \leq C_F w_\varepsilon$, recall (4.32). Hence, the only possible bound is $\|G\|_{L^2(w_\beta)} \leq C \left(\|p_t^1\|_{L^2(w_\alpha)} + \|p_t^2\|_{L^2(w_\alpha)} \right)$ as long as $\beta + 2\varepsilon < \alpha$. It is therefore necessary to work under an a priori control on the $L^2(w_\alpha)$ -norm of p_t^1 and p_t^2 . This a priori estimate is essentially contained in (4.33) and proven in [L9, Lem. 5.3] using the dissipativity hypotheses on F and the exponential moments on X_t . Finally putting all these considerations together in (4.39), one obtains then, using the Poincaré inequality (4.41)

$$\frac{d}{dt} \|\pi_t\|_{L^2(w_\beta)}^2 \leq -(k_\beta - C\delta) \|\pi_t\|_{L^2}^2 + C\delta \sup_{s \leq t} |m_s^1 - m_s^2|^2. \quad (4.42)$$

Choosing now $\delta > 0$ sufficiently small so that $\chi(\delta) := k_\beta - C\delta > 0$, one obtains from standard extensions of Grönwall inequalities ([L9, Lem. A.2]) a typical estimate of the form

$$\begin{cases} \|p_t^1 - p_t^2\|_{L^2(w_\beta)}^2 & \leq \max \left(\|p_0^1 - p_0^2\|_{L^2(w_\beta)}^2, C\delta \sup_{s \leq t} |m_s^1 - m_s^2|^2 \right), \\ \|p_t^1 - p_t^2\|_{L^2(w_\beta)}^2 & \leq e^{-\chi(\delta)t} \|p_0^1 - p_0^2\|_{L^2(w_\beta)}^2 + C\delta \sup_{s \leq t} |m_s^1 - m_s^2|^2. \end{cases} \quad (4.43)$$

The control of the part concerning the mean-values $|m_s^1 - m_s^2|$ is treated separately, with similar techniques. Hence, the previous calculation tells us that it is necessary to perform the claimed fixed-point procedure on a adapted version of the distance functional X_t (as in § 4.1.3) using both spaces $L^2(w_\alpha)$ and $L^2(w_\beta)$. We refer to [L9, § 5.3] for more details. At this point, a second important issue remains: once the fixed-point procedure carried-out, one only obtains the existence of a SNHM \mathcal{M}^δ in a δ -neighborhood of \mathcal{M}^0 . However, in order to obtain parameterisation (4.35), one needs to make sure that \mathcal{M}^δ has sufficient regularity (i.e. the fact that the previous fixed-point f is \mathcal{C}^1), which is not ensured as this point of the proof. A significant part of the work of [L9] is to obtain this required regularity, see [L9, § 6]: the strategy is first to establish a formal equation that the derivative of f should satisfy, secondly to perform a similar fixed-point procedure for this derivative and finally to identify formally this second fixed-point as the derivative of f . With this at and, (4.35) is then a simple by-product of the construction of f .

4.4 FHN model: the kinetic case

We now review [L11] where we analyse the emergence of similar periodicity patterns in (4.18) for the FHN model in the kinetic case: that is $d = 2$, F given by (1.29) with $k_1, \sigma_1 > 0$ but with

$k_2 = \sigma_2 = 0$. This corresponds to a modelling more closely related to biological applications as the recovery variable w in (1.29) captures nonlinear mechanisms that are intrinsic to each neuron, so that imposing interaction between recovery variables between neurons does not really make sense from a biological perspective. The McKean-Vlasov diffusion (4.19) here becomes

$$\begin{cases} dX_t = \delta \left(X_t - \frac{X_t^3}{3} - Y_t \right) dt - K(X_t - \mathbb{E}[X_t]) dt + \sqrt{2}\sigma dB_t \\ dY_t = \frac{\delta}{c} (X_t + a - bY_t) dt \end{cases}, \quad t \geq 0, \quad (4.44)$$

Let us sum up here the main heuristics, that relies again on a slow-fast approach: under Hypothesis 4.3.2, for δ small, in (4.44), the interaction and noise terms constitute a fast part of the dynamics, while the intrinsic dynamics term δF constitutes a slow one. When $\delta = 0$, the fast dynamics part of (4.44) simply reduces an Ornstein-Uhlenbeck process (of constant expectation) with Gaussian invariant measure. The only difference is that this Gaussian has singular variance w.r.t. the Y_t -coordinate. When δ is now positive but small, the distribution of X_t should be approximated (at first order in δ) by $\mathcal{N}(\mathbb{E}[X_t], \sigma^2/K)$, where $\mathbb{E}[X_t]$ evolves slowly in time. Now if X_t is at first order Gaussian, so is Y_t since its dynamics is linear. So, at first order in δ , (X_t, Y_t) should have a Gaussian distribution $\mathcal{N}(m_t, \Gamma_\delta)$, where Γ_δ is a symmetric covariance matrix. Obtaining the first order in δ for $m_t = (\mathbb{E}[X_t], \mathbb{E}[Y_t])$ is easy: taking the expectation in (4.44), m_t satisfies $\dot{m}_t = \delta \left(\begin{array}{c} \mathbb{E}[X_t] - \frac{\mathbb{E}[X_t^3]}{3} - \mathbb{E}[Y_t] \\ \frac{1}{c} (\mathbb{E}[X_t] + a - b\mathbb{E}[Y_t]) \end{array} \right)$ and considering that $X_t \approx \mathcal{N}(x_t, \sigma^2/K)$, we obtain $\mathbb{E}[X_t^3] \approx \mathbb{E}[X_t]^3 + 3\frac{\sigma^2}{K}$, which leads to the same dynamics (4.36) for the expectation as in the elliptic case. To compute Γ_δ , denote by (Z_t^x, Z_t^y) as the first order approximation of $(X_t - \mathbf{E}(X_t), Y_t - \mathbf{E}(Y_t))$ defined as

$$\begin{cases} dZ_t^x &= -KZ_t^x dt + \sqrt{2}\sigma dB_t, \\ dZ_t^y &= \frac{\delta}{c} (Z_t^x - bZ_t^y) dt. \end{cases} \quad (4.45)$$

Straightforward covariance considerations show that the equilibrium of (4.45) is $\mathcal{N}(0, \Gamma_\delta)$ with Γ_δ necessarily given by

$$\Gamma_\delta = \frac{\sigma^2}{K} \begin{pmatrix} 1 & \frac{\delta}{Kc+b\delta} \\ \frac{\delta}{Kc+b\delta} & \frac{\delta}{b(Kc+b\delta)} \end{pmatrix}. \quad (4.46)$$

4.4.1 Main results

In view of the analysis made for Theorem 4.3.4, fix the parameters (a, b, c, σ, K) in (1.29) such that the system (4.36) admits a (stable) limit cycle. Note that the existence of such limit cycle does not depend on δ : if, for $\delta = 1$, this stable periodic solution is $(\gamma_t)_{t \in [0, T_\gamma]}$ (with period T_γ), the corresponding periodic orbit for $\delta > 0$ becomes $(\gamma_t^\delta) := (\gamma_{\delta t})_{t \in [0, T_\gamma/\delta]}$, with period $\frac{T_\gamma}{\delta}$. The previous heuristics suggest that one should compare any solution μ_t of (4.18) to Gaussian distributions driven by the above limit cycle $(\gamma_t)_{t \in [0, T_\gamma]}$. More precisely, denote q_m the Gaussian distribution with mean m and covariance Γ_δ given by (4.46),

$$q_m(z) = q_m^\delta(z) := \frac{1}{2\pi \det(\Gamma_\delta)} \exp\left(\frac{1}{2}(z - m) \cdot \Gamma_\delta^{-1}(z - m)\right), \quad z \in \mathbb{R}^2, \quad (4.47)$$

and define the Gaussian limit cycle

$$\mathcal{G}_\delta := \left\{ q_{\gamma_t^\delta}^\delta : t \geq 0 \right\}. \quad (4.48)$$

Of course, \mathcal{G}_δ is not a limit cycle for (4.18) (it is only approximately invariant for the dynamics, see § 4.5 below). The point of [L11] is to prove that there exists a limit cycle \mathcal{C}_δ for (4.18) in a neighborhood of \mathcal{G}_δ . Proximity is here measured in terms of an adequate Wasserstein-type distance (more details below). The main result of [L11] is the following:

Theorem 4.4.1. [Th. 1.4 in [L11]] *Choose parameters a, b, c, K and σ^2 such that (4.36) admits an stable limit cycle. Then there exists a $\delta_c > 0$ such that for all $\delta \leq \delta_c$ there exists a periodic*

solution ν_t^{per} to (4.18), defining an invariant cycle \mathcal{C}_δ which satisfies $\sup_t \text{dist}_{W_2}(\nu_t^{per}, \mathcal{G}_\delta) = O(\delta)$. Moreover there exist $C_1, C_2 > 0$ that do not depend on δ , $C(\delta) > 0$ and a rate $\lambda(\delta) > 0$ such that if $\mu_0 \in \mathcal{P}_2$ satisfies

$$\int_{\mathbb{R}^2} z^6 d\mu_0(z) \leq C_1, \quad \text{and} \quad \text{dist}_{W_2}(\mu_0, \mathcal{C}_\delta) \leq C_2 \delta, \quad (4.49)$$

then for $t \mapsto \mu_t$ the solution to (4.18) with initial condition μ_0 we have:

$$\text{dist}_{W_2}(\mu_t, \mathcal{C}_\delta) \leq C(\delta) e^{-\lambda(\delta)t} \text{dist}_{W_2}(\mu_0, \mathcal{C}_\delta). \quad (4.50)$$

4.4.2 Elliptic vs kinetic: main difficulties and lines of proof

The main difficulty of Theorem 4.4.1 is of course the lack of control one has on the Y -coordinate in (4.44). In particular, it is impossible to derive Theorem 4.4.1 from its elliptic counterpart Theorem 4.3.3 while making $k_2, \sigma_2 \rightarrow 0$: stability analysis leading to Theorem 4.3.3 crucially requires the interaction and noise to be nontrivial on every coordinates. One can however note that the dynamics of m_t in the elliptic case does not depend on k_2, σ_2 (at least at first order in δ): the system (4.36) only involves the ratio $\varpi_1 = \frac{\sigma_1^2}{k_1}$. It is therefore not surprising to retrieve the same dynamics for m_t in the kinetic case. This implies in particular that one observes the emergence of oscillations for the same parameters as in the elliptic case. The only signature of the lack of ellipticity in the kinetic case lies in the covariance structure of the problem: due to the absence of noise on the Y -coordinate, the variance along the Y -direction is here significantly smaller than in the elliptic case and Γ_δ given in (4.46) becomes singular as $\delta \rightarrow 0$ (see Figure 4.5).

This consideration comes along with a technical issue: the lack of control on Y in the kinetic case prevents us from obtaining some strong L^2 -control as one has in the elliptic case: there is no similar spectral gap for the Ornstein-Uhlenbeck operator as in (4.41) in the kinetic case. We circumvent this difficulty by working with a weaker topology, that is with some suitable Wasserstein metric, that is adapted to our needs: if $\mathcal{P}_2 = \mathcal{P}_2(\mathbb{R}^2)$ is the set of probability measures with second finite moment, for $\beta \in (0, 1)$ and $b > 0$, define for $\nu_1, \nu_2 \in \mathcal{P}_2$,

$$W(\nu_1, \nu_2) = \max \left(\delta^\beta \left| \int_{\mathbb{R}^2} z \nu_1(dz) - \int_{\mathbb{R}^2} z \nu_2(dz) \right|, \inf \sqrt{\mathbf{E}[|\tilde{X}_1 - \tilde{X}_2|^2]}, b \inf \sqrt{\mathbf{E}[|\tilde{Y}_1 - \tilde{Y}_2|^2]} \right), \quad (4.51)$$

where the infimum is taken over all couplings (X_i, Y_i) , $i = 1, 2$ such that $(X_i, Y_i) \sim \nu_i$ and where $\tilde{X} := X - \mathbf{E}(X)$. This decomposition into mean-value and centered process allows us to treat the dynamics of the centered process solution to (4.18) and the dynamics of its expectation in a relatively separate way. The point of the proof of Theorem 4.4.1 is again to proceed with contraction arguments of the dynamics of (4.18) for δ small. The control of the expectation m is ensured by the local stability of the limit cycle $(\gamma_t)_{t \in [0, T_\gamma)}$ of (4.36) (via Floquet estimates), whereas the control on the centered variables $(\tilde{X}_t, \tilde{Y}_t)$ is made via a two-step procedure: first the exponential stability on the X -coordinate and second the fact that this stability propagates to the Y variable through the linearity of the interaction. The calculations in [L11] being quite technical, let us only illustrate the use of the Wasserstein metric (4.51) on a preliminary result (see [L11, Lem. 2.4, Lem. 3.2]), expressing the proximity of the centered measure $\tilde{\mu}_t$ with the Gaussian q_0 defined in (4.47): there exist $\delta_1 > 0$, $\kappa_0, \kappa_1 > 0$ such that whenever $\int (x^6 + y^6) \mu_0(dx, dy) \leq \kappa_0$, if $\delta \in (0, \delta_1)$ and $W(\tilde{\mu}_0, q_0) \leq \kappa_1 \delta$ then

$$\sup_{t \geq 0} W(\tilde{\mu}_t, q_0) \leq \kappa_1 \delta. \quad (4.52)$$

Indeed, for any $\varepsilon > 0$, take a coupling $\{(\tilde{X}_0, \tilde{Y}_0), (Z_0^x, Z_0^y)\}$ between $\tilde{\mu}_0$ and q_0 such that $\mathbf{E}(|\tilde{X}_0 - Z_0^x|^2) + \mathbf{E}(|\tilde{Y}_0 - Z_0^y|^2) < W(\tilde{\mu}_0, q_0)^2 + \varepsilon$. Then Ito formula applied to the (centered version of the) first components of (4.44) and (4.45) gives $\frac{1}{2} d(\tilde{X}_t - Z_t^x) = -K(\tilde{X}_t - Z_t^x)^2 + (\tilde{X}_t^x - Z_t^x) \left(\delta \left(X_t - \frac{X_t^3}{3} - Y_t \right) - \dot{x}_t \right)$, with $m_t = (x_t, y_t)$. A priori bounds on X and x show that

$\mathbf{E} \left(\left(\delta \left(X_t - \frac{X_t^3}{3} - Y_t \right) - \dot{x}_t \right)^2 \right)^{\frac{1}{2}} \leq C\delta$, using the control on moments of order 6 we have assumed on the initial condition. Hence one obtains

$$\frac{1}{2} \frac{d}{dt} \mathbf{E} \left((\tilde{X}_t - Z_t^x)^2 \right) \leq C\delta \left(\mathbf{E} \left((\tilde{X}_t - Z_t^x)^2 \right) \right)^{\frac{1}{2}} - K \mathbf{E} \left((\tilde{X}_t - Z_t^x)^2 \right). \quad (4.53)$$

Then a Grönwall argument shows that for δ small, $\sup_{t \geq 0} \mathbf{E} \left((\tilde{X}_t - Z_t^x)^2 \right) \leq C\delta + \varepsilon$. But then this estimates propagates to the Y - variable of (4.44): since $\frac{c}{2\delta} d(\tilde{Y}_t - Z_t^y)^2 = (\tilde{Y}_t - Z_t^y) (\tilde{X}_t - Z_t^x) dt - b(\tilde{Y}_t - Z_t^y)^2 dt$, plugging the above estimate on \tilde{X} , we obtain in a same way that $\sup_{t \geq 0} \mathbf{E} \left((\tilde{Y}_t - Z_t^y)^2 \right) \leq \frac{1}{b^2} (C\delta^2 + \varepsilon)$. Estimates are in particular true for the infimum over all couplings $W(\tilde{\mu}_t, q_0)$ and letting $\varepsilon \rightarrow 0$ gives the result (4.52). Note we have not been precise here on the control on the constants C appearing above, but the dissipativity on the X -variable in (4.53) indeed ensures that once $W(\tilde{\mu}_0, q_0) \leq \kappa_1 \delta$ then (4.52) is indeed true for the *same constant* κ_1 . We refer to [L11, Lem. 2.4] for details.

The prefactor δ^β in the definition of the distance W accounts for the typical difference of scaling between the dynamics of the expectation m_t and the recentered version $\tilde{\mu}_t$ of a solution μ_t to (4.18): a generic situation to have in mind is the case of some distribution μ_0 with mean m_0 such that $|m_0 - \gamma_\theta^\delta|$ of order $\delta^{1-\epsilon}$ ($\epsilon > 0$) and $W(\tilde{\mu}_0, q_0)$ of order δ (and hence smaller than $|m_0 - \gamma_\theta^\delta|$). The presence of δ^β in the definition of W is here to ensure that $W(\mu_0, q_{\gamma_\theta^\delta})$ remains of order δ : in other words, the contraction of m around γ^δ given by the approximation (4.36) is not too much perturbed by the dynamics of $\tilde{\mu}_t$. Note however that this choice of rescaling is mostly transparent as W is equivalent to the standard Wasserstein-2 distance W_2 : for δ small enough $cW \leq W_2 \leq 3\delta^{-\beta}W$. We refer to [L11] for more details.

4.5 On regularity of limit cycles and isochrons

4.5.1 Isochrons maps

Theorem 4.3.3 addresses the existence of a stable limit cycle related to the PDE (4.18)). A related question concerns the existence of pseudo-periodic behaviors for the empirical measure $\mu_{n,t}$ of the particle system (4.20) (pseudo in the sense that there cannot be any periodic behavior for (4.20) as recalled in § 1.3.1). This issue will be the purpose of Chapter 5. Having in mind the possible connections between the PDE (4.18) and its microscopic counterpart (4.20), one may find Theorem 4.3.3 unsatisfying for several reasons. A technical issue first: the existence of limit cycles for (4.18) is stated in a L^2 framework, that is perfectly suitable for the study of the smooth probability-valued process μ_t solution to (4.18) but is certainly not adapted to its microscopic companion $\mu_{n,t}$. We circumvent this by working rather in some spaces more natural for empirical processes, that is suitably weighted Sobolev space with negative index H^{-k} .

Secondly and more importantly, as we will see in Chapter 5, the study of $\mu_{n,t}$ requires to have a precise understanding of the regularity of the trajectories of (4.18) in a neighborhood of the limit cycle. The analysis of Chapter 5 relies on the existence of a smooth (mostly C^2 , as we will want to apply Ito's formula) *isochron map* related to (4.18), which is not addressed by Theorem 4.3.3. Let us briefly recall informally the basic notions here, which relates to Floquet Theory, originally set in finite dimension (see [Tes12, HPS77, GW79]): suppose that some dynamics $\dot{x}_t = F(x_t)$ in \mathbb{R}^d , with F smooth, admits a stable limit cycle \mathcal{M} , with period T . Introduce the linearized equation around some $q_t \in \mathcal{M}$, $\dot{z}_t = DF(q_t)z_t$. Then, it can be shown that the principal matrix solution Π associated to the periodic solution q , solving $\partial_t \Pi(t, s) = DF(q_t) \Pi(t, s)$ can be written as $\Pi(s+t, s) = N(s+t, s) e^{-tQ(s)}$, with $t \mapsto N(s+t, s)$, T -periodic. Any $Q(s)$ and $Q(t)$ are similar and hence have the same spectrum. Hyperbolic stability of \mathcal{M} can be expressed in terms of the spectrum of $Q(0)$: 0 is an eigenvalue to $Q(0)$ with multiplicity 1 and the rest of the spectrum has strictly negative real part. Denoting as $\Phi(x, t)$ as the solution x_t with $x_0 = x$, this stability can be equivalently expressed as $\lim_{t \rightarrow \infty} e^{\gamma t} \text{dist}(\Phi(x, t), \mathcal{M}) = 0$, as long as $\gamma < \gamma_0$, where γ_0 is the spectral gap of $Q(0)$. In this setting, the *isochron* $W(x)$ of

a point $x \in \mathcal{M}$ is defined as $W(x) := \left\{ y \in \mathbb{R}^d, \text{dist}(\Phi(x, t), \Phi(y, t)) \xrightarrow[t \rightarrow \infty]{} 0 \right\}$: $W(x)$ consists of all initial values leading to a trajectory that ultimately approaches the one initiated at x on the limit cycle. Standard arguments (see e.g. [HPS77]) show that for every $x \in \mathcal{M}$, $W(x)$ is of dimension $(d-1)$, is transverse to $F(x)$ at x , with the same regularity and $\{W(x), x \in \mathcal{M}\}$ is a foliation of the stable manifold W of \mathcal{M} . One can finally define the *isochron map*: for every $y \in \mathbb{R}^d$, define $\theta(y)$ as the unique $t \in [0, T[$ such that $y \in W(q_t)$. In other words, $\theta : W \rightarrow \mathbb{R} \bmod T$ gives a notion of phase along the limit cycle for every point in W .

