



**HAL**  
open science

# Hawkes processes on random graphs in neuroscience : large population limit and long time stability

Zoé Agathe-Nerine

► **To cite this version:**

Zoé Agathe-Nerine. Hawkes processes on random graphs in neuroscience : large population limit and long time stability. General Mathematics [math.GM]. Université Paris Cité, 2023. English. NNT : 2023UNIP7028 . tel-04502131

**HAL Id: tel-04502131**

**<https://theses.hal.science/tel-04502131v1>**

Submitted on 13 Mar 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

Ecole doctorale de Science Mathématiques de Paris Centre  
(ED 386)

*Laboratoire de Mathématiques Appliquées de Paris 5 (MAP5)  
UMR 8145 CNRS*

---

# Hawkes processes on random graphs in neuroscience: large population limit and long time stability

---

Par Zoé AGATHE-NERINE

Thèse de doctorat de mathématiques appliquées  
Dirigée par Ellen SAADA et co-encadrée par Eric LUÇON

Soutenue publiquement le 27 novembre 2023

Devant un jury composé de :

Ellen SAADA	Directrice de recherche CNRS, Université Paris Cité	Directrice
Eric LUÇON	Maître de conférence, Université Paris Cité	Co-encadrant
Eva LÖCHERBACH	Professeure, Université Paris 1 Panthéon-Sorbonne	Présidente du jury
Wilhelm STANNAT	Professeur, Technische Universität Berlin	Rapporteur
Denis TALAY	Directeur de recherche INRIA, Sophia Antipolis	Rapporteur
Manon COSTA	Maîtresse de conférence, Université Paul Sabatier	Examinatrice
Giambattista GIACOMIN	Professeur, Université Paris Cité	Examinateur
Viet Chi TRAN	Professeur, Université Gustave Eiffel	Examinateur



Au hasard une délivrance,  
Au hasard l'étoile filante  
Et l'éternel ciel de ma tête  
S'ouvre plus large à son soleil,  
A l'éternité du hasard.

---

Paul ELUARD, *Au hasard*.



# Processus de Hawkes sur des graphes aléatoires en neurosciences : limite en grande population et stabilité en temps long

## Résumé

Dans cette thèse, nous nous concentrons sur un modèle de population de neurones, où l'activité neuronale de la population est représentée par un processus de Hawkes multivarié. L'activité de chaque neurone est décrite par un processus de comptage, où chaque saut du processus représente l'émission d'un potentiel d'action (également appelé spike) par le neurone. Une étiquette est attribuée à chaque neurone, liée à sa fonction ou à sa localisation spatiale. La connectivité entre les neurones est donnée par un graphe aléatoire éventuellement dilué et inhomogène, où la probabilité de présence de chaque arête dépend des étiquettes de ses sommets. Le taux d'excitation de chaque neurone dépend du passé des neurones connectés : il augmente en cas d'excitation par d'autres neurones, ou diminue en cas d'inhibition.

La motivation principale est d'étudier le comportement du système dynamique lorsque la taille de la population tend vers l'infini, lorsque le graphe d'interaction est fixé. Une question importante est de comprendre comment l'inhomogénéité spatiale (dans l'interaction) influence le comportement en temps long du système. Une autre question consiste à relier le modèle microscopique à la *Neural Field Equation* (NFE), qui modélise une dynamique neuronale à grande échelle avec des interactions non locales.

Après une introduction en français et une autre introduction plus détaillée en anglais, ce document contient trois chapitres. Le chapitre 3 introduit notre modèle principal, et décrit sa limite en grande population par des processus ponctuels de Poisson inhomogènes. Nous prouvons un résultat de propagation du chaos, valable sur des intervalles de temps finis. Nous nous concentrons aussi sur le comportement en temps long de la limite en grande population dans le cas linéaire. Les chapitres suivants explorent des cas où il est possible d'étudier le comportement en temps long du système microscopique, plus précisément jusqu'à des temps polynomiaux en la taille de la population. Dans le chapitre 4, la dynamique de la limite en grande population est attirée par une solution stationnaire unique de la NFE.

Nous prouvons la stabilité en temps long de la dynamique du système microscopique autour de cette solution. Dans le chapitre 5, la NFE admet un ensemble de solutions stationnaires stables qui forment une variété. Nous montrons la proximité en temps long entre le système microscopique et cette variété. En particulier, nous prouvons que sur une échelle de temps de l'ordre de la taille de la population, la dynamique du système peut être décrite par un mouvement brownien sur la variété des solutions stationnaires.

## Mots-clés

processus de Hawkes non linéaire multivarié, processus ponctuel, champ-moyen, système spatialement étendu, propagation du chaos, graphes aléatoires, convergence de graphe, équation de champ neuronal

# Hawkes processes on random graphs in neuroscience: large population limit and long time stability

## Abstract

In this thesis, we are interested in a model of a population of neurons represented by a multivariate Hawkes process. The activity of each neuron is described by a counting process, where each jump represents a spike of the neuron. A label is attributed to each neuron, linked to its function or to its spatial location. The connectivity between neurons is given by a random possibly diluted and inhomogeneous graph where the probability of presence of each edge depends on the labels of its vertices. The firing rate of each neuron depends on the history of the connected neurons: this rate increases in case of excitation by others neurons, or decreases in case of inhibition.

The main motivation of this work is to study the behavior of the dynamical system when the size of the population goes to infinity, for fixed realization of the interaction: we focus on the large population limit in a quenched setting. One main issue is the understanding of how the spatial inhomogeneity (in the interaction) influences the large time behavior of the system. Another issue is to link the microscopic model to the Neural Field Equation (NFE), which models large-scale neural dynamics with non-local interactions.

This document contains, after an introduction in French and one (more detailed) in English, three chapters. Chapter 3 sets out the formal framework of the main model and gives its large population limit, described by inhomogeneous Poisson point processes. We prove a propagation of chaos result, valid on finite time intervals. A focus is also made on the large time behavior of the large population limit in the linear case. The following chapters explore cases where one can study the long-time behavior (up to time horizons that are polynomial in the size of the population) of the microscopic system. In Chapter 4 the dynamics is attracted to a unique stationary solution of the NFE. We prove the long-time stability of the dynamics of the system around this solution. In Chapter 5 the neural field equation admits a manifold of stable stationary solutions. We show the long-time proximity of the microscopic system to this manifold. In particular we prove that the dynamics of the system can be illustrated by a Brownian motion on the



manifold, when the time scale is of order the size of the population.

## **Key words**

multivariate nonlinear Hawkes processes, point processes, mean-field systems, spatially extended systems, propagation of chaos, random graphs, graph convergence, neural field equation

# Remerciements

— *la partie où le mot **merci** apparaît vraiment beaucoup de fois, mais chaque occurrence est amplement méritée !*

Je tiens d’abord à remercier ma superbe direction de thèse, merci merci merci Ellen et Eric. Je considère que j’ai énormément de chance de vous avoir comme directrice et directeur de thèse. Merci pour ce très beau sujet, pour votre encadrement, pour votre implication, pour vos conseils, pour vos encouragements, et pour votre patience aussi... Merci Ellen pour toutes nos discussions et tes retours. Ce n’était pas ton sujet de prédilection au début mais tu as été essentielle pour que cette thèse aboutisse. Merci Eric, merci pour tout. Dans ce tout, merci de m’avoir fait confiance comme première doctorante, les prochain.e.s auront également beaucoup de chance de t’avoir.

I am very grateful to both my reviewers Wilhelm Stannat and Denis Talay for having accepted to read my manuscript. Thank you for your input and kind words about this work.

Je remercie également Manon Costa, Giacomini Giambattista, Viet Chi Tran et Eva Löcherbach d’avoir accepté de faire partie de mon jury. Chi, merci de m’avoir invitée à parler à Lisbonne, je garde de très bons souvenirs de cette conférence (et de la visite de la ville). Merci Eva pour m’avoir initiée aux mathématiques en neurosciences, ton cours de M2 m’a suivi pendant toute cette thèse. Merci également de m’avoir dirigée vers le MAP5 (quel bon conseil !) et d’avoir fait partie de mon comité de suivi. Ça a été toujours un plaisir de te revoir pour diverses conférences (et soutenances!).