This notion of isochron was extended to an infinite dimensional setting: in the present result [L12], we rely heavily on an abstract result developed by Bates, Lu and Zheng [BLZ98], in the context of SNHM as described in § 4.1.1. It is shown in [BLZ98] that if a system admits a manifold that is approximately invariant and approximately normally hyperbolic (in a way to be precised later), then the system possesses an actual normally hyperbolic invariant manifold in a neighborhood of the approximately invariant one. As in the finite dimensional setting, it is in addition proven in [BLZ98] that the stable manifold of the actual SNHM (in our case \mathcal{M} is attractive, the stable manifold is in fact a neighborhood \mathcal{W} of \mathcal{M}) is foliated by invariant foliations: $\mathcal{W} = \cup_{m \in \mathcal{M}} \mathcal{W}_m$ and that there exists an isochron map $\Theta : \mathcal{W} \rightarrow \mathbb{R} \bmod T$. Note however that the deep general result of [BLZ98] only ensures that Θ is Hölder continuous, which is not entirely satisfactory in view of the analysis of Chapter 5: we will require Θ to be C^2 . One challenge of [L12] is precisely to ensure that Θ will have the required regularity.

4.5.2 Main hypotheses and settings

Transposed to our setting of (4.18), one main drawback of the method developed in [BLZ98] is that it requires the function F and its derivatives to be bounded (hypothesis that is not met in the case of the FitzHugh Nagumo model). However replacing F in (4.18) by some $\tilde{F} := x \mapsto F(x)\psi(\varepsilon|x|)$, \tilde{F} becomes bounded and the existence of a limit cycle for (4.18) with \tilde{F} will imply the same for F .

Empirical measures within weighted Sobolev spaces

The point now is to work in Sobolev spaces with negative indices: we consider again norms with Gaussian weights w_θ given by (4.28) for various choices of θ , supposing ellipticity of the coefficients: $k_i, \sigma_i > 0$ for all $i = 1, \dots, d$. Denote again as $L_\theta^2 = L^2(w_\theta)$ and $\mathcal{L}_\theta^* f = \nabla \cdot (\sigma^2 \nabla f) - \theta K x \cdot \nabla f$ the adjoint of the Ornstein-Uhlenbeck operator (4.40). It is well-know (see for example [BGL14]) that \mathcal{L}_θ^* is diagonalisable $\mathcal{L}_\theta^* \psi_l = -\lambda_l \psi_l$ along the orthonormal basis $(\psi_l)_{l \in \mathbb{N}^d}$ given for all $l \in \mathbb{N}^d$ by $\psi_l(x) := \psi_{l, \theta}(x) = \prod_{i=1}^d h_{l_i} \left(\sqrt{\frac{\theta k_i}{\sigma_i^2}} x_i \right)$ with $\lambda_l = \theta \sum_{i=1}^d k_i l_i$. Here h_n is the n^{th} renormalized Hermite polynomial: $h_n(x) = \frac{(-1)^n}{\sqrt{n!} (2\pi)^{\frac{1}{4}}} e^{\frac{x^2}{2}} \frac{d^n}{dx^n} \left\{ e^{-\frac{x^2}{2}} \right\}$. For f, g with decompositions $f = \sum_{l \in \mathbb{N}^d} f_l \psi_l$ and $g = \sum_{l \in \mathbb{N}^d} g_l \psi_l$, define the scalar products

$$\langle f, g \rangle_{H_\theta^r} = \langle (a_\theta - \mathcal{L}_\theta^*)^r f, \bar{g} \rangle_{L_\theta^2} = \sum_{l \in \mathbb{N}^d} (a_\theta + \lambda_l)^r f_l \bar{g}_l, \quad (4.54)$$

where $a_\theta = \theta \text{Tr} K$ (this choice of a_θ is made to simplify some technical proofs in [L12], another choice of positive constant would produce an equivalent norm) and denote by H_θ^r the completion of the space of smooth function u satisfying $\|u\|_{H_\theta^r} < \infty$. We work in [L12] in H_θ^{-r} the dual of H_θ^r , which is in particular a proper space to deal with empirical measures: any Dirac mass δ_x in \mathbb{R}^d belongs to H_θ^{-r} as long as $r > d/2$. Note here that special attention has to be put on a technical issue: the control on the norm of δ_x is exponential in x , $\|\delta_x\|_{H_\theta^{-r}} \leq C w_{\frac{\theta}{4-\eta}}(x)$ for $\eta > 0$ [L15, Lemma 2.1] (this is a significant difficulty compared to the case of phase oscillators, where the compactness of the state space induces a uniform bound $\sup_{x \in \mathbb{T}} \|\delta_x\|_{H_\theta^{-r}} \leq C$, which is a major simplification to the analysis). Note here another significant technical issue: anticipating on Chapter 5 and [L15] which addresses the behavior of the particle system (4.20) $(X_{i,t})_{i=1, \dots, N}$, a key argument in [L15] will be to obtain bounds on exponential moments of the form $\mathbb{E}[w_{m\theta}(X_{i,t})]$ with m possibly very large. Therefore, we will have to consider small values of θ in order to

maintain proper controls on the previous expectation, so that we need to work in H_θ^{-r} for general $\theta > 0$ and not only for the standard choice $\theta = 1$. A great deal of effort is spend in [L12] in order to look at the dependence in various parameters θ, θ' of the contraction properties of the Ornstein-Uhlenbeck operator $\mathcal{L}_{\theta'}$ in H_θ^{-r} (see [L12, App. A] for more details): a typical result and the key to the present analysis is the existence of $C > 0$ such that for r sufficiently large (independent of θ), for $0 < \theta \leq \theta'$, for any $\lambda < \theta \min(k_1, \dots, k_d)$, for all $u \in H_\theta^{-r}$ such that $\int u = 0$,

$$\|e^{t\mathcal{L}_{\theta'}}\|_{H_\theta^{-r}} \leq Ce^{-\lambda t} \|u\|_{H_\theta^{-r}}. \quad (4.55)$$

We refer to [L12, Prop. A.3] for other estimates and a (particularly technical) proof of these results.

4.5.3 Main results

The starting point is to suppose that the ordinary differential equation (4.36) admits a stable periodic solution $(\gamma_t^\delta)_{t \in [0, \frac{T}{\delta}]}$ that is hyperbolically stable (in the sense of Floquet formalism as previously stated). Under this hypothesis, $(\gamma_t^\delta)_{t \in [0, \frac{T}{\delta}]}$ is a simple example of a d -dimensional SNHM in the sense of § 4.1.2. Rewrite again (4.18) in the slow-fast formulation (4.23), whose solution is $\mu_t = (p_t, m_t)$. We work in $\mathbf{H}_\theta^{-r} = H_\theta^{-r} \times \mathbb{R}$. A simplified statement of the main result of [L12] is (the main interesting point being item 3. below):

Theorem 4.5.1. [Thms 1.2, 1.4 and 1.6 in [L12]]

1. The slow fast system (4.23) is well-posed: for any $\mu_0 = (p_0, m_0) \in \mathbf{H}_\theta^{-r}$ with $\int_{\mathbb{R}^d} p_0 = 1$, there exists a unique mild solution $\mu_t := (p_t, m_t) = T^t(\mu)$ to (4.23) defined for all $t \geq 0$ and $\mu \mapsto T^t(\mu)$ is C^2 .
2. Periodicity: there exists $\delta_0 > 0$ and $r_0 \geq 1$ such that for all $r \geq r_0$, $\delta \in (0, \delta_0)$ and $\theta \in (0, 1]$, equation (4.23) admits a periodic solution

$$(\Gamma_t^\delta)_{t \in [0, T_\delta]} := (q_t^\delta, \gamma_t^\delta)_{t \in [0, T_\delta]} \quad (4.56)$$

in \mathbf{H}_θ^{-r} with period $T_\delta > 0$. Moreover q_t^δ is a probability distribution for all $t \geq 0$, and $t \mapsto \partial_t \Gamma_t^\delta$ and $t \mapsto \partial_t^2 \Gamma_t^\delta$ are in $C([0, T_\delta], \mathbf{H}_\theta^{-r})$.

3. Regularity: if

$$\mathcal{M}^\delta := \{\Gamma_t^\delta : t \in [0, T_\delta]\} \quad (4.57)$$

there exists a neighborhood $\mathcal{W}^\delta \in \mathbf{H}_\theta^{-r}$ of \mathcal{M}^δ and a C^2 mapping $\Theta^\delta : \mathcal{W}^\delta \rightarrow \mathbb{R}/T_\delta\mathbb{Z}$ that satisfies, for all $\mu \in \mathcal{W}^\delta$, denoting $\mu_t = T^t\mu$,

$$\Theta^\delta(\mu_t) = \Theta^\delta(\mu) + t \pmod{T_\delta}, \quad (4.58)$$

and there exists a positive constant $C_{\Theta, \delta}$ such that, for all $\mu \in \mathcal{W}^\delta$ with $\mu_t = T^t\mu$,

$$\|\mu_t - \Gamma_{\Theta^\delta(\mu) + t}^\delta\|_{\mathbf{H}_\theta^{-r}} \leq C_{\Theta, \delta} e^{-\lambda \delta t} \|\mu - \Gamma_{\Theta^\delta(\mu)}^\delta\|_{\mathbf{H}_\theta^{-r}}. \quad (4.59)$$

Remark 4.5.2. The existence of the manifold \mathcal{M}^δ is accompanied with the existence of smooth projections on the tangent space to \mathcal{M}^δ with appropriate contraction properties along the stable direction (see [L12, Th. 1.4] for more detailed statements on this point). The stability result (4.59) comes along also with similar controls on the Fréchet derivatives of the isochron map Θ (see [L12, Th. 1.6] for more details).

4.5.4 Comments and main strategy of proof

The slow-fast viewpoint described in the previous paragraphs suggests that for δ small, the manifold (recall the definition of $q_{0, \sigma^2 K^{-1}}$ in (4.25))

$$\tilde{\mathcal{M}}^\delta := \{(q_{0, \sigma^2 K^{-1}}, \gamma_t) : t \in [0, T]\} \quad (4.60)$$

is an approximately invariant manifold which is approximately normally hyperbolic (without unstable direction). The notion of approximately invariant manifold goes back to the work of Bates, Lu and Zeng [BLZ08] and is highly technical. Let us only retain here the main type of estimate here: there exists a positive constant κ_1 such that for all $u \in \mathbb{R}/\frac{T}{\delta}\mathbb{Z}$, $\|T^{\frac{\tau}{\delta}}(q_{0,\sigma^2 K^{-1}}, \gamma_u) - (q_{0,\sigma^2 K^{-1}}, \gamma_{u+\frac{\tau}{\delta}})\|_{\mathbf{H}_{\theta}^{-r}} \leq \kappa_1 \delta$: $\tilde{\mathcal{M}}^\delta$ is approximately invariant for the discrete semigroup $(T^{n\frac{\tau}{\delta}})_{n \geq 0}$. The other conditions (mostly technical, see [BLZ08] and [L12, § 1.5] for details) essentially require regularity assumptions on $\tilde{\mathcal{M}}^\delta$ along the stable and tangent direction, ensuring that the estimates one has on the discrete semigroup propagates on continuous time. The deep result of [BLZ08] then implies the existence of an actual invariant manifold \mathcal{M}^δ for (4.23) that is located at a distance of order δ from $\tilde{\mathcal{M}}^\delta$ given in (4.60). In [BLZ08] it is in addition proven that the stable manifold of the actual SNHM (in our case \mathcal{M}^δ is attractive, the stable manifold is in fact a neighborhood \mathcal{W}^δ of \mathcal{M}^δ) is foliated by invariant foliations: $\mathcal{W}^\delta = \cup_{m \in \mathcal{M}^\delta} \mathcal{W}_m^\delta$, where $\nu \in \mathcal{W}_m^\delta$ if and only if $T^t(\nu) - T^t(m)$ converges to 0 exponentially fast. This implies the existence of an isochron map $\Theta^\delta : \mathcal{W}^\delta \rightarrow \mathbb{R}/T_\delta\mathbb{Z}$ that satisfies $\Theta^\delta(\nu) = t$ if $\nu \in \mathcal{W}_{\Gamma_t^\delta}^\delta$. The second main result of [L12] is then to prove the regularity of the isochron map Θ , following similar arguments in the finite-dimensional case, see e.g. [Guc75], that we briefly comment here: the strategy is simply based on the following identity: the isochron map satisfies $\Gamma_{\Theta(\mu)}^\delta = S(\mu) := \lim_{n \rightarrow \infty} T^{nT_\delta} \mu$, where we recall that Γ^δ is the limit cycle obtained from Theorem 4.5.1, with period T_δ . Implicit theorem considerations show that the regularity of Θ is the same as the regularity of S , so that it remains to prove that S is C^2 . This is done by writing the proper equations for $DT^t(\mu)$ and $D^2T^t(\mu)$ and proving that the sequences $(DT^{nT_\delta})_{n \geq 0}$ and $(D^2T^{nT_\delta})_n$ are Cauchy sequences. We refer to [L12, § 5] for details.

4.6 Perspectives

4.6.1 Electrical vs chemical synapses

The FHN system (4.18) described in § 4.3 corresponds to the so-called *electrical synapse transmission* between neurons [BFFT12]. Electrical synapses are often found in neural systems that require very fast connections between neurons (e.g. defensive reflexes or in neurons involved in the visual system). Postsynaptic potential in electrical synapses is not caused by the opening of ion channels by chemical transmitters, but rather by direct electrical coupling between pre and post-synaptic neurons [Luo20]. These connections represent however a minority w.r.t. the more common *chemical synapses*. A simplified model for chemical synapses would then write (see [BFFT12, § 2.5]) as $(V_i, w_i, y_i)_{i=1, \dots, n}$ solving

$$\begin{cases} dV_{i,t} &= \left(\delta \left(V_{i,t} - \frac{V_{i,t}^3}{3} - w_{i,t} \right) - \frac{K}{n} \sum_{j=1}^n (V_{i,t} - V_{j,t}) y_{j,t} \right) dt + \sqrt{2\sigma} dB_{i,t}, \\ dw_{i,t} &= \frac{\delta}{c} (V_{i,t} + a - bw_{i,t}) dt, \\ dy_{i,t} &= (S(V_{i,t})(1 - y_{i,t}) - y_{i,t}) dt + \sqrt{2\sigma} (V_{i,t}, y_{t,i}) dW_{i,t} \end{cases}, \quad i = 1, \dots, n, \quad t \geq 0, \quad (4.61)$$

for suitably chosen functions S and σ (see [BFFT12, BFT15] for further details). The only (major) change from (4.20) is in the presence of the variable $y_i \in [0, 1]$ which accounts for the fraction of open channels in the presynaptic neuron j . A natural question in this case would be to ask whether similar emergence or stability of periodic behaviors hold in the chemical case, as suggested by numerical simulations (see e.g. [BFFT12, § 4]). The system we have studied so far corresponding to the situation where each y_t is equally 1. It is certainly very likely to obtain similar results for modifications of (4.61) where each $y_{i,t}$ is a.s away from 0. Extending the present results to the chemical case (4.61) would require to have precise estimates on the possibility that the gating variables $y_{i,t}$ get close to 0. A technical difficulty in handling the multiplicative nature of the noise in (4.61) is expected. It is certainly a worthy perspective of research.

4.6.2 And beyond FHN dynamics?

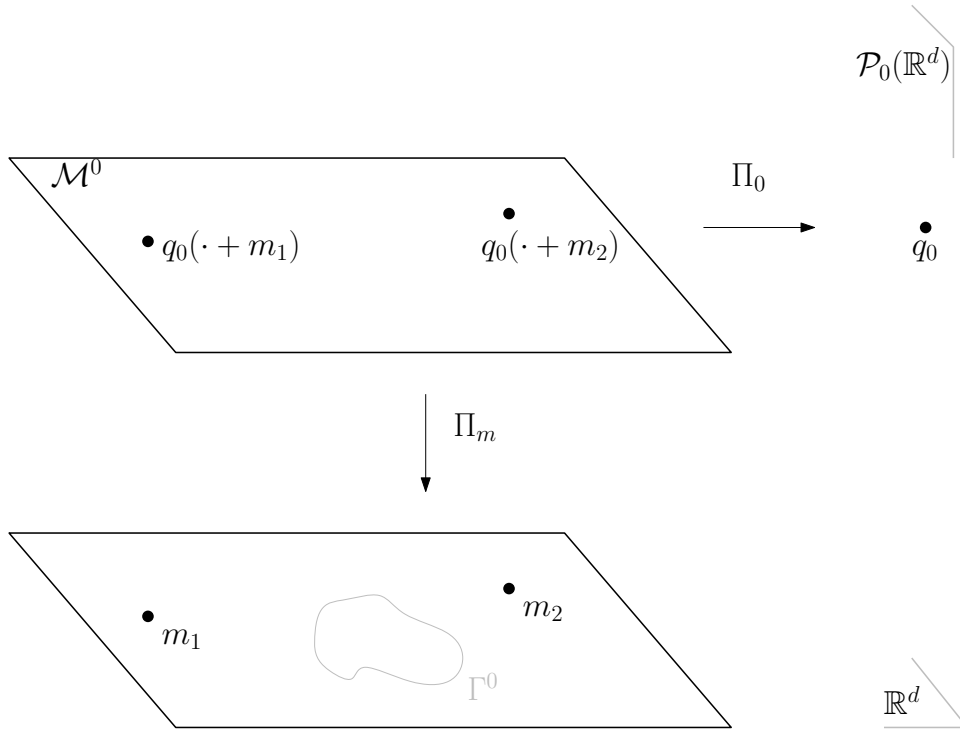
The same techniques developed here are certainly applicable of generalisable to more general dynamics (e.g. the Morris-Lecar model, also exhibiting excitable properties [New14]). The recent work [BFO19] have analysed synchronisation properties of mean-field Hodgkin-Huxley oscillators (with also numerical evidence that oscillations are present in this system, see [BFO19, § 3]). The question of applying the present techniques to these models is a natural perspective.

4.6.3 Quasi-cycles and the role of inhibition in oscillations

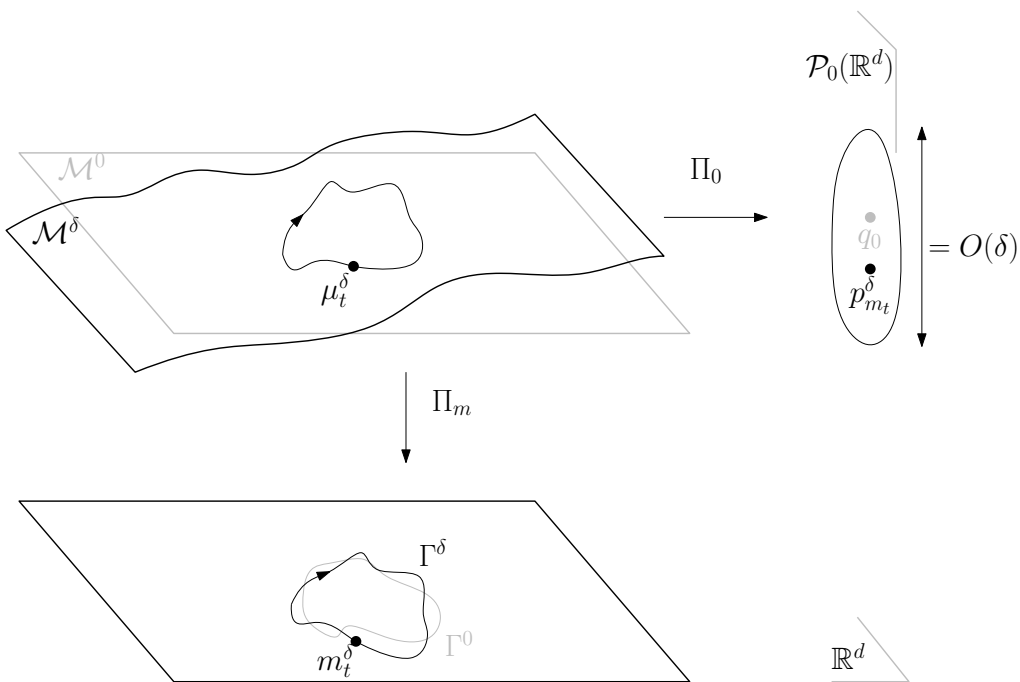
As already said, oscillations are ubiquitous in neuronal systems. One particular key idea is that neural oscillations influence information transfer between different areas of the brain (see [GW22] and references therein). In particular, the role of inhibition in neuronal activity in the emergence of oscillations is definitely not fully elucidated so far. A simple model in this direction may be given by EI oscillators [GW22, BG10]: the model consists of a pair of Excitatory (E) and Inhibitory (I) oscillators with linear cross-interaction under noise. Under specific choice of parameters, the deterministic system possesses a unique fixed-point with conjugated complex eigenvalues: we have damped-oscillations. It has been shown in [BG10] that these oscillations are sustained under the presence of noise (so called *quasi-cycles*). A precise rigorous estimate on the dynamics of the radial part of the process is given in [BG10] in terms of an Ornstein-Uhlenbeck process. It is very natural to inquire about the nature of these quasi-cycles in a mean-field setting. What is the influence of the interaction on the emergence or stability of quasi-cycles in this case? It is natural to expect that the formalism developed in this chapter might apply to this setting.