Non présents dans mon jury (déjà nombreux), j’aimerais remercier Christophe Poquet et Julien Chevallier, vos travaux m’ont beaucoup inspirée !

J’en profite pour remercier mes nombreux profs de mathématiques (et autres matières) pour tout ce qu’ils et qu’elles m’ont transmis, ce n’est pas étonnant maintenant que je souhaite enseigner. Merci tout particulièrement à M. Demaret pour m’avoir initiée aux maths au collège, je repense toujours avec affection à ces premières démonstrations sur le tableau latéral de Cluny ; et merci à M. Juan pour la rigueur inculquée en prépa. Merci Sylvie Méléard pour ce M2 que j’ai beaucoup apprécié.

J’ai eu la chance de faire cette thèse dans un excellent labo de maths avec une superbe ambiance. Merci Anne pour ta bienveillance et ton accueil. Merci Marie-Hélène pour tout ce que tu gères. Merci à l’équipe des permanent.e.s pour les discussions sur la terrasse ; Marie et sa bonne humeur perpétuelle, Angèle et ses supers conseils (pour le tarot comme pour l’EN)... Merci Nathalie d’avoir fait partie de mon comité de suivi et pour toutes les discussions très intéressantes que nous

avons eues. Merci Georges, Lionel et Antoine de m'avoir confiance pour les TDs de MC1 (que j'ai beaucoup aimé et aime encore donner !) et autres enseignements. Merci aussi Antoine pour ta bonne humeur contagieuse dans la cuisine.

Merci à tous les "éphémères" de ce labo. Entre le tarot, le laser game, les séances ciné, les soirées jeux, les dégustations culinaires et plus encore... Merci pour cette ambiance de folie, on en oublie presque qu'on vient pour travailler !

A mes camarades du bureau 725, quel plaisir de partager ce bureau ensemble ! Merci à vous tous et toutes ; Anton pour avoir verdi et ambiencé notre bureau avec tes plantes et tes poiriers ; Rémi & Rémi pour vos explications de sports, les recommandations littéraires et vos blagues (et merci Rémi L pour tes supers conseils de resto et nos discussions de tout, t'es vraiment un mec chouette !), le bureau m'a paru un peu vide lorsque vous vous êtes exilés ; Pierre Louis mon dracénois surprise, j'ai hâte de rencontrer Maxine ; Herb je ne sais pas si je dois te remercier pour tes partages culinaires douteux, mais merci pour les fous rires ; Marie pour ton bon goût en décoration (tmtc) ; Loïc on retourne ensemble visiter le quartier et ses belles adresses quand tu veux, merci pour tes énigmes (enfin peut être pas toutes !) et tes délicieuses dégustations culinaires, tu restes le maître du tarot ; Eloi merci merci pour le génialissime TarotBot, merci d'être toujours présent ; Alexander et ta quête de l'ergonomie absolue (j'ai appris beaucoup !) ; Guillaume pour tes conseils lillois ; Antoine S et notre rush final commun ; Adélie ma nouvelle voisine super sympa.

Parce qu'il y a quand même d'autres bureaux, merci ; Sonia, j'aime voir ta QR-Calaveras tous les jours chez moi ; Mehdi merci pour tous tes tips qui m'ont facilité la vie ; Ivan et ton excellent fromage ; Charlie – bientôt sur les poteaux, désolée pour Astérix mais pas pour Claude ; Ariane pour les débriefings KL et plus encore ! Diala, Mariem, on a commencé nos thèses ensemble, et je suis heureuse de les finir ensemble également. Diala j'ai hâte de découvrir ton futur travail sur les processus de Hawkes et Mariem j'admire à l'avance tes futures photographies. Merci également à Yen, Adrien, Chabane, Charles, Elisa, Keanu, Laurent, Léna, Sisi, et plus récemment Clémence, Lucie, Beatriz, Ariel, Arwan, Yassine, Bianca ; profitez bien !

Je n'oublie pas celles et ceux qui ont déjà soutenu ou qui étaient de passage et ont contribué à cette belle ambiance : Antoine M (c'était un plaisir de décortiquer des films avec toi), Angie, Alessandro, Alexandre, Apolline (et tes jolis fantômes), Bernardin, Cécile, Florian (merci pour tes tips EN, et d'avoir inauguré la tradition du tarot, elle a de beaux jours devant elle !), Juliana, Kendall, Léonard (et tes escargots colorés), Lucia, Martin (je rigole toujours en pensant à ton installation d'ordi), Michel, Ousmane, Safa, Sergio, Sinda, Thaïs, Thibaut.

Pendant cette thèse, j'ai rencontré également en conférences ou séminaires nombre de doctorant.e.s. Merci pour ces chouettes moments passés ensembles ; Claudia, Emilie (dès le M2), Emeline, Branda, Nhi, Nadia, Pierre, Perrine, Thomas, Rémi, Rita, Stefano, Zach, Louise...

Je garde de très bons souvenirs de ma première réelle expérience de recherche auprès de Michel Benaïm à Neuchâtel. Merci encore à Aïsha, Ana, Carsten, David, Katie, Luc, Lucas, Sanaz, Thiebout, Tom, Valentin et Edouard pour m'avoir si bien intégrée à leur monde et pour m'avoir montré en premier qu'être doctorante,

ça pouvait être fun !

J'aimerais remercier tout particulièrement mes piliers normaliens Laetitia et Michel, depuis la première année à Cachan jusqu'à la thèse (et après !), quel plaisir de se retrouver en conf et ailleurs ! Merci de votre présence et de votre soutien.

J'ai rencontré tellement de chouettes personnes dans cette école, je ne vais pas être exhaustive mais merci à Balthazar, Clément, Diane, Etienne, Guillaume, Maud, Mathilde, Pierre, les Rémi, Thomas et bien d'autres pour ces belles années. C'est là aussi que j'ai rencontré Jordy (& Marine) et Steven (puis Grace & Louis), c'est toujours un plaisir de se retrouver.

Pendant mon M2, je jouais déjà au tarot avec une équipe au top, merci Audrey, Gaël, Loane, Vianney et Laura pour cette belle année passée ensemble.

J'ai de la chance d'avoir reçu énormément de soutien, parfois de très très loin. Merci Julie, Nathalie & Eric, Jacques & Annick d'être toujours présent.e.s même à l'autre bout du monde. Julie, j'ai hâte de voyager de nouveau ensemble. Nathalie je connais plein de nouveaux jeux sympas à te faire découvrir ! Jacques, merci pour tout ce que j'ai appris au CSANC, l'escrime n'en fait qu'une petite partie.

Merci à la bande, Violaine & Seb, Anaëlle & Pierre, Marguerite & Valentin et Mathias & Priscille de m'avoir si bien intégrée. Passer du temps ensemble est vraiment un plaisir. Merci les calédo-parisiennes pour tous les repas à Paris (et les tisanes), merci les calédo-lillois.e.s pour nos (nombreuses) excursions et journées jeux ensemble !

Merci à mes amies calédoniennes Laura, Marie, Morgane et Perrine, avec qui l'expression "tout se dire" est si vraie. C'est une telle force de se savoir aussi écoutée auprès de vous. Merci pour votre soutien, nos discussions et nos fous rires.

Merci à ma famille phénoménale, mes grands-parents, mes cousines et mes cousins, mes tantes, mes oncles et mes tontons. Merci pour les cousinades ; repas, brunch, parcs, jeux, spectacles ou expos ; toutes les occasions sont bonnes pour passer des moments ensemble.

Merci aussi à ma belle famille calédonienne. Merci Sonia et Thierry de toujours m'accueillir comme votre fille ; et merci pour les cours de pêche à Boatpass ! Merci François et Rémi de me faire autant rigoler.

Merci à mes parents que j'aime et que j'adore. Vous m'avez tout donné et tant appris, merci pour tout. Merci à mes soeurs, notre complicité est l'une des choses les plus importantes de ma vie. Merci d'être les meilleurs piliers de mon monde.

Et enfin mon David. Merci d'être là pour moi, toujours et encore. Tu es mon meilleur soutien tout en me poussant à toujours mieux, tu me fais rire tous les jours et ton sourire m'aide à tout traverser. Merci de rendre ma vie aussi belle.