4.6.4 Systems with delays

On a more general perspective, it is of crucial importance w.r.t. neuroscience applications to consider versions of (4.20) where delay is present in both the dynamics and the interaction. Delay is in particular present within the transport of information through axons. We refer to [Tou14, Ami97, CL09] for the biological motivations of the importance of delays for the modelling of neuronal populations. From a mathematical point of view, it is well-known that delay may be responsible for the emergence of oscillations in deterministic ODEs. In a mean-field context, we are not aware of any rigorous results on the emergence of oscillations in such a context. The only rigorous result known to the author concerns only propagation of chaos on a bounded interval for mean-field dynamics with delay in e.g. [Tou14], with numerical evidence of oscillations induced by delay, see [Tou14, Fig. 2 and 3]. Since the very principle of the methods developed in this chapter is to reduce an infinite dimensional dynamics of a nonlinear Fokker-Planck PDE in terms of a finite dimensional ODEs, the question of deriving the same procedure in a context with delay seems a very promising perspective of research.



(a) The case $\delta = 0$: the invariant manifold \mathcal{M}^0 for (4.18) consists of translations of the same Gaussian measure $q_0 = q_{0, \sigma^2 K^{-1}}$ given by (4.25). Hence the projection of \mathcal{M}^0 onto $\mathcal{P}_0(\mathbb{R}^d)$ is trivial, reduced to $\{q_0\}$. We have reproduced in gray the ineffective limit cycle Γ^0 for (4.36), see for the limit cycle Γ^δ when $\delta > 0$ (recall Remark 4.3.5).



(b) The case $\delta > 0$: \mathcal{M}^δ is a perturbation of order δ of \mathcal{M}^0 and the dynamics onto \mathcal{M}^δ is parameterised by the mean-value, driven by a limit cycle Γ^δ , perturbation of Γ^0 , recall Remark 4.3.5.

Figure 4.3 – The perturbation argument leading to Theorem 4.3.3: we move from a trivial Euclidean invariant manifold \mathcal{M}^0 for (4.18) when $\delta = 0$ (Figure 4.3a) to some stable \mathcal{M}^δ with a nontrivial dynamics parameterised by its mean-value (4.35) (Figure 4.3b).

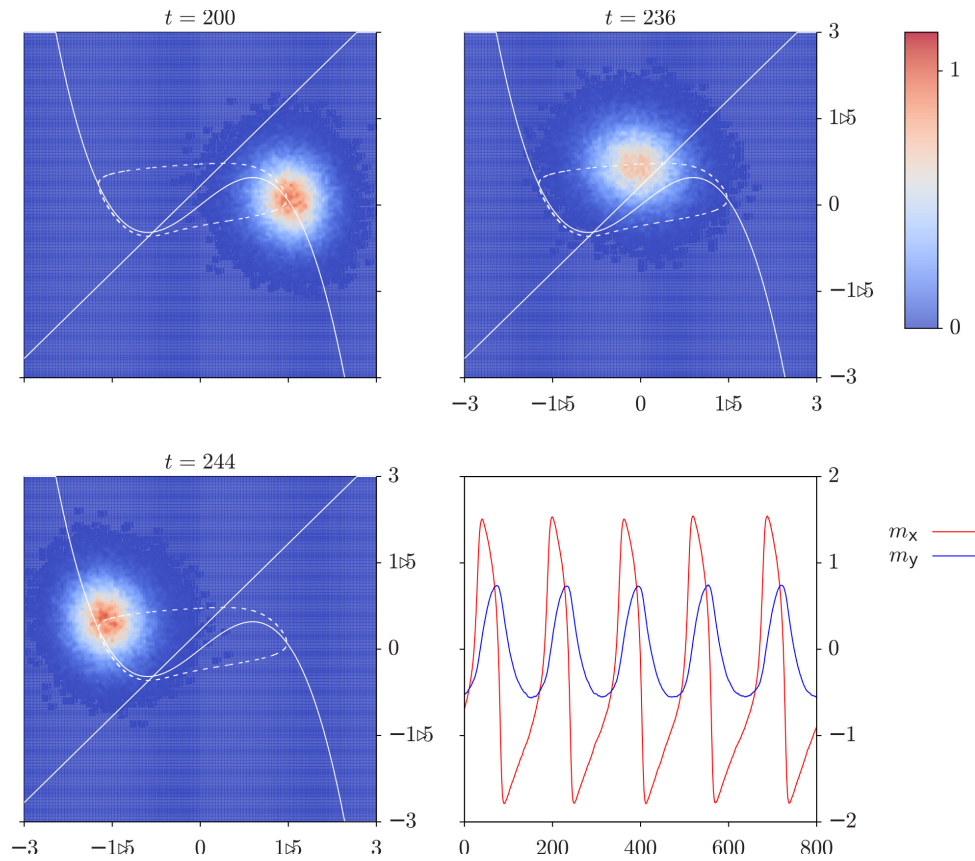


Figure 4.4 – From left to right: evolution of the empirical measure of the system (4.20) for the FitzHugh-Nagumo model, in the excitable regime given by $a = \frac{1}{3}$, $b = 1$ and $c = 10$. The theoretical limit-cycle is represented in dotted-lines. Lower-right: trajectories of the corresponding mean-values on $[0, 800]$. Simulations are made for $n = 50000$ particles and parameters $k_1 = k_2 = 1$, $\sigma_1^2 = \sigma_2^2 = 0.2$ and $\delta = 0.2$.

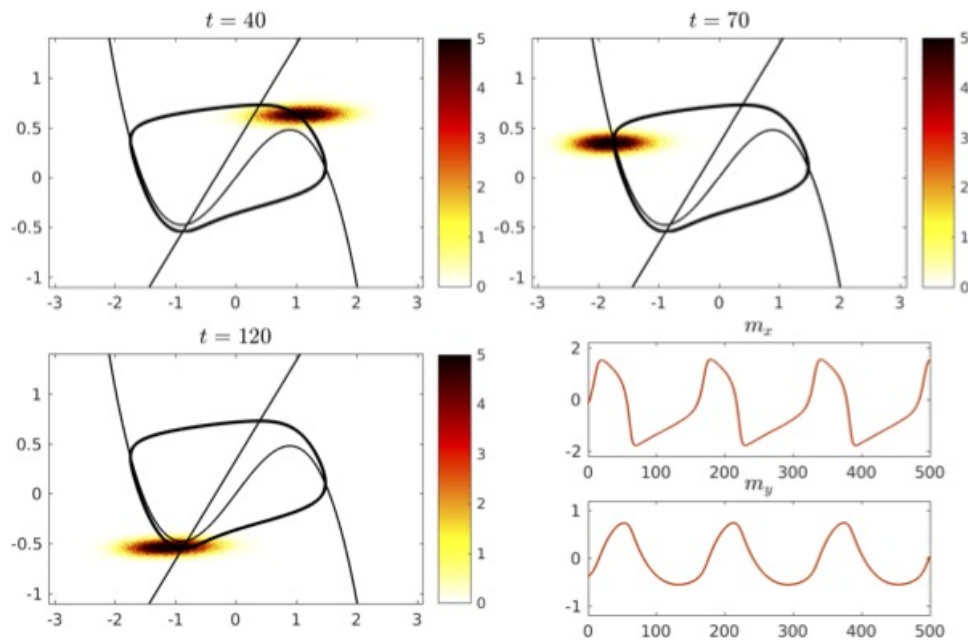


Figure 4.5 – Time evolution of the empirical density of the particle system (4.20) in the kinetic case, for FHN oscillators with $\delta = 0.2$, $a = \frac{1}{3}$, $b = 1$, $c = 10$, $k_1 = 1$, $\sigma_1^2 = 0.2$ and $n = 10^5$. The corresponding evolution of the empirical mean-value is represented.

Chapter 5

Long-time analysis of inhomogeneous diffusions

We are concerned in the present chapter with the following question: how to transpose the existence of periodic solutions for the Fokker-Planck equations (1.4) (Chapter 4) into pseudo-periodic behaviors for the empirical measure $\mu_{n,t}$ (1.2) of the microscopic system (1.1) on unbounded times scales? On an informal level, the general statement is:

Informal statement 5.0.1. *Suppose that the Fokker-Planck equation (1.4) admits a stable periodic orbit $\mathcal{M} := (q_t)_{t \in [0, T_0]}$. Then, there exists a time scale α_n with $\alpha_n \xrightarrow[n \rightarrow \infty]{} +\infty$ and a random phase process $\Theta_{n,t}$ such that, provided the initial condition $\mu_{n,0}$ is close enough to \mathcal{M} , then the rescaled process $(\mu_{n, \alpha_n t})$ converges in probability to $(q_{\Theta_{n,t}})$ with $(\Theta_{n,t})$ having a nontrivial limit as $n \rightarrow \infty$.*

Typically α_n will be \sqrt{n} or n . This informal statement was originally proven in the case of the Kuramoto model without disorder (recall Example 1.2.2, for $\nu \equiv \delta_0$), in the seminal work of Bertini, Giacomin, Poquet [BGP14]: the main result of [BGP14] is that Statement 5.0.1 is true for the choice of the scaling $\alpha_n = n$ (\mathcal{M} being the stable circle of stationary solutions M_0 defined in (1.21)). Here the rescaled empirical measure $\mu_{n, \alpha_n t}$ wanders along \mathcal{M} with a phase that is essentially diffusive: $\Theta_{n,t}$ converges in law to some Brownian motion on the circle. We refer to [BGP14] for details. Note that similar issues have been considered in the context of SPDE models for phase separation, in particular for Ginzburg-Landau dynamics [BDMP95, BB98, Fun95] and for Cahn-Hilliard equation [ABK12].

5.1 Random traveling waves in the disordered Kuramoto model

The first aim of the chapter is to review [L7] which addresses Statement 5.0.1 in the case of the Kuramoto model with nontrivial disorder (Example 1.2.2). Recall § 1.2.2 and § 4.2.2: the Fokker-Planck (1.15) admits a circle of stationary solution M that is linearly stable. Recall also the analysis made in § 4.2 (in particular Theorem 4.2.1): in case the law of the disorder ν is symmetric, Theorem 4.2.1 is essentially empty, as the macroscopic speed of rotation is equally zero. The relevant phenomenon happens on a microscopic level. Suppose here that $(\omega_i)_{i \geq 1}$ is an i.i.d. sequence with symmetric law $\nu = \frac{1}{2}(\delta_{-\omega_0} + \delta_{+\omega_0})$ (note that the result is quenched and does not really depend on the underlying mechanism that produced the sequence $(\omega_i)_{i \geq 1}$, it only depends on the asymmetry of this sequence, see [L7, Def. 2.1] for straightforward generalisations). Even though ν is symmetric, fluctuations of any *finite sample* of the disorder $(\omega_1, \dots, \omega_n)$ induce a microscopic asymmetry (of size $\frac{1}{\sqrt{n}}$) in favour of one population ($+\omega_0$, say). This suggests that on a time scale $\alpha_n = \sqrt{n}$, the empirical measure $\mu_{n, \sqrt{n}t}$ will follow this asymmetry and perform a traveling wave along the manifold of stationary solutions M with speed deterministic in the sample $(\omega_i)_{i \geq 1}$. The asymmetry of the sample $(\omega_1, \dots, \omega_n)$ is characterised for all n by the quantity (ξ_n^-, ξ_n^+) defined by: $\xi_n^\pm := n^{1/2} \left(\frac{n^\pm}{n} - \frac{1}{2} \right)$, where n^\pm is the number of frequencies in $(\omega_1, \dots, \omega_n)$ taking the value $\pm\omega_0$. Note that $\xi_n^- + \xi_n^+ = 0$ for all $n \geq 1$. The main result of [L7] is the following.

Theorem 5.1.1 (Th. 2.4 in [L7]). *For all $K > 1$, there exists $\delta(K)$ such that, for all $\delta \leq \delta(K)$, there exists $v > 0$ (depending on K , δ and ω_0) and a real number $\varepsilon_0 > 0$ such that the following holds: for any vector of probability measures p_0 satisfying $\text{dist}_{H^{-1}}(p_0, M) \leq \varepsilon_0$ such that for all $\varepsilon > 0$,*

$$\mathbb{P}(\|\mu_{n,0} - p_0\|_{-1} \geq \varepsilon) \rightarrow 0, \text{ as } n \rightarrow \infty, \quad (5.1)$$

there exists $\theta_0 \in \mathbb{T}$ (depending on p_0) and a constant c such that for each finite time $t_f > 0$ and all $\varepsilon > 0$, denoting $t_0^n = cn^{-1/2} \log n$, we have

$$\mathbb{P}\left(\sup_{t \in [t_0^n, t_f]} \|\mu_{n, n^{1/2}t} - q_{\theta_0 + v(\xi_n^+ - \xi_n^-)t}\|_{-1} \geq \varepsilon\right) \rightarrow 0, \text{ as } n \rightarrow \infty. \quad (5.2)$$

Moreover, we have the following expansion in δ : $v = \delta\omega_0 + O(\delta^2)$.

Theorem 5.1.1 is simply saying that, on a time scale of order $n^{1/2}$, the empirical measure μ_n is asymptotically close to a synchronized profile $q \in M$, traveling at speed v along M . This drift depends on the asymmetry ξ_n of the quenched disorder $(\omega_i)_{i \geq 1}$. In (5.2), t_0^n represents the time necessary for the system to get sufficiently close to the manifold M .

Main lines of proof

In Theorem 5.1.1, we work with a disordered extension of the H^{-1} topology introduced in (1.25). The strategy of proof (which borrows from [BGP14]) is the following: the starting point is to consider the dynamics of the difference between the empirical measure $\mu_{n,t}$ and any generic element of $q \in M$, $\nu_{n,t} := \mu_{n,t} - q$. We write here a mild formulation of $\nu_{n,t}$ (where all the terms make sense as elements of $\mathcal{C}([0, \infty), H^{-1})$), which is a direct consequence of Ito's formula and (1.15):

$$\nu_{n,t} = e^{tL}\nu_{n,0} + \int_0^t e^{(t-s)L} (D_n - \partial_\theta R_n(\nu_{n,s})) ds + Z_{n,t}, \quad n \geq 1, \quad t \geq 0. \quad (5.3)$$

Here, e^{tL} is the semigroup of the linearized operator around $q \in M$ defined by (4.17), $D_n := -\partial_\theta \left(q \left(\left(\frac{n^+}{n} - \frac{1}{2} \right) (J * q^+) \right) + \left(\frac{n^-}{n} - \frac{1}{2} \right) (J * q^-) \right)$ reflects the asymmetry of the disorder, R_n is the quadratic error induced by linearisation around q and $Z_n = \frac{1}{2} (Z_n^- + Z_n^+)$ is the noise term where

$$Z_{n,t}^\pm(h) = \frac{1}{n^\pm} \sum_{j=1}^{n^\pm} \int_0^t \partial_\theta \left[\left(e^{(t-s)L^*} h \right) \right] (\theta_j(s)) dB_j(s), \quad (5.4)$$

We refer to [L7] for the precise definitions of the previous objects. Equation (5.3) reflects the competition of the two main antagonist phenomena. The first effect is deterministic: from the spectral decomposition of L , we know that the semigroup e^{tL} is contracting along the orthogonal direction at $q \in M$, whereas the tangent direction remains neutral (by invariance by rotation of the problem). This simply means that the deterministic part of (5.3) (in absence of the other terms D_n , R_n and Z_n) leads to the convergence of μ_n to some $q \in M$. On the other hand, the noise term Z_n and the drift term D_n tend to deviate μ_n from q . But thanks to the orthogonal stability of M , this deviation can essentially only occur w.r.t. the tangential direction. Therefore the strategy is to proceed with a time discretisation: we divide $[0, \sqrt{n}t]$ into \sqrt{n} intervals of length T and proceed by recursion. Starting from $\nu_{n,0} = \mu_{n,0} - q_{\psi_0}$, where ψ_0 is a proper projection of $\mu_{n,0}$ on M , since one can easily prove that the noise term Z_n is of order at most $N^{-1/2+\zeta}$ (for small $\zeta > 0$), if one supposes that initially $\nu_{n,0}$ is of order $N^{-1/2+2\zeta}$, it will remain of order $2N^{-1/2+2\zeta}$ on $[0, T]$ (since R_n is quadratic, if ν_n is small, R_n will be even smaller). The end of the recursion is then to project back the process $\mu_{n,T}$ onto M (defining a new projection q_{ψ_T}) and to proceed further the recursion with $\nu_{n,t} = \mu_{n,t} - q_{\psi_T}$ for $t \in [T, 2T]$. The key argument is here to make sure that $\nu_{n,T} = \mu_{n,T} - q_{\psi_T}$ is not only of order $2N^{-1/2+2\zeta}$ but goes back exactly to order $N^{-1/2+2\zeta}$. We use here the contraction of the dynamics around M by using that $\|e^{TL} P_1 u\| \leq e^{-\gamma T} \|u\|$ and choosing T sufficiently large. The end of the argument is to look at the dynamics of the projected process q_{ψ_n} and to check that, suitably rescaled, this gives the desired drift term. Since the strategy is similar to the one followed in the next paragraph, we leave the details to [L7].

5.2 Long-time diffusive behavior for general mean-field diffusions

We have recently transposed the above techniques and results to the general case of diffusions in \mathbb{R}^d with linear interaction as described in § 4.3. The work [L15] is the companion paper of [L12] mentioned in § 4.5: we work in the same set of hypotheses, (especially the same regime of small $\delta > 0$) ensuring in particular that the Fokker-Planck PDE (4.18) (or equivalently its slow-fast formulation (4.23)) admits a stable periodic orbit (recall Theorem 4.5.1). For the analysis of the empirical measure of (4.20), it is convenient to adopt a slow-fast reformulation that is similar to (4.23): the knowledge of the empirical measure $\mu_{n,t}$ of (4.20) is exactly equivalent to the knowledge of $\mu_{n,t} := (p_{n,t}, m_{n,t})$ $t \geq 0$, where $p_{n,t} := \frac{1}{n} \sum_{k=1}^n \delta_{Y_{i,t}}$ is the empirical measure of the rescaled process

$$Y_{i,t} := X_{i,t} - m_{n,t}, \quad i = 1, \dots, n \quad (5.5)$$

and

$$m_{n,t} = \frac{1}{n} \sum_{j=1}^n X_{j,t}, \quad t \geq 0. \quad (5.6)$$

Standard propagation of chaos result gives that $(p_{n,t}, m_{n,t})$ is close as $n \rightarrow \infty$ to (p_t, m_t) solution to (4.23). Placing ourselves under the hypotheses of Theorem 4.5.1, we know that the deterministic system (4.23) admits a stable limit cycle $(q_t^\delta, \gamma_t^\delta)_{t \in [0, T_\delta]}$. In particular, we know from Theorem 4.5.1 that q^δ is close to be Gaussian. Therefore, its microscopic equivalent p_n should be close to a Gaussian too: defining $g_N(x) = \frac{1}{(2\pi)^{\frac{d}{2}} (1 - \frac{1}{N})^d \sqrt{\det(K^{-1}\sigma^2)}} w_{\frac{N}{N-1}}(x)$ (recall the definition of w_θ in (4.28)), we can state the main result of [L15]:

Theorem 5.2.1. *Suppose that the hypotheses of Theorem 4.5.1 are satisfied. There exist $\gamma \in (0, 1)$, some sufficiently small $\theta > 0$ and $r \geq 1$ sufficiently large and $\delta_1 > 0$ such that for all $\delta \in (0, \delta_1)$ the following holds: if we suppose that there exist $\kappa_0 > 0$, $u_0 \in [0, T_\delta)$ such that for all $\varepsilon > 0$*

$$\sup_{n \geq 1} \mathbf{E} [\langle p_{n,0}, w_{-\gamma} \rangle] < \infty, \quad (5.7)$$

$$\mathbf{P} \left(\|p_{n,0} - g_n\|_{H_\theta^{-r+2}} \leq \kappa_0 \right) \xrightarrow{n \rightarrow \infty} 1, \quad (5.8)$$

$$\mathbf{P} \left(\|\mu_{n,0} - \Gamma_{u_0}\|_{\mathbf{H}_\theta^{-r}} \leq \varepsilon \right) \xrightarrow{n \rightarrow \infty} 1, \quad (5.9)$$

then for all $\varepsilon > 0$, we have

$$\mathbf{P} \left(\sup_{t \in [0, t_f]} \|\mu_{n,nt} - \Gamma_{u_0+nt+v_{n,t}}\|_{\mathbf{H}_\theta^{-r}} \leq \varepsilon \right) \xrightarrow{n \rightarrow \infty} 1, \quad (5.10)$$

where the random process $v_{n,t}$ satisfies $v_{n,0} = 0$ and converges weakly to $v_t = bt + a^2 w_t$, where w is a standard Brownian motion and b and a are constant coefficients.