# Sommaire

<b>Notations</b>	<b>xiii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Contexte biologique . . . . .	1
1.2 Comment modéliser l'activité neuronale? . . . . .	4
1.3 Le modèle général . . . . .	11
<b>2 Introduction</b>	<b>17</b>
<i>Sections 2.1 and 2.2 consists in an English translation of Sections 1.1 and 1.2. Section 2.3 develops the mathematical background and our main model. Section 2.4 contains a detailed summary of the results obtained for each chapter.</i>	
2.1 Biological context . . . . .	18
2.2 How to model neural activity? . . . . .	20
2.3 Our general model . . . . .	26
2.4 Chapter organisation and contents . . . . .	33
<b>3 Multivariate Hawkes processes on inhomogeneous random graphs</b>	<b>51</b>
<i>This chapter consists in a modified version of my article Agathe-Nerine (2022) (expanded with some computations), published in Stochastic Processes and their Applications, with minor wording and notation changes for harmonization purposes.</i>	
3.1 Introduction . . . . .	53
3.2 A system of $N$ interacting particles on a graph and its limit . . . . .	56
3.3 Convergence of the model in large population . . . . .	59
3.4 Large time behavior of the limit process in the linear case . . . . .	65
3.5 Applications . . . . .	70
3.6 Possible extensions . . . . .	76
3.7 Existence and uniqueness of the model and its limit (proofs) . . . . .	78
3.8 Convergence of the mean-field process (proofs) . . . . .	85
3.9 The empirical measure and the spatial profile (proofs) . . . . .	107
3.10 Behavior in large time limit - Linear case (proofs) . . . . .	111
3.A Useful results . . . . .	116
3.B About matrices, graphs and graphon . . . . .	119

<b>4</b>	<b>Long-term stability of interacting Hawkes processes on random graphs</b>	<b>123</b>
	<i>This chapter consists in a modified version of my article <i>Agathe-Nerine (2023a)</i> (expanded with some computations), published in <i>Electronic Journal of Probability</i>, with minor wording and notation changes for harmonization purposes.</i>	
4.1	Introduction . . . . .	124
4.2	Hypotheses and main results . . . . .	130
4.3	Asymptotic behavior of $(X_t)$ . . . . .	141
4.4	Large time behavior of $X_N(t)$ . . . . .	144
4.5	Noise perturbation (proofs) . . . . .	149
4.6	Drift term (proofs) . . . . .	155
4.7	About the finite time behavior (proofs) . . . . .	165
4.A	Auxiliary results . . . . .	168
<b>5</b>	<b>Stability of wandering bumps for Hawkes processes interacting on the circle</b>	<b>173</b>
	<i>This chapter consists in a modified version of my article <i>Agathe-Nerine (2023b)</i> (expanded with some computations), with minor wording and notation changes for harmonization purposes. It has been submitted.</i>	
5.1	Introduction . . . . .	174
5.2	Stability of wandering bumps for interacting Hawkes processes . . . . .	180
5.3	Stationary solutions (proofs) . . . . .	191
5.4	Long time behavior (proofs) . . . . .	198
5.5	Fluctuations on the manifold (proofs) . . . . .	208
5.A	On the stationary solutions to the Neural Field Equation . . . . .	221
5.B	Some computations . . . . .	224
	<b>Bibliographie</b>	<b>233</b>

# Notations

## Abbreviations

a.s.	almost surely
w.r.t.	with respect to
i.i.d.	independent and identically distributed
càdlàg	right continuous with left limits
e.g.	exempli gratia
i.e.	id est

## Acronyms

ODE	Ordinary Differential Equation
SDE	Stochastic Differential Equation
NFE	Neural Field Equation
LLN	Law of Large Numbers
CLT	Central Limit Theorem

## Sets

$\mathbb{N}$	the set of all natural numbers
$\mathbb{R}_+$	the set of all nonnegative real numbers

## Functional spaces

Let  $(E, \mathcal{A}, \mu)$  be a measured space.

$L^p(E, \mu)$	the Lebesgue space, with $p \geq 1$
$\ \cdot\ _{E, \mu, p}$	the norm in $L^p(E, \mu) : \ h\ _{E, \mu, p} = \left(\int_E h^p d\mu\right)^{\frac{1}{p}}$
$\ \cdot\ _\infty$	the infinity norm of a real-valued bounded function on $E$
$\ \cdot\ _L$	the Lipchitz seminorm of a real-valued function on $E$
$\mathbb{D}(E, \mathbb{N})$	the space of càdlàg functions defined on $E$ and taking values in $\mathbb{N}$



## Others

$\llbracket 1, N \rrbracket$	for $N \in \mathbb{N}$ , the set $\{1, \dots, N\}$
$X \sim Y$	the variable $X$ follows the distribution of the variable $Y$
$\mathcal{B}(p)$	the Bernoulli distribution with parameter $p \in (0, 1)$
$\mathcal{U}(A)$	the uniform distribution on the set $A$
$\lfloor \cdot \rfloor$	the floor function
$\lceil \cdot \rceil$	the ceiling function
$a \wedge b$	the minimum between $a$ and $b$
$C_{p_1, p_2}$	a positive constant that depends on parameters $(p_1, p_2)$
$\propto$	proportional to
$\equiv$	identical to

# Chapitre 1

## Introduction

Dans ce chapitre nous présentons les principales motivations de cette thèse. En particulier, nous faisons un bref rappel des modèles mathématiques issus des neurosciences et des notions mathématiques utilisées, avant de présenter notre principal modèle étudié. Une version anglaise des deux premières parties de ce chapitre se trouve au Chapitre 2. La dernière partie présente le modèle étudié durant cette thèse et rassemble les résultats principaux de chaque chapitre. Une version anglaise plus complète, présentant le contexte mathématique et un résumé très détaillé des résultats démontrés se trouve dans le Chapitre 2.

### Sommaire

---

1.1	Contexte biologique . . . . .	1
1.1.1	Activité neuronale . . . . .	1
1.1.2	Le b.a.-ba du neurone . . . . .	2
1.2	Comment modéliser l'activité neuronale? . . . . .	4
1.2.1	Systèmes dynamiques et modèles continus . . . . .	4
1.2.2	Processus ponctuel . . . . .	6
1.3	Le modèle général . . . . .	11
1.3.1	Chapitre 3 : Processus de Hawkes multivariés sur des graphes aléatoires inhomogènes . . . . .	12
1.3.2	Chapitre 4 : Stabilité en temps long des processus de Hawkes en interaction sur des graphes aléatoires . . . . .	13
1.3.3	Chapter 5 : Stabilité des <i>bumps</i> pour les processus de Hawkes en interaction sur le cercle . . . . .	15

---

## 1.1 Contexte biologique

### 1.1.1 Activité neuronale

Les neurones sont les principaux composants du système nerveux, en tant que cellules spécialisées dans la réception, l'intégration et la transmission d'informations. Si le cerveau humain compte en moyenne 86 milliards de neurones, il existe une

grande diversité dans le règne animal : les insectes ont environ un million de neurones, et l'un des plus petits cerveaux du monde (un ver microscopique) en contient environ 300.

Le comportement neuronal est très complexe, et l'étude des mécanismes à l'oeuvre est assez récente. **Galvani** découvre que la propagation nerveuse est liée à une activité électrique au 18<sup>ème</sup> siècle, lors d'expériences sur des grenouilles. Vers 1900, **Bernstein** découvre l'élément clé du codage de l'information, le potentiel d'action. **Golgi** et **Ramón y Cajal** décrivent anatomiquement les neurones pendant cette même période et observent des réseaux complexes avec une forte hétérogénéité. **Broca** s'intéresse à l'organisation spatiale du cortex. L'avancée des techniques d'observation telles que l'EEG (*électroencéphalographie*) et l'IRM (*imagerie par résonance magnétique*) — utilisées en science à partir des années 1950 — ont profondément contribué à la compréhension de la physiologie d'un neurone, et à l'établissement des premiers modèles mathématiques décrivant l'activité neuronale.