Here, we work once again in Sobolev spaces with negative indexes $\mathbf{H}_\theta^{-r} = H_\theta^{-r} \times \mathbb{R}$ defined in § 4.5.2. Hypotheses (5.7) and (5.8) ensure sufficient control on the moments of the initial condition and (5.9) require the initial condition to be sufficiently close to the periodic orbit. Hence, (5.10) is nothing else than a reformulation of Statement 5.0.1: with high probability, on a time scale of order $\alpha_n = n$, the empirical measure $\mu_{n,nt}$ performs a Brownian motion on the periodic orbit Γ . (Not so) explicit formulae for b and a^2 are given in [L15] in terms of the first two derivatives of the isochron Θ taken on $(\Gamma_u)_{u \in [0, T_\delta]}$. For simplicity, the result of Theorem 5.2.1 is proved for initial conditions close to a point $\Gamma_{u_0}^\delta$ of the stable periodic solution, but as it was done in [BGP14] and [L7] one could start close to a point lying in the basin of attraction of the periodic solution, and show that the empirical measures first reaches a neighborhood of size ε of the periodic solution (after a time interval of length independent from n).

5.2.1 Main lines of proof

The proof relies on a similar procedure than for [L7], but with the numerous extra difficulties coming from the fact that we work in the non compact space \mathbb{R}^d . The first step is to prove that the proximity (5.8) of the re-centered process $p_{n,t}$ with the Gaussian g_n propagates from $t = 0$ to any time of order n . This is carried out by writing the semi-martingale decomposition of $p_{n,t}^{(k)} - g_n := p_{n,kT+t} - g_n$ for $k = 0, \dots, n$

$$p_{n,t}^{(k)} - g_n = e^{\mathcal{U}_n} \left(p_{n,0}^{(k)} - g_n \right) - \delta \int_0^t e^{(t-s)\mathcal{U}_n} \nabla \cdot \left\{ \left(F_{m_n, kT+s} - \langle p_{n,s}^{(k)}, F_{m_n, kT+s} \rangle \right) p_{n,s}^{(k)} \right\} ds + V_{n,t}^{(k)}, \quad (5.11)$$

Here, $\mathcal{U}_n = \left(\frac{n-1}{n} \right) (\nabla \cdot (\sigma^2 \nabla p) + \frac{n}{n-1} \nabla \cdot (pKx)) = \left(\frac{n-1}{n} \right) \mathcal{L}_{\frac{n}{n-1}}$ where $\mathcal{L}_{\frac{n}{n-1}}$ is the Ornstein-Uhlenbeck operator given in (4.40) and $V_n^{(k)}$ is the noise term

$$V_{n,t}^{(k)}(f) = \sum_{i=1}^n \int_0^t \nabla e^{(t-s)\mathcal{U}_n^*} f(Y_{i, kT+s}) \cdot \sigma \left(dB_{i, kT+s} - \frac{1}{n} \sum_{j=1}^n dB_{j, kT+s} \right). \quad (5.12)$$

Here we exploit the stability properties of the Ornstein-Uhlenbeck operator (note here the importance to obtain spectral gap estimates such as (4.55) that are uniform in the parameter θ' as $\theta' = \frac{n}{n-1}$ depends on n). The main difficulty lies in the control of the noise term $V_n^{(k)}$: one is typically interested in controls of the following type (for a sufficiently large $m \geq 1$, see [L15, Prop. 3.6] details)

$$\mathbf{E} \left[\left\| V_{n,t}^{(k)} - V_{n,s}^{(k)} \right\|_{H_{\theta}^{-r+2}}^{2m} \right] \leq C \frac{(t-s)^m}{n^m} \quad (5.13)$$

Estimate (5.13) is a consequence of the Burkholder-Davis-Gundy inequality in Hilbert spaces [MR16] and essentially boils down to estimating norms of Dirac masses $\|\delta_{Y_{i,s}}\|_{H_{\theta}^{-r+2}}$. But this last quantity may only be bounded by exponential terms $Cw_{\frac{\theta}{4-\eta}}(Y_{i,s})$ for any $\eta > 0$. Hence, proving (5.13) requires to have uniform a priori bounds on $\mathbf{E} \left[w_{\frac{m\theta}{4-\eta}}(Y_{i,s}) \right]$. Since we consider cases where $Y_{i,s}$ has a distribution close to be Gaussian, the last expectation is bounded provided θ is sufficiently small. Therefore, a special attention has to be put on the spectral properties of $\mathcal{L}_{\theta'}$ in H_{θ}^{-r} not only when $\theta = \theta' = 1$ but uniformly for $\theta \rightarrow 0$. This is one of the main difficulty in both [L12, L15], we refer e.g. to [L12, App. A] for more details on this point. Applying now Garsia-Rademich-Rumsey Lemma (see [SV06] and [L15, Lem. 3.7]), we deduce from (5.13) that the event

$$\mathcal{A}_n := \left\{ \max_{k=0, \dots, k_f} \sup_{t \in [0, T]} \left\| V_{n,t}^{(k)} \right\|_{H_{\theta}^{-r}} \leq n^{-\frac{1}{2} + \zeta} \right\} \quad (5.14)$$

is such that, for δ, ζ small and appropriate $r \geq 1$, $\mathbb{P}(\mathcal{A}_n) \xrightarrow[n \rightarrow \infty]{} 1$ (see [L15, Lem. 3.7]). Hence, from now on, we are allowed to place ourselves on this event \mathcal{A}_n and essentially perform deterministic estimates. The second point of the proof is to show that the proximity (5.8) propagates from $t = 0$ to times of order n (see [L15, Prop. 4.1]): for small $\delta > 0$ and appropriate $r \geq 1$ and $\kappa_0, \theta, \gamma > 0$, if $\sup_{n \geq 1} \mathbf{E}(\langle p_{n,0}, w_{-\gamma} \rangle) \leq \infty$ and $\mathbf{P} \left(\|p_{n,0} - g_n\|_{H_{\theta}^{-r+2}} \leq \kappa_0 \right) \xrightarrow[n \rightarrow \infty]{} 1$, then there exists $\kappa_1 > 0$ such that

$$\mathbf{P} \left(\sup_{0 \leq t \leq nt_f} \|p_{n,t} - g_n\|_{H_{\theta}^{-r+2}} \leq \kappa_1 \right) \xrightarrow[n \rightarrow \infty]{} 1 \quad (5.15)$$

Proving (5.15) relies again on the semimartingale decomposition (5.11) and a recursive argument on k , using the spectral gap estimates we have on the Ornstein-Uhlenbeck operator (4.55), in a very similar way to the proof of Theorem 5.1.1. The next point is then to prove the main proximity result (5.10): the difficulty is here to properly write a mild formulation of a semi-martingale decomposition similar to (5.11) of μ_n in the vicinity of the stable orbit Γ and using a local stability argument of Γ expressed in terms of Floquet operators (see [L15, § 4.2] for more details).

The third and last point of the proof is to look at the dynamics of the isochron phase $\Theta(\mu_{n,t})$ along the manifold. The use of the isochron to derive the proper asymptotics of the empirical measure along the periodic orbit goes back to the work of Giacomini, Poquet and Shapira [GPS18] who used similar techniques for finite-dimensional SDEs with vanishing noise. The difficulty is here again to extend these techniques to an infinite-dimensional setting. In a word, the point is to apply Ito's formula to the phase process $\Theta(\mu_{n,t})$, on a time scale of order n . Ito's formula gives that $\Theta(\mu_{n,nt})$ is basically of order nt (that is, the contribution of the deterministic dynamics along Γ), plus supplementary terms coming from the Ito correction (involving the first and second derivatives $D\Theta$ and $D^2\Theta$) that become only apparent on the time scale of order n . We refer to [L15, § 5] for more details.

5.3 Perspectives

5.3.1 Interaction on a no longer complete graph

We discussed at length in Chapter 3 on the large population dynamics of particles interacting on a graph that is no longer complete. This discussion only concerned dynamics on a bounded time interval $[0, T]$ and a natural question is of course about the possibility of transposing the techniques of the present chapter to look at the dynamics on a longer time scale. This is actually a difficult question, that has raised only recent partial answers so far. The main step in this direction was made by Coppini [Cop22] for both subcritical and supercritical Kuramoto model on Erdős-Rényi graphs. In the subcritical case, it is shown in [Cop22] that the empirical measure $\mu_{n,\alpha_n t}$ remains close to the uniform measure $\frac{1}{2\pi}$ up to sub-exponential times scales $\alpha_n = \exp(o(n))$, which is a very satisfactory result. In the supercritical case, however, [Cop22] only proves a partial result: $\mu_{n,\alpha_n t}$ remains in a vicinity of the synchronized manifold for $\alpha_n = \exp(o(n))$. The precise dynamics of the empirical measure along this manifold on a time scale of order n (as in § 5.1 or [BGP14]) remains largely open for now. Extending the previous result to Erdős-Rényi graphs is a real challenge as it requires to identify the terms of order $\frac{1}{n}$ in the expansion of the empirical measure around its mean-field limit (whereas the CLT result of Chapter 3 only required to go up to $\frac{1}{\sqrt{n}}$). A question would then be about the possibility of applying Grothendieck inequalities in this case, as we have done in § 3.2.

5.3.2 Other types of dynamics

From what we have seen in the present chapter, the main ingredients leading to Statement 5.0.1 (for both Kuramoto model and FHN-type dynamics) is informally the combination of 1) an invariant manifold for the mean-field limit with sufficient stability properties and 2) a random microscopic dynamics seen as a perturbation of the deterministic dynamics along this invariant manifold, with sufficient control on the noise part. An important remark is that absolutely nothing indicates that the dynamics should be of diffusive type. There are in particular many examples of mean-field systems with jump dynamics that also present oscillatory behaviors. We have in particular in mind the case of mean-field Hawkes processes, that have been intensively studied recently to model neuronal dynamics (we refer to the classical references [DFH16, BM96] and the next chapter of this manuscript for further definitions and references). Many extensions of Hawkes processes present oscillatory behaviors under the additional presence of inhibition [DL17], [L13] or spatial extension [DL17].

The PhD thesis of Zoé Agathe-Nerine

The previous remark is the starting point of the PhD thesis of Zoé Agathe-Nerine I have been supervising in co-direction with Ellen Saada since sept. 2020. The point of the work of Zoé is to look at mean-field Hawkes dynamics interacting on possibly inhomogeneous W -random graphs (hence, this is in connection with Chapters 3, 5 and 6 of this manuscript). Informally speaking,

the point is to look at (Z_1, \dots, Z_n) counting processes with conditional intensities given by

$$\lambda_t^{i,n} = \Phi\left(X_t^{n,i}\right), \quad X_t^{n,i} = \frac{\kappa_i^n}{n} \sum_{j=1}^n \xi_{i,j}^n \int_0^t h(t-s) dZ_{j,u}^i \quad (5.16)$$

where the $\xi_{i,j}^n$ are sampled according to an inhomogeneous graphon W , as in Chapter 3. We refer to Chapter 6 for insights on (5.16). This work has led so far to one publication [AN22b] and one preprint [AN22a]. It is proven in [AN22b] that the system described by (5.16) converges as $n \rightarrow \infty$ to a family of inhomogeneous Poisson process whose deterministic intensity solves a nonlinear spatially-extended convolution equation. In case of exponential memory kernels, this convolution equation presents strong links with neural field equations (NFE) which are of prime importance in the modeling of mesoscopic neural activity in the brain [WC72, Ama77, Bre12, KE13]:

$$\partial_t u_t(x) = -u_t(x) + \int_I W(x,y) f(u_t(u)) \rho(dy) + I_t(x). \quad (5.17)$$

We refer to [Bre12, Ama77] for precise definitions and biological background on (5.17). The work of Zoé stems from the seminal article of Chevallier, Duarte, Löcherbach and Ost [CDLO19] who first gave a microscopic interpretation of (5.17) as the macroscopic limit of interacting Hawkes processes. [AN22b] generalizes the result of [CDLO19] to general random connections. The main result of the second work [AN22a] is to prove long-term stability of such Hawkes processes in case (5.17) possesses a unique equilibrium (using techniques similar to the present chapter). Zoé is currently working on the long-term stability of the so-called *ring model* (that would be the equivalent of (5.17) on the circle $I = \mathbb{S}^1$, see [KE13]). Long-term stability and the asymptotic description of the particle system along this circle (wandering bumps) is expected, using techniques similar to the ones exposed in the present chapter. One point is here to give a new rigorous interpretation (on a long time scale) of fluctuations results around (5.17) derived on a bounded time interval by Chevallier and Ost [CO20].

Looking at the neural field equation on the real line

The results of [CDLO19, AN22b] are only valid when the spatial domain I in (5.17) is endowed with a probability measure ρ . This in particular does not include the historical NFE (5.17) [Ama77, WC72] that corresponds to $\rho = \text{Lebesgue}$ on $I = \mathbb{R}^d$. This restriction is by no means a technical issue, as, by a crucial use of the translation invariance of the Lebesgue measure, it is possible to prove the existence of traveling waves [EM93] in case $(I, \rho(dx)) = (\mathbb{R}^d, dx)$ (whereas this translation invariance is broken for any probability measure ρ). In this setting, a promising perspective of research is 1) to give a microscopic interpretation of (5.17) in terms of appropriate spatially-extended Hawkes processes and 2) to investigate long-term stability of this microscopic system w.r.t. these traveling patterns. Based on recent results on the stability of these traveling waves [LS16, Lan16, KS14, BW12], it is natural to expect significant progress in this direction, based on extensions of the techniques developed in the present chapter.

Chapter 6

Hawkes processes with multiplicative inhibition

The last chapter of this manuscript concerns a recent interest [L13] in mean-field interacting Hawkes processes, modelling the activity of a population of interacting neurons with inhibition. The nature of the dynamics is different from diffusions type models as discussed in the rest of the document, but the philosophy and main objectives remain the same: although the work presented here could be viewed as preliminary, the techniques developed in this manuscript (especially the long-term dynamics around periodic solutions) are also likely to be adaptable to this case.

6.1 Hawkes processes with excitation and inhibition

Following the seminal paper [HO74], the interest in the use of Hawkes processes in the modelling of neuron dynamics has been renewed by the important paper of Delattre, Fournier, Hoffmann [DFH16], who especially looked to homogeneous mean-field interacting Hawkes processes and their limit in large population, expressed in terms of inhomogeneous Poisson processes. One typically looks at interacting counting processes $(Z_{1,t}, \dots, Z_{n,t})$ whose conditional intensities $(\lambda_{1,t}, \dots, \lambda_{n,t})$ satisfy

$$\lambda_t^{i,n} = \lambda_t^n = \Phi \left(\frac{1}{n} \sum_{j=1}^n \int_0^{t^-} h(t-s) dZ_{j,s} \right). \quad (6.1)$$

The Hawkes framework has led to a variety of extensions and applications, including Hawkes processes with age [Che17b, Che17a, RDL19], spatial extension [CDLO19, AN22b, AN22a]. We do not aim to be exhaustive here and refer to the previous references for further details. Here Φ represents synaptic integration and the kernel h models time interdependence between neurons. A generic limit as $n \rightarrow \infty$ of (6.1) (see [DFH16] for details) is then given in terms of inhomogeneous Poisson processes \bar{Z}_t with deterministic intensity $\lambda(t)$ solving

$$\lambda(t) = \Phi \left(\int_0^t h(t-s) \lambda(s) ds \right). \quad (6.2)$$

The convolution equation (6.2) is nothing else than the *Nonlinear Renewal Equation* (NRE) that has many applications that encompass largely the case of Hawkes processes. Its longtime behavior was first studied in the linear case ($\Phi(x) = \mu + x$) by Feller [Fel41, ARM76] and further developed in the fully nonlinear case in [Lon74]. The linear case has a nice interpretation in terms of Galton-Watson processes with immigration [HO74]. Assuming h to be positive and Φ nondecreasing in (6.1) corresponds to an *excitatory regime*: each spike contributes to an increase (of size $\frac{1}{n}$, modulated by the kernel h) of the intensity of each counting process Z_i . The difficulty comes when modelling *inhibition*: a common framework that has been introduced so far in the literature (see e.g. or inhibition [DL17, RL20, RBRTM13, CGMT20]) is to allow h to take negative values (it is therefore possible that a spike of a neuron j may *decrease* the intensity of neuron i , this what we could call *additive inhibition*). A crucial remark is that one necessarily

needs to modulate the intensity by a positive kernel Φ that is necessarily nonlinear (a typical example being $\Phi(x) = x^+$) in order to preserve the positivity of the intensity in (6.1). This causes a major difficulty of the long-time analysis of (6.1)- (6.2): the nonlinearity and the fact that h may take negative values break the monotonicity that is naturally present for (6.2) in the excitatory case.

6.2 Our model: multiplicative inhibition

The main contribution of the model of [L13] is to bypass the difficulty of the previous models with additive inhibition (with negative kernels h) by proposing a *multiplicative* structure for inhibition, thus allowing to conserve positive kernels and preserve monotonicity in the model.

6.2.1 The model

We give a restrictive version of the model considered in [L13] for simplicity of exposition and refer to the paper for details (notably the fact that one can work with generic intrinsic functional for the dynamics of each population). Consider a population of neurons of size $n = n_A + n_B \geq 1$, that is divided into population A (that is considered to be excitatory) with size $n_A := \lfloor \alpha n \rfloor$ with $\alpha \in (0, 1)$ fixed (typically $\alpha = 0.8$, see [BS98]) and a population B with size $n_B = n - n_A$ which is considered to be inhibitory. Define the family of counting processes $(Z_{1,t}, \dots, Z_{n_A,t})$ (population A) and $(Z_{n_A+1,t}, \dots, Z_{n,t})$ (population B) with coupled conditional stochastic intensities given respectively by λ^A and λ^B as follows

$$\begin{cases} \lambda_t^{A,n} : & = \left(\mu_A + \frac{1}{n} \sum_{1 \leq j \leq n_A} \int_0^{t^-} h_1(t-u) dZ_{j,u} \right) \Phi_{B \rightarrow A} \left(\frac{1}{n} \sum_{n_A+1 \leq j \leq n} \int_0^{t^-} h_2(t-u) dZ_{j,u} \right), \\ \lambda_t^{B,n} : & = \mu_B + \frac{1}{n} \sum_{1 \leq j \leq n_A} \int_0^{t^-} h_3(t-u) dZ_{j,u}. \end{cases} \quad (6.3)$$

with $\mu_A, \mu_B \geq 0$ and $h_i \geq 0$ for $i = 1, \dots, 3$ with $h_i(u) \xrightarrow{u \rightarrow \infty} 0$, $i = 1, \dots, 3$. Suppose the h_i , $i = 1, \dots, 3$ to be locally integrable on $[0, +\infty)$. Suppose also that $\Phi_{B \rightarrow A}$ is globally Lipschitz and non-increasing with $0 \leq \Phi_{B \rightarrow A} \leq 1$, $\Phi_{B \rightarrow A}(0) = 1$ and $\Phi_{B \rightarrow A}(x) \rightarrow 0$ as $x \rightarrow \infty$.

The dynamics (6.3) is still of mean-field nature (each neuron interacts with a positive proportion of the whole population) but inhomogeneous, as the dynamics of populations A and B are different: the dynamics of A is essentially a linear Hawkes process (*excitation* within population A), but where *inhibition* acts as a multiplicative factor between 0 and 1 applied to the excitatory inputs, through the kernel $\Phi_{B \rightarrow A}$. The dynamics of B incorporates *retroaction* from population A onto population B , that is supposed to be mostly additive (although considering nonlinear feedback kernels $\Phi_{A \rightarrow B}$ is possible). Here, the crucial point is that the multiplicative structure of inhibition in (6.3) no longer requires the synaptic kernels to take negative values: we only consider nonnegative kernels h_i , $i = 1, \dots, 3$. In this way, the non-negativity of the intensities is automatically preserved, together with the nice properties of the canonical linear Hawkes process: we may still take advantage of the positivity of the synaptic kernels and the monotonicity of $\Phi_{B \rightarrow A}$ to obtain asymptotic results similar to the linear case.

6.2.2 Main results

The results of [L13] are twofold: first the analysis of (6.1) as $n \rightarrow \infty$ and second, the analysis of the limiting system as $t \rightarrow \infty$. Informally, in the limit $n \rightarrow \infty$, (6.3) is correctly described in terms of a couple of inhomogeneous Poisson processes (\bar{Z}^A, \bar{Z}^B) with intensities (λ^A, λ^B) solving

$$\begin{cases} \lambda_t^A & = \left(\mu_A + \alpha \int_0^t h_1(t-u) \lambda_u^A du \right) \Phi_{B \rightarrow A} \left((1-\alpha) \int_0^t h_2(t-u) \lambda_u^B du \right), \\ \lambda_t^B & = \mu_B + \alpha \int_0^t h_3(t-u) \lambda_u^A du. \end{cases} \quad (6.4)$$

6.2.3 Large population dynamics

This section is a simple derivation around standard coupling techniques introduced in [DFH16]: the first result of [L13] is

Proposition 6.2.1 (Prop. 2.7, 2.8 and 2.9 in [L13]). *Under the previous assumptions, both systems (6.3) and (6.4) are well posed. If in addition, the h_i are locally square integrable, there exists a coupling $(Z_{i,t}, \bar{Z}_{i,t})$ between (6.3) and (6.4) such that, for all $T > 0$, there exists some constant $C > 0$ (depending on T and the parameters of the model) such that*

$$\sup_{i=1,\dots,n} \mathbf{E} \left[\sup_{t \in [0, T]} |Z_{i,t} - \bar{Z}_{i,t}| \right] \leq \frac{C}{\sqrt{n}}. \quad (6.5)$$

The coupling mentioned is standard: one simply constructs both processes upon a thinning procedure on the same i.i.d. Poisson measures on $[0, +\infty)^2$ (see e.g. [DFH16]). Note however a small technical difficulty here: the same result in [DFH16] relies on the fact that the synaptic kernel Φ in (6.1) is globally Lipschitz. This is not the case here: even though $\mu_A + x$ and $\Phi_{B \rightarrow A}$ are globally Lipschitz, the product $(\mu_A + x)\Phi_{B \rightarrow A}(x)$ is not and the result of [DFH16] no longer applies. The key point to note is that, as $0 \leq \Phi_{B \rightarrow A} \leq 1$, one can stochastically dominate our process by a *linear Hawkes process* that is known to exist. It remains to perform on a thinning procedure on this dominating process to conclude (see [RL20, Prop. 1.4] or [CGMT20, Prop. 2] for similar ideas).

6.2.4 Long time dynamics of the mean-field limit

The nice thing about the mean-field limit (6.4) (compared to additive inhibition) is that one can easily analyse rigorously the large-time behavior of (6.4). In the linear case without inhibition (i.e. (6.1) with $\Phi(x) = \mu + x$), it is well-known that there is a phase transition, expressed in terms of $\kappa := \|h\|_{L^1} = \int_0^{+\infty} h(u) du$ (see [DFH16]): if $\kappa < 1$, the system is subcritical and $\lambda(t) \xrightarrow[t \rightarrow \infty]{} \frac{\mu}{1-\kappa}$ and if $\kappa > 1$, $\lambda(t) \xrightarrow[t \rightarrow \infty]{} +\infty$ (supercritical case). In our case, the relevant parameters become

$$\kappa_1 := \alpha \|h_1\|_{L^1}, \quad \kappa_2 := (1 - \alpha) \|h_2\|_{L^1}, \quad \kappa_3 := \alpha \|h_3\|_{L^1}. \quad (6.6)$$

The longtime behavior of system (6.4) depends strongly on the connectivity between populations A and B : crucial criteria are the absence/presence of inhibition from B to A ($\kappa_2 = 0$ or $\kappa_2 > 0$) and the absence/presence of retroaction from A to B ($\kappa_3 = 0$ or $\kappa_3 > 0$). We will not address here the easy three cases where the system is not fully-connected (one among κ_2 or κ_3 is zero) (see [L13] for a detailed account). We rather restrict to the more interesting case with full connectivity that reveals richer dynamical patterns.

Assumption 6.2.2 (Full connectivity). *Suppose that the system (6.3) is fully connected that is $\kappa_2 > 0$ (inhibition from B to A) and $\kappa_3 > 0$ (retroaction from A to B).*

The first main result is

Theorem 6.2.3 (Th. 3.6 in [L13]). *Under the previous assumptions and in particular Assumption 6.2.2, then, whatever the values of κ_i , $i = 1, \dots, 3$, λ_t^A does not tend to $+\infty$ as $t \rightarrow \infty$.*

Proof of Theorem 6.2.3 is elementary and basically relies on a contradiction argument: to put it simply, if one would have that $\lambda_t^A \xrightarrow[t \rightarrow \infty]{} +\infty$ then from the second line of (6.4), we see that λ_t^B would also go to ∞ . Using now that $\Phi_{B \rightarrow A}(x) \xrightarrow[x \rightarrow \infty]{} 0$, the first line of (6.4) goes to 0 as $t \rightarrow \infty$, which is in contradiction with the fact that $\lambda_t^A \xrightarrow[t \rightarrow \infty]{} +\infty$ (this is a bit more complicated than that, since the first line is actually an indeterminate form $\infty = \infty \times 0$, but this can be rigorously solved, see [L13, § 8.2] for details). Note here that the present argument uses in a crucial way that the kernels h are nonnegative. Hence, population A cannot be supercritical, even for arbitrary large values of κ_1 : it may be possible that in absence of inhibition, population

A is supercritical, but inhibition and retroaction brings back population A into subcriticality. This result reflects biological observations: episodes of *epileptic seizures*, involving abnormal synchronous firing of large groups of neurons [Luo20, p. 541], constitute an actual example of population A becoming supercritical when recurrent inhibition from population B is altered or suppressed.

The second and main result of [L13] is concerned with the existence of a limit of $(\lambda_t^A, \lambda_t^B)$ as $t \rightarrow \infty$. The result being technical, we only give a weak simplified statement of the main result and refer to [L13] for details.

Theorem 6.2.4 (Th. 3.10 in [L13]). *Under the previous assumptions and in particular Assumption 6.2.2, if $\mu_A > 0$, then, if one among $(\lambda_t^A, \lambda_t^B)$ converges as $t \rightarrow \infty$, both do, and their limit are necessarily given by $\ell_A = \frac{\mu_A \Phi_{B \rightarrow A}(\kappa_2 \ell_B)}{1 - \kappa_1 \Phi_{B \rightarrow A}(\kappa_2 \ell_B)}$ and $\ell_B = \ell$, unique fixed-point of Φ defined by*

$$\Phi(\ell) := \mu_B + \frac{\kappa_3 \mu_A \Phi_{B \rightarrow A}(\kappa_2 \ell)}{1 - \kappa_1 \Phi_{B \rightarrow A}(\kappa_2 \ell)}. \quad (6.7)$$

Moreover, there are sufficient conditions (depending on $\mu_A, \mu_B, \Phi_{B \rightarrow A}, \kappa_1, \kappa_2$ and κ_3) such that this convergence holds.

The proof of Theorem 6.2.4 is based on the computation of both $\underline{\lambda}^t := \liminf_{t \rightarrow \infty} \lambda_t^t$ and $\bar{\lambda}^t := \limsup_{t \rightarrow \infty} \lambda_t^t$ for $\iota \in \{A, B\}$ in (6.4). The point is to take advantage of the positivity of the kernels h : in the dynamics of λ^A in (6.4), the first part $\lambda^A \mapsto \mu_A + \alpha \int_0^t h_1(t-u) \lambda_u^A du$ is nondecreasing in λ^A , whereas the second term $\Phi_{B \rightarrow A}((1-\alpha) \int_0^t h_2(t-u) \lambda_u^B du)$ is nonincreasing in λ^B . Hence, $\underline{\lambda}^A, \underline{\lambda}^B, \bar{\lambda}^A, \bar{\lambda}^B$ are linked in an intricate way: one has

$$\begin{cases} \bar{\lambda}^A & \leq (\mu_A + \kappa_1 \bar{\lambda}^A) \Phi_{B \rightarrow A}(\kappa_2 \underline{\lambda}^B), \\ \underline{\lambda}^A & \geq (\mu_A + \kappa_1 \underline{\lambda}^A) \Phi_{B \rightarrow A}(\kappa_2 \bar{\lambda}^B) \\ \bar{\lambda}^B & \leq \mu_B + \kappa_3 \bar{\lambda}^A, \\ \underline{\lambda}^B & \geq \mu_B + \kappa_3 \underline{\lambda}^A \end{cases} \quad (6.8)$$

A great deal of effort is spent in [L13] to verify that the quantities in (6.8) are actually *finite* (otherwise (6.8) is essentially empty). Letting this point aside, putting these inequalities together show that

$$\Phi(\bar{\lambda}^B) \leq \underline{\lambda}^B \leq \bar{\lambda}^B \leq \Phi(\underline{\lambda}^B), \quad (6.9)$$

so that the convergence of λ^B (and hence λ^A) mentioned in Theorem 6.2.4 holds if and only if the previous inequality (6.9) is an actual equality. A technical point here: in order to even *define* Φ in (6.7), one has to make sure that the denominator of (6.7) does not vanish as $\ell = \lim \lambda^B$. A point of the proof of [L13] is precisely to show that $\kappa_1 \Phi_{B \rightarrow A}(\kappa_2 \bar{\lambda}^B) < 1$ is always true. This may however not be the case for $\underline{\lambda}^B$: one may have that $\kappa_1 \Phi_{B \rightarrow A}(\kappa_2 \underline{\lambda}^B) \geq 1$, but this corresponds to a case where $\underline{\lambda}^B < \bar{\lambda}^B$: we would then not have convergence. In a word, *any possible limit of λ_t^B concerns cases where $\kappa_1 \Phi_{B \rightarrow A}(\kappa_2 \underline{\lambda}^B) < 1$ and is the fixed-point of (6.7)*. A priori bounds give that a sufficient condition for $\kappa_1 \Phi_{B \rightarrow A}(\kappa_2 \underline{\lambda}^B) < 1$ is

$$\kappa_1 \Phi_{B \rightarrow A} \left(\frac{\mu_B \kappa_2}{1 - \kappa_3} \right) < 1. \quad (6.10)$$

Condition (6.10) has an interesting biological interpretation: in the case where A would be disconnected from B ($\kappa_2 = 0$), A is then a standard linear Hawkes process, so that the transition between sub and supercritical behavior of A is at $\kappa_1^{\text{isolated}} := 1$. With nontrivial inhibition and retroaction, we see from (6.10) that the domain of validity for the convergence of λ^A becomes larger: it may happen that $\kappa_1 > \kappa_1^{\text{isolated}} = 1$ but that (6.10) remains valid: this is another illustration that inhibition has a tendency to quench supercriticality in neuronal systems.

Considering (6.9), it is natural to introduce the domain $\mathcal{U} := \{(u, v), \Phi(u) \leq u \leq v \leq \Phi(v)\}$: we have that $(\underline{\lambda}^B, \bar{\lambda}^B) \in \mathcal{U}$. In particular, an obvious sufficient condition (the one mentioned in Theorem 6.2.4 which can be found in [L13, § 5]) to turn (6.9) in an equality is to suppose that \mathcal{U} reduces to the singleton $\{(\ell, \ell)\}$. This last claim is in particular true under either of the following conditions: (i) Φ is a contraction or (ii) $\Phi \circ \Phi$ has a unique fixed-point. We refer to [L13, § 5] for a detailed analysis of these conditions and applications to concrete examples.

Remark 6.2.5. We have applied (see [L13, § 5.2]) the previous result to the case of a polynomial inhibition: fix $\beta, R > 0$ and consider

$$\Phi_{B \rightarrow A}(x) = \frac{1}{1 + \left(\frac{x}{R}\right)^\beta} \quad x \geq 0. \quad (6.11)$$

It is shown in [L13], Prop. 5.3, that the convergence result of Theorem 6.2.4 is valid under the following conditions: either (i) $\beta \in (0, 1]$ ^[1], or (ii) $\beta > 1, \kappa_1 < 1$, at least when R is sufficiently large (weak inhibition) or when R is sufficiently small (strong inhibition) or (iii) $\beta > 1, \kappa_1 \geq 1$, at least when R is sufficiently small (strong inhibition).

The difficulty here is that, at least numerically, convergence of λ^B seems to remain true even if $\{(\ell, \ell)\} \not\subseteq \mathcal{U}$: giving a necessary condition for convergence is out of reach of the present result.

6.2.5 Fluctuations and a test of inhibition

Fluctuation theorems for Hawkes processes (both in large population $n \rightarrow \infty$ and in large time $t \rightarrow \infty$) have already been considered in the literature [DFH16, DL17] and some of the existing results directly apply to our case: applying [DL17, Th 2], one obtains that under the present assumptions and some sufficient subcriticality condition, for any subsets $I_{k_A}^A \subset \{1, \dots, n_A\}$ and $I_{k_B}^B \subset \{n_A + 1, \dots, n\}$ of fixed cardinality k_A and k_B (independent of n)

$$\left(\sqrt{m_t^A} \left(\frac{Z_t^{i,A}}{m_t^A} - 1 \right), \sqrt{m_t^B} \left(\frac{Z_t^{j,B}}{m_t^B} - 1 \right) \right)_{i \in I_{k_A}^A, j \in I_{k_B}^B} \xrightarrow[\substack{(t,n) \rightarrow (\infty, \infty), \\ \frac{t}{n} \rightarrow 0}]{d} \mathcal{N}(0, I_{k_A + k_B}). \quad (6.12)$$

Having in mind testing procedures, an estimator $\widehat{\ell}_{n,T}$ for the limiting intensity of population A may be given through the following approximations: $\ell \underset{T \rightarrow \infty}{\approx} \lambda_T \underset{n \rightarrow \infty}{\approx} \frac{Z_{i,t}}{T} := \widehat{\ell}_{n,T}$, $1 \leq i \leq n$. $\widehat{\ell}_{n,T}$ can be computed from the observation of one neuron (recall that in the mean field limit, neurons within the same population are interchangeable) in a n -particle system of Hawkes processes observed in a large time T . From a statistical viewpoint, a statement as in (6.12) cannot be exploited directly as the unknown parameter ℓ does not explicitly appear and secondly the normalization of $Z_{i,t}$ by m_t is unknown.

What is instead required here (for instance to derive testing procedures and confidence intervals) is a result of the form $\sqrt{T}(\widehat{\ell}_{n,T} - \ell) \xrightarrow[(T,n) \rightarrow (\infty, \infty)]{d} K$ for some distribution K . This last result is at least true in the simple case of an isolated subcritical population A . Here, $\ell = \frac{\mu_A}{1 - \kappa_1}$ and the following decomposition holds

$$\sqrt{T}(\widehat{\ell}_{n,T} - \ell) = \sqrt{\frac{m_T}{T}} \sqrt{m_T} \left(\frac{Z_{i,T}}{m_T} - 1 \right) + \sqrt{T} \left(\frac{m_T}{T} - \frac{\mu_A}{1 - \kappa_1} \right) := \sqrt{\frac{m_T}{T}} I_{n,T}(1) + I_T(2). \quad (6.13)$$

From (6.12), it immediately follows that $I_{n,T}(1) \xrightarrow{d} \mathcal{N}(0, 1)$ as $(T, n) \rightarrow (\infty, \infty)$, $\frac{T}{n} \rightarrow 0$. The difficulty is here to control the remaining deterministic term $I_2(T)$, which requires to evaluate the rate at which m_T/T converges to its limit $\mu_A/(1 - \kappa_1)$. In this linear case, the answer may be found in [BDHM13, Lemma 5]: $I_T(2) = o(1)$ if $\int_0^\infty \sqrt{t} h(t) dt < \infty$, so that, we derive

$$\sqrt{\frac{T}{\widehat{\ell}_{n,T}}} (\widehat{\ell}_{n,T} - \ell) \xrightarrow[\substack{(T,n) \rightarrow (\infty, \infty), \\ \frac{T}{n} \rightarrow 0}]{d} \mathcal{N}(0, 1). \quad (6.14)$$

A test of inhibition

The convergence (6.14) is of practical use for building a testing procedure for the presence/absence of inhibition in a population of neurons. This is simply based on the fact that

[1]. Note that $\beta \in [0, 1]$ if and only if $\Phi_{B \rightarrow A}$ is convex.

there is a hierarchy of limits concerning the intensity of population A, depending on the presence/absence of inhibition and retroaction in the system (see [L13] for details): we have

$$\ell_A^{(I)} > \ell_A^{(II)} > \ell_A^{(III)}, \quad (6.15)$$

where $\ell_A^\nu := \lim_{t \rightarrow \infty} \lambda_t^A$, $\nu \in \{I, II, III\}$. Here (I) (resp. (II) and (III)) corresponds to the isolated case (no inhibition and retroaction), resp. inhibition without retroaction and the fully-coupled case. This hierarchy, together with the fluctuation result (6.14) allows to build a procedure detailed in [L13] testing H_0 : *Population B has no inhibitive effect on population A* against H_1 : *Population B has an inhibitive effect on population A*. In practice, neurophysiologists studying synaptic function routinely isolate one type of synaptic coupling with specific toxins, so that they can for instance suppress inhibition. In a experimental preparation, neurons from the A population can be continuously recorded in three successive conditions: *control* (no toxin applied); *Toxin* (inhibition blocked) and *wash* (the physiologists' term to designate a return to the control condition after removal of a toxin). Collecting for a duration T the activities of the neurons of population A leads to the estimators for the intensity of A in the three cases: $\widehat{\ell}_{n,T}^{Control}$, $\widehat{\ell}_{n,T}^{Toxin}$ and $\widehat{\ell}_{n,T}^{Wash}$. Denoting by $(\ell^{Control}, \ell^{Toxin})$ the *true* intensities of population A with and without inhibition blocked, the test becomes

$$H_0 : \ell^{Control} = \ell^{Toxin} \text{ against } H_1 : \ell^{Control} > \ell^{Toxin}.$$

Considering the test statistics $R_{n,T} = \widehat{\ell}_{n,T}^{Control} - \widehat{\ell}_{n,T}^{Toxin}$, where $\widehat{\ell}_{n,T}^{Control}$ and $\widehat{\ell}_{n,T}^{Toxin}$ are independent, we easily derive from (6.14), that under H_0

$$\sqrt{\frac{T}{\widehat{\ell}_{n,T}^{Control} + \widehat{\ell}_{n,T}^{Toxin}}} (\widehat{\ell}_{n,T}^{Control} - \widehat{\ell}_{n,T}^{Toxin}) \xrightarrow[\frac{T}{n} \rightarrow 0]{d, H_0, (T,n) \rightarrow (\infty, \infty)} \mathcal{N}(0, 1),$$

from which one can easily derive a consistent asymptotic test for inhibition (see [L13] for details).

6.2.6 Towards oscillations: inhibition through a sigmoidal kernel

Theorem 6.2.4 is concerned with convergence of the intensities of both populations A and B as $t \rightarrow \infty$. Contrary to the standard excitatory linear case, inhibition is also responsible for oscillations: oscillations is ubiquitous in neural systems, as it can be observed e.g. in the motor system but also for the activity of some parts of the cerebral cortex and hippocampus [Luo20, Chap. 8]. Oscillations are believed to play an important role in perception and cognition. We give in [L13] detailed non-rigorous intuition and numerical simulations indicating that system (6.3)-(6.4) presents oscillations. We only give a numerical illustration of this phenomenon in Figure 6.1 and refer to the paper for details. Note only oscillations are typically observed when the inhibition kernel $\Phi_{B \rightarrow A}$ is of sigmoid type (think for example to Remark 6.2.5, where $\Phi_{B \rightarrow A}$ in (6.11) approaches $\mathbf{1}_{[0,R]}$ as $\beta \rightarrow \infty$).

6.3 Perspectives

Recall that in the case of interacting diffusions, the microscopic particle system (1.1) is Markovian and its mean-field limit (1.5) is not. The situation is somehow converse in the Hawkes setting (think e.g. of the homogeneous case (6.1) for simplicity): the particle system (6.3) is generically not Markov (apart from particular cases where h is exponential or Erlang [CDLO19]) whereas the mean-field McKean-Vlasov process \bar{Z}_t is a (Markovian) time-inhomogeneous Poisson point process, whose intensity is given by (6.2). A nice feature of Hawkes processes is that the dynamics of the mean-field process \bar{Z} is intrinsically *finite-dimensional* (as its reduces to (6.2)), that is in itself a considerable simplification w.r.t. the considerations of the previous chapters in the diffusion case. These two observations indicate that there is much more structure for the analysis of mean-field nonlinear processes in the Hawkes setting and we are inclined to think of Hawkes processes as easy examples where to apply the long-term stability techniques that we reviewed in the previous chapters of this manuscript.

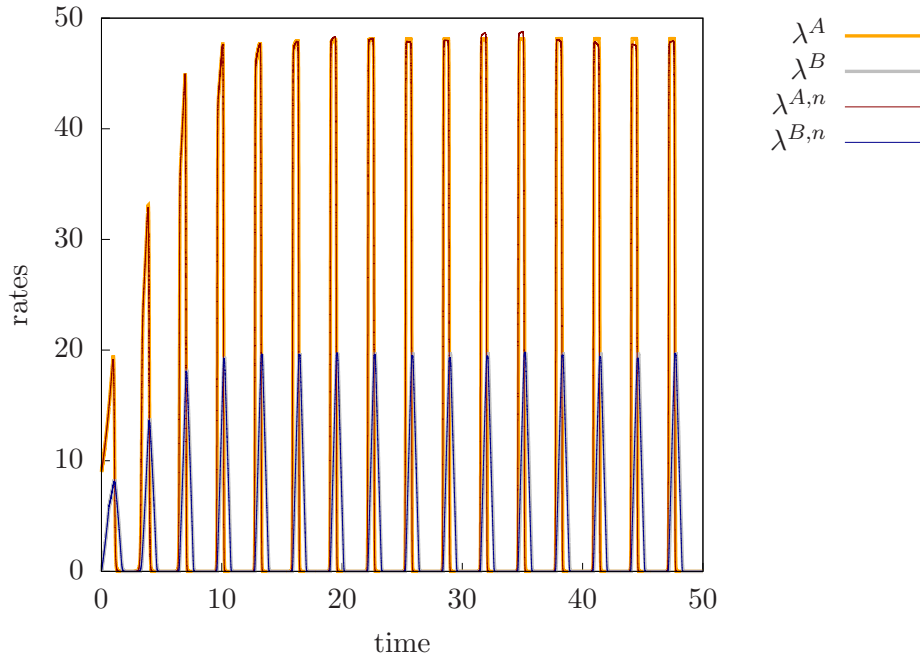


Figure 6.1 – Trajectories of the microscopic intensities $(\lambda_t^{A,n}, \lambda_t^{B,n})$ (6.3) and their mean-field counterparts $(\lambda_t^A, \lambda_t^B)$ (6.4). The inhibition kernel $\Phi_{B \rightarrow A}$ is given by (6.11) with $R = 1$, $\beta = 20$. The synaptic kernels are with compact support $h_i = \mathbf{1}_{[0, \theta_i]}$, $i = 1, 2, 3$. The parameters are $\alpha = 0.8$, $\kappa_1 = 5$, $\kappa_2 = \kappa_3 = 0.5$, $\mu_A = 9$ and $\mu_B = 0$, on $[0, T]$ with $T = 50$ and $n = 5000$.