Cependant, notre compréhension du cerveau est loin d'être complète et est toujours à l'étude. Nous présentons ici un bref résumé des mécanismes en jeu.

### 1.1.2 Le b.a.-ba du neurone

L'émission des potentiels d'action (ou *spikes*) code le transfert d'information dans le cerveau. Succinctement, un neurone émet un potentiel d'action en fonction des informations qu'il reçoit des autres neurones, et en retour ce potentiel d'action peut influencer la dynamique des autres neurones avec lesquels il interagit. Ces processus sont complexes et impliquent de nombreux éléments chimiques et électriques.

Il existe plusieurs familles de neurones (sensoriels, moteurs... voir la classification de **Luo (2020)**), mais tous partagent certaines caractéristiques communes : leur membrane est polarisée, et chaque neurone peut se décomposer en plusieurs zones morphologiques distinctes explicitées sur la Figure 1.1, chacune ayant des fonctions spécifiques. Le neurone reçoit de l'information grâce à des connexions sur son arbre dendritique, composé de *dendrites* qui sont des extensions cellulaires avec de nombreuses branches. Le *soma*, corps du neurone, contient entre autres son *noyau*. L'*axone* est un long prolongement de la cellule qui transporte les signaux nerveux loin du soma, vers une ramification étendue qui sont les régions terminales. Également appelées *terminaux de l'axone*, elles permettent la communication avec de nombreuses cellules cibles vers d'autres neurones par l'intermédiaire de *synapses*. La transmission synaptique repose sur la libération de messagers chimiques.

Des canaux ioniques présents sur la membrane plasmique du neurone la rendent perméable : des échanges d'ions (sodium  $Na^+$ , potassium  $K^+$ , calcium  $Ca^{2+}$  et chlorure  $Cl^-$ ) se font entre l'intérieur et l'extérieur de la cellule. En l'absence d'interaction, les différences de concentration ionique de part et d'autre de la membrane induisent une différence de potentiel, appelée *le potentiel de repos*. Lors des synapses, les récepteurs de l'arbre dendritique perturbent la perméabilité des canaux ioniques de la membrane et entraînent ainsi une variation du potentiel

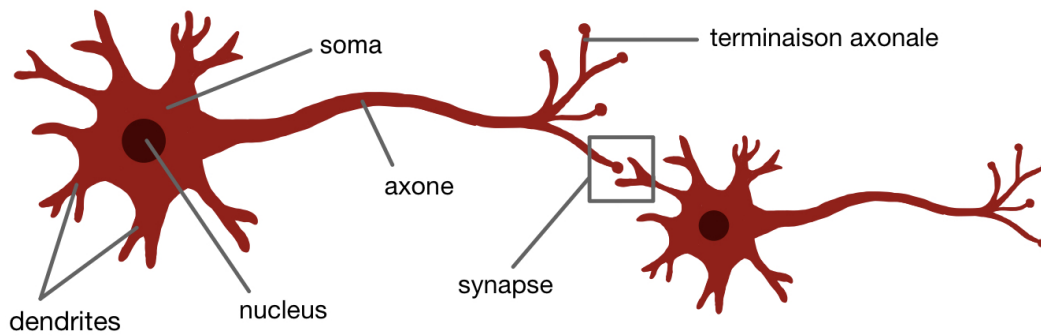


FIGURE 1.1: Schéma d'un neurone

membranaire, due notamment au courant transmembranaire de  $K^+$  et  $Na^+$ , ce que l'on peut voir sur la Figure 1.2. Si cette variation est suffisamment importante sur un court laps de temps, c'est-à-dire si elle atteint le potentiel seuil, le neurone génère un signal électrique appelé *potentiel d'action* ou *spike*. Le diagramme d'un spike typique est représenté sur la Figure 1.2. Il consiste en une dépolarisation (l'entrée d'ions sodium dans la cellule) et une repolarisation (la sortie d'ions potassium de la cellule), suivies d'une période réfractaire pendant laquelle le neurone ne peut pas déclencher un autre spike (le temps que les canaux sodiques se rétablissent), ou alors avec une difficulté inhabituelle car le seuil devient plus élevé.