6.3.1 Stability of homogeneous Hawkes processes

Even in the simple homogeneous case (6.1)-(6.2), several questions remain unanswered. Two combined difficulties arise here: the possibility that h might take negative values (in case of additive inhibition in the sense of [CGMT20]) and the possible nonlinearity of the synaptic kernel Φ .

Long-time behavior of the mean-field process

Since [Fel41, Lon74], several open questions remain about the longtime behavior of the NRE (6.2). In the purely excitatory case ($h \geq 0$), it is easy to see that if $\lambda(t) \xrightarrow{t \rightarrow \infty} \ell$, then ℓ is necessarily a fixed-point to $\ell = \Phi(\|h\|_{L^1} \ell)$. For linear Hawkes process $\Phi(x) = \mu + x$ the above convergence is a standard result, under the optimal condition $\kappa := \|h\|_{L^1} < 1$ [DFH16]. This convergence has been generalised in [AN22a] in the case Φ is Lipschitz under the generic condition $\|\Phi\|_{Lip} \|h\|_{L^1} < 1$. It is however very likely that this condition may not be optimal (this is quite clear when one thinks of the case where Φ is sigmoid: $\ell = \Phi(\|h\|_{L^1} \ell)$ may have one or several solutions whereas $\|\Phi\|_{Lip}$ is very large). Let us here mention a few open problems: (i) in case $\ell = \Phi(\|h\|_{L^1} \ell)$ has a unique solution, [Lon74] readily implies the convergence of $\lambda(t)$ to ℓ . Is it possible to obtain rates of convergence in this case? (ii) what about the case where the previous fixed-point relation has several solutions (think of a sigmoid Φ)? can we characterise convergence in this case? This is quite clear when h is exponential, but the same question with general h is open. The situation remains even less clear in situations with additive inhibition (i.e. when one relaxes the hypothesis $h \geq 0$ with $\Phi \geq 0$ nonlinear). Numerical simulation suggest the possibility of oscillations but the conditions for convergence or oscillations for $\lambda(t)$ remain largely open.

Fluctuations

The issue of fluctuations (w.r.t. both time t and size of population n) for various functionals of Hawkes processes is not fully elucidated. A first question concerns the extension of the

statistically useful CLT mentioned in § 6.2.5 to the nonlinear case: extending (6.12) to (6.14) requires a proper control on the rate of convergence of $\frac{m_t}{t}$ to ℓ (recall $I_T(2)$ in (6.13)). This has been done only so far in the linear case, but extensions to the general nonlinear case are expected.

A key argument for (6.14) was the original CLT results of [DFH16, DL17] of the form $\sqrt{m_t} \left(\frac{Z_{1,t}}{m_t} - 1 \right) \rightarrow \mathcal{N}(0, 1)$. The domain of validity (w.r.t. t and n) of these results remain unclear: [DFH16] restrict to linear Hawkes processes (and the convergence holds with no restriction on $(t, n) \rightarrow (+\infty, +\infty)$) whereas the result of [DL17] is valid for generic nonlinear Hawkes processes but under the restrictive condition $(t, n) \rightarrow (+\infty, +\infty)$ with $\frac{t}{n} \rightarrow \infty$. What is the regime of validity of fluctuations in the fully nonlinear case? is the condition $\frac{t}{n} \rightarrow 0$ really relevant? is there something specific happening on a time scale proportional to n^α for some $\alpha \geq 1$? what is the specificity of the nonlinearity of Φ on this matter?

An alternative approach for fluctuation results was carried out by Heesen and Stannat [HS21] by looking at $\sqrt{n} \left(\lambda_t^{(n)} - \lambda_t \right)$ as $n \rightarrow \infty$ on bounded time interval $[0, T]$. The question of the long-time behavior of the resulting fluctuation process and its connection with the CLT results of the previous paragraph are unclear.

Long term stability and KMT coupling

The previous questions on fluctuations have actually strong connections with the problem of stability of Hawkes processes on long time scales. The general setting is similar to Statement 5.0.1 in Chapter 5: suppose that one is able to characterise the behavior of the mean-field intensity $\lambda(t)$ as $t \rightarrow \infty$. It may be the convergence to some $\ell = \Phi(\|h\|_{L^1} \ell)$ but one is also interested in situations when $\lambda(t)$ has a periodic behavior. We believe that it is indeed the case for the multiplicative model (6.4) (Figure 6.1) but we have also in mind the result of Ditlevsen and Löcherbach [DL17] where the existence of limit cycles for the mean-field limit of the potential $X^{n,i}$ as in (5.16) in case the graph of interaction has circular connectivity and with Erlang kernels h is proven. Provided these macroscopic structures have sufficiently stability properties, one is left with proving long-term stability results in a same way as done in Chapter 5 for diffusions. A first step of this programme has been carried-out by Zoé Agathe-Nerine: in case h is exponential and under the condition $\|\Phi\|_{Lip} \|h\|_{L^1} < 1$ (hence a unique limit ℓ for $\lambda(t)$), [AN22a] proves the proximity of λ_n to ℓ up to arbitrary polynomial times in n .

An important question on this matter concerns the specificity of exponential kernels h for the dynamics of (6.1). Choosing exponential (or Erlang) kernels is a considerable and nice simplification to the analysis of (6.1), as it allows to use SDEs and Markov structure [DLO19, AN22a]. But how much is it essential for the intrinsic dynamics? Exponential kernels are in fact some kind of middle ground for the memory structure of an Hawkes process: the support is infinite, but the influence of past events decreases fast. Open questions are numerous: (i) in what way the stability results that exist for exponential kernels are susceptible to perturbations? (ii) in case $\ell = \Phi(\|h\|_{L^1} \ell)$ has a unique solution, can we expect long-term stability results for arbitrary h ? (iii) what can we say in the crucial example when $\ell \mapsto \Phi(\|h\|_{L^1} \ell)$ is sigmoid with three fixed-point when h is not exponential (numerical simulations in [HS21, p. 16] suggest metastability for Erlang kernels, but this has not been proven)? (iv) what is the specificity of the nonlinearity of Φ on this matter?

Another interesting point that is not fully elucidated concerns the use of Brownian approximations for Hawkes dynamics such as (6.1) (or inhomogeneous extensions of it). A diffusion approximation for Hawkes processes was first proposed in [DL17, § 5] (using Markov semigroup methods to quantify the proximity in law between the two processes on a bounded time interval). This has been further developed in [CMT20] where a pathwise coupling between a Hawkes system and its diffusive counterpart is proven, again on a bounded time interval. This coupling uses in an intricate way standard Komlós-Major-Tusnády (KMT) [KMT75] coupling between Poisson and Brownian paths. Note that similar KMT coupling has been used in [Pro23] for Markov chains. As far as long-term stability is concerned, we see here two equally interesting ways of addressing this KMT coupling: (i) either apply the previous long-term stability

methods to the Brownian approximation (but in this case the question of the stability of the original Hawkes process remains) or (ii) to take advantage of the known long-term stability for the original Hawkes process (as done in [AN22a]) to derive the same stability for its diffusive approximation, hence extending on unbounded times the validity of the KMT coupling.

6.3.2 Multiplicative inhibition: oscillations and long-term stability

The multiplicative structure for inhibition introduced in the present chapter suggests the possibility of bypassing the usual difficulties of models with additive inhibition.

Existence and stability of oscillations

An obvious perspective of § 6.2.6 concerns a rigorous proof for the existence of oscillations for the intensity (λ^A, λ^B) in (6.4). One needs here to understand the role of the kernels h_i and the precise domain validity of the oscillations. The class of exponential (or Erlang) kernels h has been particularly considered in the literature [CDLO19, DLO19, DL17], as it has the nice ability to transform the convolution equation (6.2) governing the mean-field intensity $\lambda(t)$ into an ODE governing the mean activity $X(t) = \int_0^t h(t-s)\lambda(s)ds$ (see e.g. [DLO19, DL17] for further details). However, this may not be a good choice for (6.4): numerical simulations suggest that, oscillations that are present with compactly supported h_i would turn into damped oscillations with exponential kernels with the same κ_i : the presence/absence of oscillations depend crucially on the whole shape of these kernels (not simply on the values of κ_i as for the convergence case). One would be more optimistic for a proof of oscillations when $h_i = \mathbf{1}_{[0, \theta_i]}$, as (6.4) transforms into *ODEs with delays* for the mean activities X^A, X^B . Applying the machinery of delay equations [HL93, Nus79] seems here a natural perspective.

Once the existence and stability of these periodic orbits established, a natural question would be of course to address the same long-term asymptotics of the Hawkes process along these periodic orbits, adapting the techniques of Chapter 5 in the diffusion case. The question of CLT results remain also open.

Renewal properties

Several important results concerning Hawkes processes with additive inhibition have been proposed in [CGMT20, CCC22], where the renewal structure of the trajectories of one single Hawkes process Z_t with additive self-inhibition is considered. The authors derive ergodic results of the form $\frac{Z_t}{t} \xrightarrow[t \rightarrow \infty]{a.s.} \frac{\mathbb{E}(W)}{\mathbb{E}(\tau)}$, where (W, τ) are random variables capturing the renewal structure of the trajectory of Z . Very few is known on the law of (W, τ) apart from specific examples. The strategy would be to derive the same analysis but now for a Hawkes with *multiplicative self-inhibition*, the hope being that the multiplicative structure would allow for a better understanding of this renewal structure.

6.3.3 Introducing inhomogeneities

Spatially-extended systems

The previous questions concern Hawkes processes with homogeneous type interaction. We already discussed at length in § 5.3.2 of the importance of considering Hawkes processes with spatial extension and to understand properly the links of these models with the Neural Field Equation (5.17). Some advances are expected on this matter, based on the transposition of the techniques of this manuscript to the Hawkes dynamics.

The case of sparse interaction

A particular case of [AN22b] is to look at the convergence of mean-field Hawkes processes interacting on a Erdős-Rényi graph with parameter p_n , provided that $np_n \gg \log n$. The questions remains of the large population behavior of such system in the sparse regime $p_n \sim \frac{\lambda}{n}$. Extending

the existing results [LRW23, ORS20] for diffusions to the Hawkes setting seems reasonable. One would then obtain Hawkes processes interacting on Galton-Watson trees, whose dynamics is completely open so far (see [DFH16] for a similar analysis of Hawkes processes on \mathbb{Z}^d).

Bibliography

- [ABK12] D. C. Antonopoulou, D. Blömker, and G. D. Karali. Front motion in the one-dimensional stochastic Cahn-Hilliard equation. *SIAM J. Math. Anal.*, 44(5):3242–3280, 2012.
- [ABPV⁺05] J. A. Acebrón, L. L. Bonilla, Conrad J. Pérez Vicente, Félix Ritort, and Renato Spigler. The Kuramoto model: a simple paradigm for synchronization phenomena. *Rev. Mod. Phys.*, 77:137–185, Apr 2005.
- [ABR04] J. A. Acebrón, A. R. Bulsara, and W.-J. Rappel. Noisy FitzHugh-Nagumo model: from single elements to globally coupled networks. *Phys. Rev. E (3)*, 69(2):026202, 9, 2004.
- [AF03] Robert A. Adams and John J. F. Fournier. *Sobolev spaces*, volume 140 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, second edition, 2003.
- [AM19a] Michele Aleandri and Ida G. Minelli. Opinion dynamics with lotka-volterra type interactions. *Electron. J. Probab.*, 24:31 pp., 2019.
- [AM19b] Marc Arnaudon and Pierre Del Moral. A variational approach to nonlinear and interacting diffusions. *Stochastic Analysis and Applications*, 37(5):717–748, May 2019.
- [Ama77] Shun-Ichi Amari. Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.*, 27(2):77–87, June 1977.
- [Ami97] D. Amit. Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cerebral Cortex*, 7(3):237–252, April 1997.
- [AN06] Noga Alon and Assaf Naor. Approximating the cut-norm via grothendieck’s inequality. *SIAM Journal on Computing*, 35(4):787–803, January 2006.
- [AN22a] Zoé Agathe-Nerine. Long-term stability of interacting Hawkes processes on random graphs. <https://arxiv.org/abs/2207.13942>, 2022.
- [AN22b] Zoé Agathe-Nerine. Multivariate Hawkes processes on inhomogeneous random graphs. *Stochastic Processes and their Applications*, 152:86–148, October 2022.
- [APF18] Luisa Andreis, Paolo Dai Pra, and Markus Fischer. McKean–vlasov limit for interacting systems with simultaneous jumps. *Stochastic Analysis and Applications*, 36(6):960–995, November 2018.
- [APV10] V. Araujo, M.J. Pacifico, and M. Viana. *Three-Dimensional Flows*. Springer Berlin Heidelberg, 2010.
- [ARM76] K B Athreya and K Rama Murthy. Feller’s renewal theorem for systems of renewal equations. *Journal of the Indian Institute of Science*, 58(10), 1976.
- [Aro68] D. G. Aronson. Non-negative solutions of linear parabolic equations. *Ann. Scuola Norm. Sup. Pisa (3)*, 22:607–694, 1968.
- [AS06] Daniel M. Abrams and Steven H. Strogatz. Chimera states in a ring of nonlocally coupled oscillators. *International Journal of Bifurcation and Chaos*, 16(01):21–37, 2006.
- [BAG97] G Ben Arous and A Guionnet. Symmetric Langevin spin glass dynamics. *The Annals of Probability*, pages 1367–1422, 1997.

- [BB98] Stella Brassesco and Paolo Buttà. Interface fluctuations for the $D = 1$ stochastic Ginzburg-Landau equation with nonsymmetric reaction term. *J. Statist. Phys.*, 93(5-6):1111–1142, 1998.
- [BBW19] Shankar Bhamidi, Amarjit Budhiraja, and Ruoyu Wu. Weakly interacting particle systems on inhomogeneous random graphs. *Stochastic Processes and their Applications*, 129(6):2174 – 2206, 2019.
- [BCCZ18] Christian Borgs, Jennifer T. Chayes, Henry Cohn, and Yufei Zhao. An L^p theory of sparse graph convergence II: LD convergence, quotients and right convergence. *Ann. Probab.*, 46(1):337–396, 01 2018.
- [BCCZ19] Christian Borgs, Jennifer T. Chayes, Henry Cohn, and Yufei Zhao. An L^p theory of sparse graph convergence I: Limits, sparse random graph models, and power law distributions. *Transactions of the American Mathematical Society*, 372(5):3019–3062, May 2019.
- [BCD20] I. Bailleul, R. Catellier, and F. Delarue. Propagation of chaos for mean field rough differential equations. *ArXiv e-print 1907.00578*, 2020.
- [BCL⁺08] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi. Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing. *Adv. Math.*, 219(6):1801–1851, 2008.
- [BCL⁺11] Christian Borgs, Jennifer Chayes, László Lovász, Vera Sós, and Katalin Vesztergombi. Limits of randomly grown graph sequences. *European J. Combin.*, 32(7):985–999, 2011.
- [BCL⁺12] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi. Convergent sequences of dense graphs II. Multiway cuts and statistical physics. *Ann. of Math. (2)*, 176(1):151–219, 2012.
- [BCN20] Gianmarco Bet, Fabio Coppini, and Francesca R. Nardi. Weakly interacting oscillators on dense random graphs. *ArXiv e-print 2006.07670*, 2020.
- [BCnCD16] Alethea B. T. Barbaro, José A. Cañizo, José A. Carrillo, and Pierre Degond. Phase transitions in a kinetic flocking model of Cucker-Smale type. *Multiscale Model. Simul.*, 14(3):1063–1088, 2016.
- [BCW20] Erhan Bayraktar, Suman Chakraborty, and Ruoyu Wu. Graphon mean field systems. *ArXiv e-print 2003.13180*, 2020.
- [BDF12] Amarjit Budhiraja, Paul Dupuis, and Markus Fischer. Large deviation properties of weakly interacting processes via weak convergence methods. *Ann. Probab.*, 40(1):74–102, 2012.
- [BDHM13] E. Bacry, S. Delattre, M. Hoffmann, and J. F. Muzy. Some limit theorems for Hawkes processes and application to financial statistics. *Stochastic Process. Appl.*, 123(7):2475–2499, 2013.
- [BDMP95] S. Brassesco, A. De Masi, and E. Presutti. Brownian fluctuations of the interface in the $D = 1$ Ginzburg-Landau equation with noise. *Ann. Inst. H. Poincaré Probab. Statist.*, 31(1):81–118, 1995.
- [BFFT12] Javier Baladron, Diego Fasoli, Olivier Faugeras, and Jonathan Touboul. Mean-field description and propagation of chaos in networks of Hodgkin-Huxley and FitzHugh-Nagumo neurons. *The Journal of Mathematical Neuroscience*, 2(1):10, 2012.
- [BFO19] Mireille Bossy, Joaquín Fontbona, and Héctor Olivero. Synchronization of stochastic mean field networks of Hodgkin–Huxley neurons with noisy channels. *Journal of Mathematical Biology*, 78(6):1771–1820, May 2019.
- [BFT15] Mireille Bossy, Olivier Faugeras, and Denis Talay. Clarification and complement to “Mean-field description and propagation of chaos in networks of Hodgkin-Huxley and FitzHugh-Nagumo neurons”. *J. Math. Neurosci.*, 5:Art. 19, 23, 2015.

- [BG10] Peter H. Baxendale and Priscilla E. Greenwood. Sustained oscillations for density dependent markov processes. *Journal of Mathematical Biology*, 63(3):433–457, November 2010.
- [BGG12] François Bolley, Ivan Gentil, and Arnaud Guillin. Convergence to equilibrium in wasserstein distance for fokker–planck equations. *Journal of Functional Analysis*, 263(8):2430 – 2457, 2012.
- [BGL14] Dominique Bakry, Ivan Gentil, and Michel Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Cham, 2014.
- [BGM10] François Bolley, Arnaud Guillin, and Florent Malrieu. Trend to equilibrium and particle approximation for a weakly selfconsistent Vlasov-Fokker-Planck equation. *M2AN Math. Model. Numer. Anal.*, 44(5):867–884, 2010.
- [BGP10] Lorenzo Bertini, Giambattista Giacomin, and Khashayar Pakdaman. Dynamical aspects of mean field plane rotators and the Kuramoto model. *J. Statist. Phys.*, 138:270–290, 2010.
- [BGP14] Lorenzo Bertini, Giambattista Giacomin, and Christophe Poquet. Synchronization and random long time dynamics for mean-field plane rotators. *Probability Theory and Related Fields*, 160(3):593–653, Dec 2014.
- [BH99] N Brunel and V Hakim. Fast global oscillations in networks of integrate-and-fire neurons with low firing rates. *Neural Computation*, 11(7):1621–1671, Oct 1999.
- [BHS21] Anna Bonnet, Miguel Martinez Herrera, and Maxime Sangnier. Maximum likelihood estimation for Hawkes processes with self-excitation or inhibition. *Statistics & Probability Letters*, 179:109214, December 2021.
- [Bil99] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999.
- [Ble11] Ron Blei. The Grothendieck Inequality Revisited. *arXiv e-prints*, page arXiv:1111.7304, November 2011.
- [BLZ98] P.W. Bates, K. Lu, and C. Zeng. *Existence and persistence of invariant manifolds for semiflows in Banach space.*, volume 135. Mem. Amer. Math. Soc., 1998.
- [BLZ08] Peter W. Bates, Kening Lu, and Chongchun Zeng. Approximately invariant manifolds and global dynamics of spike states. *Inventiones mathematicae*, 174(2):355–433, July 2008.
- [BM96] Pierre Brémaud and Laurent Massoulié. Stability of nonlinear Hawkes processes. *Ann. Probab.*, 24(3):1563–1588, 1996.
- [Bol86] E. Bolthausen. Laplace approximations for sums of independent random vectors. *Probab. Theory Relat. Fields*, 72(2):305–318, 1986.
- [Bol01] Béla Bollobás. *Random graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2001.
- [Bor16] Charles Bordenave. Lecture notes on random graphs and probabilistic combinatorial optimization. <https://www.math.univ-toulouse.fr/~bordenave/coursRG.pdf>, April 2016.
- [BPG⁺03] Stéphane Barland, Oreste Piro, Massimo Giudici, Jorge R. Tredicce, and Salvador Balle. Experimental evidence of van der Pol–FitzHugh–Nagumo dynamics in semiconductor optical amplifiers. *Phys. Rev. E*, 68:036209, Sep 2003.
- [BPR22] Rangel Baldasso, Alan Pereira, and Guilherme Reis. Large deviations for interacting diffusions with path-dependent McKean–vlasov limit. *The Annals of Applied Probability*, 32(1), February 2022.
- [Bre83] J. Bretagnolle. Lois limites du bootstrap de certaines fonctionnelles. *Annales de l’institut Henri Poincaré (B) Probabilités et Statistiques*, 19(3):281–296, 1983.

- [Bre12] Paul C. Bressloff. Spatiotemporal dynamics of continuum neural fields. *J. Phys. A*, 45(3):033001, 109, 2012.
- [Bre14a] Paul C. Bressloff. *Stochastic processes in cell biology*, volume 41 of *Interdisciplinary Applied Mathematics*. Springer, Cham, 2014.
- [Bre14b] Paul C. Bressloff. *Waves in neural media*. Lecture Notes on Mathematical Modelling in the Life Sciences. Springer, New York, 2014. From single neurons to neural fields.
- [Bri88] D. R. Brillinger. Maximum likelihood analysis of spike trains of interacting nerve cells. *Biological Cybernetics*, 59(3):189–200, August 1988.
- [BS98] Valentino Braitenberg and Almut Schüz. *Cortex: Statistics and Geometry of Neuronal Connectivity*. Springer Berlin Heidelberg, 1998.
- [BS00] Neil J. Balmforth and Roberto Sassi. A shocking display of synchrony. *Physica D: Nonlinear Phenomena*, 143(1-4):21–55, 2000.
- [Buc88] John Buck. Synchronous rhythmic flashing of fireflies. ii. *The Quarterly Review of Biology*, 63(3):pp. 265–289, 1988.
- [Buz04] G. Buzsáki. Neuronal oscillations in cortical networks. *Science*, 304(5679):1926–1929, June 2004.
- [Buz06] György Buzsáki. *Rhythms of the Brain*. Oxford University Press, October 2006.
- [BW12] Paul C. Bressloff and Matthew A. Webber. Front propagation in stochastic neural fields. *SIAM Journal on Applied Dynamical Systems*, 11(2):708–740, January 2012.
- [CCC22] Patrick Cattiaux, Laetitia Colombani, and Manon Costa. Limit theorems for Hawkes processes including inhibition. *Stochastic Processes and their Applications*, 149:404–426, July 2022.
- [CD22] Louis-Pierre Chaintron and Antoine Diez. Propagation of chaos: A review of models, methods and applications. i. models and methods. *Kinetic and Related Models*, 15(6):895, 2022.
- [CDF15] Francesca Collet, Paolo Dai Pra, and Marco Formentin. Collective periodicity in mean-field models of cooperative behavior. *Nonlinear Differential Equations and Applications NoDEA*, 22(5):1461–1482, 10 2015.
- [CDFM20] Michele Coghi, Jean-Dominique Deuschel, Peter K. Friz, and Mario Maurelli. Pathwise McKean–Vlasov theory with additive noise. *The Annals of Applied Probability*, 30(5), Oct 2020.
- [CDG20] Fabio Coppini, Helge Dietert, and Giambattista Giacomin. A law of large numbers and large deviations for interacting diffusions on Erdős–Rényi graphs. *Stochastics and Dynamics*, 20(02):2050010, July 2020.
- [CDLO19] J. Chevallier, A. Duarte, E. Löcherbach, and G. Ost. Mean field limits for nonlinear spatially extended Hawkes processes with exponential memory kernels. *Stochastic Processes and their Applications*, 129(1):1 – 27, 2019.
- [CDS10] Francesca Collet, Paolo Dai Pra, and Elena Sartori. A simple mean field model for social interactions: dynamics, fluctuations, criticality. *J. Stat. Phys.*, 139(5):820–858, 2010.
- [Cer01] Sandra Cerrai. *Second order PDE’s in finite and infinite dimension*, volume 1762 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2001. A probabilistic approach.
- [CFF19] Joachim Crevat, Grégory Faye, and Francis Filbet. Rigorous Derivation of the Nonlocal Reaction-Diffusion Fitzhugh–Nagumo System. *SIAM Journal on Mathematical Analysis*, 51(1):346–373, 2019.
- [CFT16] Francesca Collet, Marco Formentin, and Daniele Tovazzi. Rhythmic behavior in a two-population mean-field Ising model. *Physical Review E*, 94(4), October 2016.

- [CG19] René Carmona and Christy V. Graves. Jet lag recovery: Synchronization of circadian oscillators as a mean field game. *Dynamic Games and Applications*, 10(1):79–99, May 2019.
- [CGM08] P. Cattiaux, A. Guillin, and F. Malrieu. Probabilistic approach for granular media equations in the non-uniformly convex case. *Probability Theory and Related Fields*, 140(1):19–40, Jan 2008.
- [CGMT20] Manon Costa, Carl Graham, Laurence Marsalle, and Viet Chi Tran. Renewal in Hawkes processes with self-excitation and inhibition. *Advances in Applied Probability*, 52(3):879–915, 2020.
- [Che17a] Julien Chevallier. Fluctuations for mean-field interacting age-dependent Hawkes processes. *Electronic Journal of Probability*, 22(0), 2017.
- [Che17b] Julien Chevallier. Mean-field limit of generalized Hawkes processes. *Stochastic Processes and their Applications*, 127(12):3870 – 3912, 2017.
- [CL94] Patrick Cattiaux and Christian Léonard. Minimization of the Kullback information of diffusion processes. *Ann. Inst. H. Poincaré Probab. Statist.*, 30(1):83–132, 1994.
- [CL95] Patrick Cattiaux and Christian Léonard. Large deviations and Nelson processes. *Forum Math.*, 7(1):95–115, 1995.
- [CL02] Fan Chung and Linyuan Lu. Connected components in random graphs with given expected degree sequences. *Annals of Combinatorics*, 6(2):125–145, November 2002.
- [CL09] Stephen Coombes and Carlo Laing. Delays in activity-based neural networks. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 367(1891):1117–1129, February 2009.
- [CM19] Hayato Chiba and Georgi S. Medvedev. The mean field analysis of the Kuramoto model on graphs I. The mean field equation and transition point formulas. *Discrete & Continuous Dynamical Systems - A*, 39(1):131–155, 2019.
- [CMT20] Julien Chevallier, Anna Melnykova, and Irene Tubikanec. Theoretical analysis and simulation methods for hawkes processes and their diffusion approximation. *ArXiv e-print 2003.10710*, 2020.
- [CMV03] José A. Carrillo, Robert J. McCann, and Cédric Villani. Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates. *Rev. Mat. Iberoamericana*, 19(3):971–1018, 12 2003.
- [CO20] Julien Chevallier and Guilherme Ost. Fluctuations for spatially extended Hawkes processes. *Stochastic Processes and their Applications*, 130(9):5510–5542, September 2020.
- [Cop22] Fabio Coppini. Long time dynamics for interacting oscillators on graphs. *The Annals of Applied Probability*, 32(1), February 2022.
- [Cor21] Quentin Cormier. A mean-field model of integrate-and-fire neurons: non-linear stability of the stationary solutions. *ArXiv e-print 2002.08649*, 2021.
- [CR16] Francesca Collet and Wioletta Ruszel. Synchronization and spin-flop transitions for a mean-field xy model in random field. *Journal of Statistical Physics*, pages 1–22, 2016.
- [CT18a] Tanguy Cabana and Jonathan D. Touboul. Large deviations for randomly connected neural networks: I. spatially extended systems. *Advances in Applied Probability*, 50(3):944–982, September 2018.
- [CT18b] Tanguy Cabana and Jonathan D. Touboul. Large deviations for randomly connected neural networks: II. state-dependent interactions. *Advances in Applied Probability*, 50(3):983–1004, September 2018.

- [CTV21] Quentin Cormier, Etienne Tanré, and Romain Veltz. Hopf bifurcation in a mean-field model of spiking neurons. *Electronic Journal of Probability*, 26:1–40, January 2021.
- [DdH96] Paolo Dai Pra and Frank den Hollander. McKean-Vlasov limit for interacting random processes in random media. *J. Statist. Phys.*, 84(3-4):735–772, 1996.
- [DDS20] Jean-François Delmas, Jean-Stéphane Dhersin, and Marion Sciauveau. Asymptotic for the cumulative distribution function of the degrees and homomorphism densities for random graphs sampled from a graphon. *Random Structures and Algorithms*, 58(1):94–149, October 2020.
- [DEGZ20] Alain Durmus, Andreas Eberle, Arnaud Guillin, and Raphael Zimmer. An Elementary Approach To Uniform In Time Propagation Of Chaos. In *Proc. Amer. Math. Soc.*, number 148.12, pages 5387–5398. American Mathematical Society, May 2020.
- [DFGV16] Helge Dietert, Bastien Fernandez, and David Gérard-Varet. Landau damping to partially locked states in the kuramoto model. *arXiv preprint arXiv:1606.04470*, 2016.
- [DFH16] Sylvain Delattre, Nicolas Fournier, and Marc Hoffmann. Hawkes processes on large networks. *Ann. Appl. Probab.*, 26(1):216–261, 2016.
- [DFL12] Pierre Degond, Amic Frouvelle, and Jian-Guo Liu. Macroscopic limits and phase transition in a system of self-propelled particles. *Journal of Nonlinear Science*, pages 1–30, 2012.
- [DFMS18] Jean-Dominique Deuschel, Peter K. Friz, Mario Aureli, and Martin Slowik. The enhanced sanov theorem and propagation of chaos. *Stochastic Processes and their Applications*, 128(7):2228–2269, July 2018.
- [DG87] Donald A. Dawson and Jürgen Gärtner. Large deviations from the McKean-Vlasov limit for weakly interacting diffusions. *Stochastics*, 20(4):247–308, 1987.
- [DG89] D. A. Dawson and J. Gärtner. Large deviations, free energy functional and quasi-potential for a mean field model of interacting diffusions. *Mem. Amer. Math. Soc.*, 78(398):iv+94, 1989.
- [dH00] Frank den Hollander. *Large deviations*, volume 14 of *Fields Institute Monographs*. American Mathematical Society, Providence, RI, 2000.
- [DIRT15] François Delarue, James Inglis, Sylvain Rubenthaler, and Etienne Tanré. Global solvability of a networked integrate-and-fire model of McKean-Vlasov type. *Ann. Appl. Probab.*, 25(4):2096–2133, 2015.
- [DJ08] Persi Diaconis and Svante Janson. Graph limits and exchangeable random graphs. *Rend. Mat. Appl., VII. Ser.*, 28(1):33–61, 2008.
- [DL17] Susanne Ditlevsen and Eva Löcherbach. Multi-class oscillating systems of interacting neurons. *Stochastic Processes and their Applications*, 127(6):1840 – 1869, 2017.
- [DLO19] Aline Duarte, Eva Löcherbach, and Guilherme Ost. Stability, convergence to equilibrium and simulation of non-linear hawkes processes with memory kernels given by the sum of erlang kernels. *ESAIM: Probability and Statistics*, 23:770–796, 2019.
- [DM10a] Amir Dembo and Andrea Montanari. Gibbs measures and phase transitions on sparse random graphs. *Braz. J. Probab. Stat.*, 24(2):137–211, 2010.
- [DM10b] Amir Dembo and Andrea Montanari. Ising models on locally tree-like graphs. *Ann. Appl. Probab.*, 20(2):565–592, 2010.
- [DMPV86] Anna De Masi, E. Presutti, and Maria Vares. Escape from the unstable equilibrium in a random process with infinitely many interacting particles. *Journal of Statistical Physics*, 44:645–696, 08 1986.

- [DPGR14] Paolo Dai Pra, Giambattista Giacomini, and Daniele Regoli. Noise-induced periodicity: Some stochastic models for complex biological systems. In *Mathematical Models and Methods for Planet Earth*, pages 25–35. Springer, 2014.
- [DS89] Jean-Dominique Deuschel and Daniel W. Stroock. *Large deviations*, volume 137 of *Pure and Applied Mathematics*. Academic Press, Inc., Boston, MA, 1989.
- [DT21] François Delarue and Alvin Tse. Uniform in time weak propagation of chaos on the torus. *ArXiv e-print 2104.14973*, 2021.
- [Dur07] Rick Durrett. *Random graph dynamics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2007.
- [DZ98] Amir Dembo and Ofer Zeitouni. *Large deviations techniques and applications*, volume 38 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, second edition, 1998.
- [Ebe16] Andreas Eberle. Reflection couplings and contraction rates for diffusions. *Probability Theory and Related Fields*, 166, 12 2016.
- [EGZ16] Andreas Eberle, Arnaud Guillin, and Raphael Zimmer. Quantitative Harris type theorems for diffusions and McKean-Vlasov processes. *ArXiv e-print 1606.06012*, 06 2016.
- [EK86] G. B. Ermentrout and N. Kopell. Parabolic bursting in an excitable system coupled with a slow oscillation. *SIAM J. Appl. Math.*, 46(2):233–253, 1986.
- [ELL22] Xavier Erny, Eva Löcherbach, and Dasha Loukianova. Mean field limits for interacting hawkes processes in a diffusive regime. *Bernoulli*, 28(1), February 2022.
- [EM93] G. Bard Ermentrout and J. Bryce McLeod. Existence and uniqueness of travelling waves for a neural network. *Proceedings of the Royal Society of Edinburgh: Section A Mathematics*, 123(3):461–478, 1993.
- [Fel41] Willy Feller. On the integral equation of renewal theory. *Ann. Math. Statist.*, 12(3):243–267, 09 1941.
- [Fen72] Neil Fenichel. Persistence and smoothness of invariant manifolds for flows. *Indiana Univ. Math. J.*, 21:193–226, 1972.
- [Fen79] Neil Fenichel. Geometric singular perturbation theory for ordinary differential equations. *Journal of Differential Equations*, 31:53–98, 1979.
- [FFG92] René Ferland, Xavier Fernique, and Gaston Giroux. Compactness of the fluctuations associated with some generalized nonlinear boltzmann equations. *Canadian Journal of Mathematics*, 44(6):1192–1205, 1992.
- [FGS17] Franco Flandoli, Benjamin Gess, and Michael Scheutzow. Synchronization by noise. *Probab. Theory Related Fields*, 168:511–556, 2017.
- [FGVG14] Bastien Fernandez, David Gérard-Varet, and Giambattista Giacomini. Landau damping in the kuramoto model. 10 2014.
- [FI15] O. Faugeras and J. Inglis. Stochastic neural field equations: a rigorous footing. *J. Math. Biol.*, 71(2):259–300, 2015.
- [Fit61] Richard FitzHugh. Impulses and physiological states in theoretical models of nerve membrane. *Biophysical Journal*, 1(6):445–466, 1961.
- [FK18] G. Faye and Z. Kilpatrick. Threshold of front propagation in neural fields: An interface dynamics approach. *SIAM Journal on Applied Mathematics*, 78(5):2575–2596, 2018.
- [FM97] Begonia Fernandez and Sylvie Méléard. A Hilbertian approach for fluctuations on the McKean-Vlasov model. *Stochastic Process. Appl.*, 71(1):33–53, 1997.
- [Föll88] Hans Föllmer. Random fields and diffusion processes. In *École d’Été de Probabilités de Saint-Flour XV–XVII, 1985–87*, volume 1362 of *Lecture Notes in Math.*, pages 101–203. Springer, Berlin, 1988.

- [FR01] Marie-Christine Firpo and Stefano Ruffo. Chaos suppression in the large size limit for long-range systems. *J. Phys. A*, 34(37):L511–L518, 2001.
- [Fri64] Avner Friedman. *Partial differential equations of parabolic type*. Prentice-Hall Inc., Englewood Cliffs, N.J., 1964.
- [Fun95] T. Funaki. The scaling limit for a stochastic PDE and the separation of phases. *Probab. Theory Related Fields*, 102(2):221–288, 1995.
- [GaOS99] Jordi Garcí a Ojalvo and José M. Sancho. *Noise in spatially extended systems*. Institute for Nonlinear Science. Springer-Verlag, New York, 1999.
- [Gär88] Jürgen Gärtner. On the McKean-Vlasov limit for interacting diffusions. *Math. Nachr.*, 137:197–248, 1988.
- [GCR12] Shamik Gupta, Alessandro Campa, and Stefano Ruffo. Overdamped dynamics of long-range systems on a one-dimensional lattice: Dominance of the mean-field mode and phase transition. *Physical Review E*, 86(6), December 2012.
- [GCR14] Shamik Gupta, Alessandro Campa, and Stefano Ruffo. Kuramoto model of synchronization: equilibrium and nonequilibrium aspects. *Journal of Statistical Mechanics: Theory and Experiment*, 2014(8):R08001, 2014.
- [GM21] Arnaud Guillin and Pierre Monmarché. Uniform long-time and propagation of chaos estimates for mean field kinetic particles in non-convex landscapes. *Journal of Statistical Physics*, 185, 11 2021.
- [GP15] Giambattista Giacomin and Christophe Poquet. Noise, interaction, nonlinear dynamics and the origin of rhythmic behaviors. *Braz. J. Probab. Stat.*, 29(2):460–493, 2015.
- [GPP12] Giambattista Giacomin, Khashayar Pakdaman, and Xavier Pellegrin. Global attractor and asymptotic dynamics in the Kuramoto model for coupled noisy phase oscillators. *Nonlinearity*, 25(5):1247, 2012.
- [GPPP12] G. Giacomin, K. Pakdaman, X. Pellegrin, and C. Poquet. Transitions in active rotator systems: Invariant hyperbolic manifold approach. *SIAM Journal on Mathematical Analysis*, 44(6):4165–4194, 2012.
- [GPR12] Shamik Gupta, Max Potters, and Stefano Ruffo. One-dimensional lattice of oscillators coupled through power-law interactions: Continuum limit and dynamics of spatial Fourier modes. *Phys. Rev. E*, 85:066201, Jun 2012.
- [GPS18] Giambattista Giacomin, Christophe Poquet, and Assaf Shapira. Small noise and long time phase diffusion in stochastic limit cycle oscillators. *Journal of Differential Equations*, 264(2):1019–1049, January 2018.
- [GT12] Alexandre Genadot and Michèle Thieullen. Averaging for a fully coupled piecewise-deterministic Markov process in infinite dimensions. *Adv. in Appl. Probab.*, 44(3):749–773, 2012.
- [GTNE10] Denis S. Goldobin, Jun-nosuke Teramae, Hiroya Nakao, and G. Bard Ermentrout. Dynamics of limit-cycle oscillators subject to general noise. *Phys. Rev. Lett.*, 105:154101, Oct 2010.
- [Guc75] J. Guckenheimer. Isochrons and phaseless sets. *Journal of Mathematical Biology*, 1(3):259–273, September 1975.
- [GW79] J. Guckenheimer and R. F. Williams. Structural stability of Lorenz attractors. *Publications Mathématiques de l’IHÉS*, 50:59–72, 1979.
- [GW22] Priscilla E. Greenwood and Lawrence M. Ward. Phase offset determines alpha modulation of gamma phase coherence and hence signal transmission. *Biosystems*, 219:104729, September 2022.
- [HL93] Jack K. Hale and Sjoerd M. Verduyn Lunel. *Introduction to Functional Differential Equations*. Springer New York, 1993.

- [HM86] Masuyuki Hitsuda and Itaru Mitoma. Tightness problem and stochastic evolution equation arising from fluctuation phenomena for interacting diffusions. *J. Multivariate Anal.*, 19(2):311–328, 1986.
- [HM14] Maxime Hauray and Stéphane Mischler. On Kac’s chaos and related problems. *J. Funct. Anal.*, 266(10):6055–6157, 2014.
- [HO74] Alan G. Hawkes and David Oakes. A cluster process representation of a self-exciting process. *J. Appl. Probability*, 11:493–503, 1974.
- [HPS77] M. W. Hirsch, C. C. Pugh, and M. Shub. *Invariant manifolds*. Lecture Notes in Mathematics, Vol. 583. Springer-Verlag, Berlin, 1977.
- [HS21] Sophie Heesen and Wilhelm Stannat. Fluctuation limits for mean-field interacting nonlinear hawkes processes. *Stochastic Processes and their Applications*, 139:280–297, September 2021.
- [IM16] J. Inglis and J. MacLaurin. A general framework for stochastic traveling waves and patterns, with application to neural field equations. *SIAM Journal on Applied Dynamical Systems*, 15(1):195–234, 2016.
- [Izh07] Eugene M. Izhikevich. *Dynamical systems in neuroscience: the geometry of excitability and bursting*. Computational Neuroscience. MIT Press, Cambridge, MA, 2007.
- [Jan13] Svante Janson. *Graphons, cut norm and distance, couplings and rearrangements*, volume 4 of *New York Journal of Mathematics*. NYJM Monographs. State University of New York, University at Albany, Albany, NY, 2013.
- [JM86] A. Joffe and M. Métivier. Weak convergence of sequences of semimartingales with applications to multitype branching processes. *Adv. in Appl. Probab.*, 18(1):20–65, 1986.
- [JM98] B. Jourdain and S. Méléard. Propagation of chaos and fluctuations for a moderate model with smooth initial data. *Ann. Inst. H. Poincaré Probab. Statist.*, 34(6):727–766, 1998.
- [JPS21] Pierre-Emmanuel Jabin, David Poyato, and Juan Soler. Mean-field limit of non-exchangeable systems. *ArXiv e-print 2112.15406*, 2021.
- [JR13] Benjamin Jourdain and Julien Reygner. Propagation of chaos for rank-based interacting diffusions and long time behaviour of a scalar quasilinear parabolic equation. *Stochastic Partial Differential Equations: Analysis and Computations*, 1(3):455–506, 2013.
- [JTT18] Jean-François Jabir, Denis Talay, and Milica Tomašević. Mean-field limit of a particle approximation of the one-dimensional parabolic-parabolic keller-segel model without smoothing. *Electronic Communications in Probability*, 23(none), January 2018.
- [Kac56] Marc Kac. Foundations of kinetic theory. In *Berkeley Symp. on Math. Statist. and Prob.*, pages 171–197. University of California Press, 1956.
- [Kat95] Tosio Kato. *Perturbation theory for linear operators*. Classics in Mathematics. Springer-Verlag, Berlin, 1995. Reprint of the 1980 edition.
- [KE13] Zachary P. Kilpatrick and Bard Ermentrout. Wandering bumps in stochastic neural fields. *SIAM Journal on Applied Dynamical Systems*, 12(1):61–94, January 2013.
- [Kha12] Rafail Khasminskii. Stability of stochastic differential equations. In *Stochastic Stability of Differential Equations*, pages 145–176. Springer Berlin Heidelberg, 2012.
- [KLN⁺95] S. Yu. Kourtchatov, V. V. Likhanskii, A. P. Napartovich, F. T. Arecchi, and A. Lapucci. Theory of phase locking of globally coupled laser arrays. *Phys. Rev. A*, 52:4089–4094, Nov 1995.

- [KMT75] J. Komlós, P. Major, and G. Tusnády. An approximation of partial sums of independent RV's, and the sample DF. i. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 32(1-2):111–131, 1975.
- [KS14] J. Krüger and W. Stannat. Front propagation in stochastic neural fields: A rigorous mathematical framework. *SIAM Journal on Applied Dynamical Systems*, 13(3):1293–1310, 2014.
- [Kur75] Yoshiki Kuramoto. Self-entrainment of a population of coupled non-linear oscillators. In *International Symposium on Mathematical Problems in Theoretical Physics (Kyoto Univ., Kyoto, 1975)*, pages 420–422. Lecture Notes in Phys., 39. Springer, Berlin, 1975.
- [KVM17] D. Kaliuzhnyi-Verbovetskyi and G. Medvedev. The semilinear heat equation on sparse random graphs. *SIAM Journal on Mathematical Analysis*, 49(2):1333–1355, 2017.
- [KVM18] D. Kaliuzhnyi-Verbovetskyi and G. Medvedev. The mean field equation for the kuramoto model on graph sequences with non-lipschitz limit. *SIAM Journal on Mathematical Analysis*, 50(3):2441–2465, 2018.
- [KX99] Thomas G. Kurtz and Jie Xiong. Particle representations for a class of nonlinear SPDEs. *Stochastic Process. Appl.*, 83(1):103–126, 1999.
- [KX04] Thomas G. Kurtz and Jie Xiong. A stochastic evolution equation arising from the fluctuations of a class of interacting particle systems. *Commun. Math. Sci.*, 2(3):325–358, 2004.
- [KX22] Christian Kuehn and Chuang Xu. Vlasov equations on digraph measures. *Journal of Differential Equations*, 339:261–349, December 2022.
- [KYW⁺10] Caroline H. Ko, Yujiro R. Yamada, David K. Welsh, Ethan D. Buhr, Andrew C. Liu, Eric E. Zhang, Martin R. Ralph, Steve A. Kay, Daniel B. Forger, and Joseph S. Takahashi. Emergence of noise-induced oscillations in the central circadian pacemaker. *PLOS Biology*, 8(10):1–19, 10 2010.
- [Lan16] E. Lang. A multiscale analysis of traveling waves in stochastic neural fields. *SIAM Journal on Applied Dynamical Systems*, 15(3):1581–1614, 2016.
- [Léo07] Christian Léonard. A large deviation approach to optimal transport. *ArXiv e-print 0710.1461*, October 2007.
- [LGONSG04] B. Lindner, J. Garcia-Ojalvo, A. Neiman, and L. Schimansky-Geier. Effects of noise in excitable systems. *Physics Reports*, 392(6):321 – 424, 2004.
- [LMS22] Matti Leimbach, Jonathan C. Mattingly, and Michael Scheutzow. Noise-induced strong stabilization. *ArXiv e-print 2009.10573*, 2022.
- [Lon74] Stig-Olof Londen. On the asymptotic behavior of the bounded solutions of a nonlinear volterra equation. *SIAM Journal on Mathematical Analysis*, 5(6):849–875, 1974.
- [LRW23] Daniel Lacker, Kavita Ramanan, and Ruoyu Wu. Local weak convergence for sparse networks of interacting processes. *The Annals of Applied Probability*, 33(2), April 2023.
- [LS06] László Lovász and Balázs Szegedy. Limits of dense graph sequences. *Journal of Combinatorial Theory, Series B*, 96(6):933 – 957, 2006.
- [LS16] Eva Lang and Wilhelm Stannat. L^2 -stability of traveling wave solutions to nonlocal evolution equations. *Journal of Differential Equations*, 261(8):4275–4297, Oct 2016.
- [LSG99] B. Lindner and L. Schimansky-Geier. Analytical approach to the stochastic fitzhugh-nagumo system and coherence resonance. *Phys. Rev. E*, 60:7270–7276, Dec 1999.
- [Luç12] Eric Luçon. *Coupled oscillators, disorder and synchronization*. PhD thesis, Université Pierre et Marie Curie, <http://tel.archives-ouvertes.fr/tel-00709998>, 2012.

- [Luo20] Liqun Luo. *Principles of Neurobiology*. Garland Science, September 2020.
- [Mal03] Florent Malrieu. Convergence to equilibrium for granular media equations and their Euler schemes. *Ann. Appl. Probab.*, 13(2):540–560, 2003.
- [McK67] H. P. McKean, Jr. Propagation of chaos for a class of non-linear parabolic equations. In *Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967)*, pages 41–57. Air Force Office Sci. Res., Arlington, Va., 1967.
- [Med14a] Georgi S. Medvedev. The nonlinear heat equation on dense graphs and graph limits. *SIAM J. Math. Anal.*, 46(4):2743–2766, 2014.
- [Med14b] Georgi S. Medvedev. The nonlinear heat equation on W -random graphs. *Arch. Ration. Mech. Anal.*, 212(3):781–803, 2014.
- [Mét87] Michel Métivier. Weak convergence of measure valued processes using Sobolev-embedding techniques. In *Stochastic partial differential equations and applications (Trento, 1985)*, volume 1236 of *Lecture Notes in Math.*, pages 172–183. Springer, Berlin, 1987.
- [Mit83] Itaru Mitoma. Tightness of probabilities on $C([0, 1]; \mathcal{S}')$ and $D([0, 1]; \mathcal{S}')$. *Ann. Probab.*, 11(4):989–999, 1983.
- [Mit85] Itaru Mitoma. An ∞ -dimensional inhomogeneous Langevin’s equation. *J. Funct. Anal.*, 61(3):342–359, 1985.
- [MMJ87] DC Michaels, EP Matyas, and J Jalife. Mechanisms of sinoatrial pacemaker synchronization: a new hypothesis. *Circulation Research*, 61(5):704–714, 1987.
- [MQT16] S. Mischler, C. Quiñinao, and J. Touboul. On a Kinetic Fitzhugh–Nagumo Model of Neuronal Network. *Communications in Mathematical Physics*, 342(3):1001–1042, 2016.
- [MR16] Carlo Marinelli and Michael Röckner. On the maximal inequalities of Burkholder, Davis and Gundy. *Expositiones Mathematicae*, 34(1):1–26, 2016.
- [MSSZ20] Sima Mehri, Michael Scheutzow, Wilhelm Stannat, and Bian Z. Zangeneh. Propagation of chaos for stochastic spatially structured neuronal networks with delay driven by jump diffusions. *Ann. Appl. Probab.*, 30(1):175–207, 02 2020.
- [MT19] Pierre Del Moral and Julian Tugaut. Uniform propagation of chaos and creation of chaos for a class of nonlinear diffusions. *Stochastic Analysis and Applications*, 37(6):909–935, June 2019.
- [Mül17] Patrick E. Müller. Path large deviations for interacting diffusions with local mean-field interactions in random environment. *Electron. J. Probab.*, 22:56 pp., 2017.
- [MVS⁺14] Yuri L. Maistrenko, Anna Vasylenko, Oleksandr Sudakov, Roman Levchenko, and Volodymyr L. Maistrenko. Cascades of multiheaded chimera states for coupled phase oscillators. *International Journal of Bifurcation and Chaos*, 24(08):1440014, August 2014.
- [Nak16] Hiroya Nakao. Phase reduction approach to synchronisation of nonlinear oscillators. *Contemporary Physics*, 57(2):188–214, 2016.
- [NAY62] J. Nagumo, S. Arimoto, and S. Yoshizawa. An active pulse transmission line simulating nerve axon. *Proceedings of the IRE*, 50(10):2061–2070, Oct 1962.
- [Neu84] H. Neunzert. An introduction to the nonlinear Boltzmann-Vlasov equation. In *Kinetic Theories and the Boltzmann Equation*, pages 60–110. Springer Berlin Heidelberg, 1984.
- [New14] Jay M. Newby. Spontaneous excitability in the morris–lecar model with ion channel noise. *SIAM Journal on Applied Dynamical Systems*, 13(4):1756–1791, January 2014.

- [nT20] Cristobal Qui ninao and Jonathan D. Touboul. Clamping and Synchronization in the Strongly Coupled FitzHugh–Nagumo Model. *SIAM Journal on Applied Dynamical Systems*, 19(2):788–827, January 2020.
- [Nus79] Roger D. Nussbaum. Periodic solutions of nonlinear autonomous functional differential equations. In *Functional Differential Equations and Approximation of Fixed Points*, pages 283–325. Springer Berlin Heidelberg, 1979.
- [OBH08] Srdjan Ostojic, Nicolas Brunel, and Vincent Hakim. Synchronization properties of networks of electrically coupled neurons in the presence of noise and heterogeneities. *Journal of Computational Neuroscience*, 26(3):369–392, 2008.
- [Oel84] Karl Oelschläger. A martingale approach to the law of large numbers for weakly interacting stochastic processes. *Ann. Probab.*, 12(2):458–479, 1984.
- [Oel87] Karl Oelschläger. A fluctuation theorem for moderately interacting diffusion processes. *Probab. Theory Related Fields*, 74(4):591–616, 1987.
- [OMHS11] Iryna Omelchenko, Yuri Maistrenko, Philipp Hövel, and Eckehard Schöll. Loss of coherence in dynamical networks: Spatial chaos and chimera states. *Phys. Rev. Lett.*, 106:234102, Jun 2011.
- [OR19] Roberto I. Oliveira and Guilherme H. Reis. Interacting diffusions on random graphs with diverging average degrees: Hydrodynamics and large deviations. *Journal of Statistical Physics*, 176(5):1057–1087, July 2019.
- [ORS20] Roberto I. Oliveira, Guilherme H. Reis, and Lucas M. Stolerman. Interacting diffusions on sparse graphs: hydrodynamics from local weak limits. *Electronic Journal of Probability*, 25:1–35, January 2020.
- [Paz83] A. Pazy. *Semigroups of linear operators and applications to partial differential equations*, volume 44 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1983.
- [Pea81] Paul A. Pearce. Mean-field bounds on the magnetization for ferromagnetic spin models. *J. Statist. Phys.*, 25(2):309–320, 1981.
- [Pis12] Gilles Pisier. Grothendieck’s theorem, past and present. *Bull. Amer. Math. Soc. (N.S.)*, 49(2):237–323, 2012.
- [Poq14] Christophe Poquet. Phase reduction in the noise induced escape problem for systems close to reversibility. *Stochastic Processes and their Applications*, 124(10):3312 – 3341, 2014.
- [Pro23] Adrien Prodhomme. Strong gaussian approximation of metastable density-dependent markov chains on large time scales. *Stochastic Processes and their Applications*, 160:218–264, June 2023.
- [PRS22] P. Pfaffelhuber, S. Rotter, and J. Stiefel. Mean-field limits for non-linear hawkes processes with excitation and inhibition. *Stochastic Processes and their Applications*, 153:57–78, November 2022.
- [PTW10] K. Pakdaman, M. Thieullen, and G. Wainrib. Fluid limit theorems for stochastic hybrid systems with application to neuron models. *Adv. in Appl. Probab.*, 42(3):761–794, 2010.
- [RAB⁺14] J. Rankin, D. Avitabile, J. Baladron, G. Faye, and D. Lloyd. Continuation of localized coherent structures in nonlocal neural field equations. *SIAM Journal on Scientific Computing*, 36(1):B70–B93, 2014.
- [RBRTM13] Patricia Reynaud-Bouret, Vincent Rivoirard, and Christine Tuleau-Malot. Inference of functional connectivity in neurosciences via hawkes processes. *2013 IEEE Global Conference on Signal and Information Processing*, pages 317–320, 2013.
- [RC86] Sylvie Roelly-Coppoletta. A criterion of convergence of measure-valued processes: application to measure branching processes. *Stochastics*, 17(1-2):43–65, 1986.
- [RDL19] Mads Bonde Raad, Susanne Ditlevsen, and Eva Löcherbach. Age Dependent Hawkes Process. *ArXiv e-print 1806.06370*, 2019.

- [Reb79] R. Rebolledo. La méthode des martingales appliquée à l'étude de la convergence en loi de processus. *Bull. Soc. Math. France Mém.*, (62):v+125 pp. (1980), 1979.
- [RL20] Mads Bonde Raad and Eva Löcherbach. Stability for Hawkes processes with inhibition. *Electronic Communications in Probability*, 25(none):1 – 9, 2020.
- [RsGGt00] C. Roc_soreanu, A. Georgescu, and N. Giurgi_teanu. *The FitzHugh-Nagumo model*, volume 10 of *Mathematical Modelling: Theory and Applications*. Kluwer Academic Publishers, Dordrecht, 2000. Bifurcation and dynamics.
- [RV21a] Philippe Robert and Gaëtan Vignoud. Averaging principles for markovian models of plasticity. *Journal of Statistical Physics*, 183(3), jun 2021.
- [RV21b] Philippe Robert and Gaetan Vignoud. Stochastic models of neural synaptic plasticity. *ArXiv e-print 2010.08195*, 2021.
- [Sak88] Hidetsugu Sakaguchi. Cooperative phenomena in coupled oscillator systems under external fields. *Progr. Theoret. Phys.*, 79(1):39–46, 1988.
- [Sch85] Michael Scheutzow. Noise can create periodic behavior and stabilize nonlinear diffusions. *Stochastic Process. Appl.*, 20(2):323–331, 1985.
- [Sch86] Michael Scheutzow. Periodic behavior of the stochastic Brusselator in the mean-field limit. *Probab. Theory Relat. Fields*, 72(3):425–462, 1986.
- [Sho97] R. E. Showalter. *Monotone operators in Banach space and nonlinear partial differential equations*, volume 49 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1997.
- [SK86a] Hidetsugu Sakaguchi and Yoshiki Kuramoto. A soluble active rotator model showing phase transitions via mutual entrainment. *Progr. Theoret. Phys.*, 76(3):576–581, 1986.
- [SK86b] Shigeru Shinomoto and Yoshiki Kuramoto. Phase transitions in active rotator systems. *Progress of Theoretical Physics*, 75(5):1105–1110, 1986.
- [SM88a] Steven H. Strogatz and Renato E. Mirollo. Collective synchronisation in lattices of nonlinear oscillators with randomness. *J. Phys. A*, 21(13):L699–L705, 1988.
- [SM88b] Steven H. Strogatz and Renato E. Mirollo. Phase-locking and critical phenomena in lattices of coupled nonlinear oscillators with random intrinsic frequencies. *Phys. D*, 31(2):143–168, 1988.
- [SSK88] Hidetsugu Sakaguchi, Shigeru Shinomoto, and Yoshiki Kuramoto. Phase transitions and their bifurcation analysis in a large population of active rotators with mean-field coupling. *Progr. Theoret. Phys.*, 79(3):600–607, 1988.
- [ST85] Tokuzo Shiga and Hiroshi Tanaka. Central limit theorem for a system of Markovian particles with mean field interactions. *Z. Wahrsch. Verw. Gebiete*, 69(3):439–459, 1985.
- [Sta13] Wilhelm Stannat. Stability of travelling waves in stochastic nagumo equations. *ArXiv e-print 1301.6378*, 2013.
- [Str00] Steven H. Strogatz. From Kuramoto to Crawford: exploring the onset of synchronization in populations of coupled oscillators. *Phys. D*, 143(1-4):1–20, 2000. Bifurcations, patterns and symmetry.
- [SV06] Daniel W. Stroock and S. R. Srinivasa Varadhan. *Multidimensional diffusion processes*. Classics in Mathematics. Springer-Verlag, Berlin, 2006. Reprint of the 1997 edition.
- [SV18] Michael Scheutzow and Isabell Vorkastner. Synchronization, lyapunov exponents and stable manifolds for random dynamical systems. In Andreas Eberle, Martin Grothaus, Walter Hoh, Moritz Kassmann, Wilhelm Stannat, and Gerald Trutnau, editors, *Stochastic Partial Differential Equations and Related Fields*, pages 359–366, Cham, 2018. Springer International Publishing.

- [SY13] George R. Sell and Yuncheng You. *Dynamics of evolutionary equations*, volume 143 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2013.
- [Szn84] Alain-Sol Sznitman. Nonlinear reflecting diffusion process, and the propagation of chaos and fluctuations associated. *J. Funct. Anal.*, 56(3):311–336, 1984.
- [Szn85] Alain-Sol Sznitman. A fluctuation result for nonlinear diffusions. In S. Albeverio, editor, *Infinite-dimensional analysis and stochastic processes*, volume 124 of *Research Notes in Mathematics*, pages 145–160. Pitman (Advanced Publishing Program), Boston, MA, 1985. Papers from the meeting held at the University of Bielefeld, Bielefeld, 1983.
- [Szn91] Alain-Sol 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.
- [Tao12] Terence Tao. *Topics in Random Matrix Theory*. American Mathematical Society, March 2012.
- [Tes12] Gerald Teschl. *Ordinary differential equations and dynamical systems*, volume 140 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012.
- [THF12] Jonathan Touboul, Geoffroy Hermann, and Olivier Faugeras. Noise-induced behaviors in neural mean field dynamics. *SIAM J. Appl. Dyn. Syst.*, 11(1):49–81, 2012.
- [THHI07] Tsz-Leung To, Michael A. Henson, Erik D. Herzog, and Francis J. Doyle III. A molecular model for intercellular synchronization in the mammalian circadian clock. *Biophysical Journal*, 92(11):3792 – 3803, 2007.
- [TNE09] Jun-nosuke Teramae, Hiroya Nakao, and G. Bard Ermentrout. Stochastic phase reduction for a general class of noisy limit cycle oscillators. *Phys. Rev. Lett.*, 102:194102, May 2009.
- [Tou14] Jonathan Touboul. Propagation of chaos in neural fields. *Ann. Appl. Probab.*, 24(3):1298–1328, 2014.
- [TPVE20] Jonathan D. Touboul, Charlotte Piette, Laurent Venance, and G. Bard Ermentrout. Noise-induced synchronization and antiresonance in interacting excitable systems: Applications to deep brain stimulation in parkinson’s disease. *Physical Review X*, 10(1), March 2020.
- [TT20] Milica Tomasevic and Denis Talay. A new McKean-Vlasov stochastic interpretation of the parabolic-parabolic Keller-Segel model: The one-dimensional case. *Bernoulli*, 26(2):1323–1353, 2020.
- [Tug13] Julian Tugaut. Convergence to the equilibria for self-stabilizing processes in double-well landscape. *The Annals of Probability*, 41(3A):1427 – 1460, 2013.
- [Ver18] Roman Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2018.
- [Wai10] Gilles Wainrib. *Randomness in neurons : a multiscale probabilistic analysis*. PhD thesis, École Polytechnique, 2010.
- [Wan10] Xiao-Jing Wang. Neurophysiological and computational principles of cortical rhythms in cognition. *Physiological Reviews*, 90(3):1195–1268, July 2010.
- [WC72] H R Wilson and J D Cowan. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical journal*, 12(1):1–24, 01 1972.
- [Wig13] Stephen Wiggins. *Normally hyperbolic invariant manifolds in dynamical systems*, volume 105 of *Applied Mathematical Sciences*. Springer, 2013.