Ce phénomène est bref dans le temps, de l'ordre de quelques millisecondes, et les potentiels d'action restent similaires pour un neurone donné. Une fois émis, le spike est propagé le long de l'axone jusqu'aux terminaisons et aux synapses, où le processus synaptique a lieu.

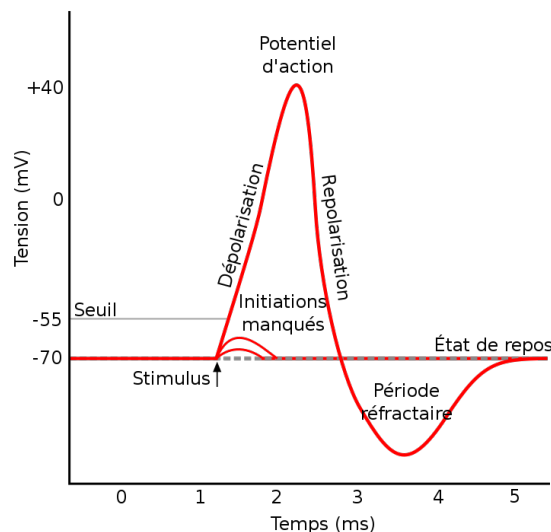


FIGURE 1.2: Schéma d'un potentiel d'action typique (from Wikimedia Commons ([https://commons.wikimedia.org/wiki/File:Potentiel\\_d%27action.svg](https://commons.wikimedia.org/wiki/File:Potentiel_d%27action.svg)), CC BY-SA 3.0)

## 1.2 Comment modéliser l'activité neuronale ?

Des modèles mathématiques simplifiés ont été introduits pour étudier l'activité neuronale. Leur diversité reflète le grand nombre de questions que les biologistes se posent, et l'échelle à laquelle se concentrer : sommes-nous plus intéressé.e.s par l'évolution du courant potentiel ? par la libération de neurotransmetteurs ? par les *trains de spikes* (la suite des instants de spikes) d'un ou de plusieurs neurones ? Nous présentons quelques-uns de ces modèles de neuroscience mathématique.

### 1.2.1 Systèmes dynamiques et modèles continus

Commençons par introduire les modèles dits continus, qui se concentrent sur la modélisation de l'évolution du potentiel membranaire par des équations différentielles ordinaires (EDO) ou des équations différentielles stochastiques (EDS). Ils ne sont pas l'objet de cette thèse, mais nous les présentons ici car ils ont été les premiers modèles historiques à aborder ces questions.

#### Un neurone ou une population finie de neurones (échelle microscopique)

Une manière de simplifier le processus biologique est de considérer des modèles de type "Intègre-et-tire". Prenons  $u$  le potentiel membranaire d'un neurone, évoluant selon une EDO comme  $\dot{u}_t = F(u_t)$ . Lorsque  $u$  atteint une valeur donnée, un seuil fixe  $U_{th}$ , on considère que le neurone émet un spike et nous réinitialisons  $u$  à sa valeur de repos  $U_{rest}$ . Ce modèle a été introduit par [Lapicque \(1907\)](#) et a été beaucoup étudié depuis, avec des variations intégrant de l'aléatoire (comme [Galves & Löcherbach \(2013\)](#); [Delarue et al. \(2015\)](#) et [Sacerdote & Giraud \(2013\)](#)). Notons que les derniers articles considèrent une population finie de neurones, où chaque neurone est modélisé par une EDO ou une EDS avec des interactions au sein de la population.

Comme il existe une certaine variabilité dans le seuil nécessaire pour émettre un spike, d'autres modèles évitent de fixer ce seuil. Le modèle d'[Hodgkin-Huxley](#) est probablement le plus célèbre concernant la modélisation déterministe d'un neurone. Basé sur l'étude d'un axone géant de calamar et à l'origine du prix Nobel de physiologie et médecine en 1963, [Hodgkin & Huxley \(1952\)](#) introduisent un système dynamique de 4 équations différentielles basées sur la conductance, représentant les interactions entre le potentiel membranaire ( $u_t$ ) et les variables décrivant les canaux ioniques de potassium ( $n_t$ ) et de sodium ( $m_t$ ), ainsi qu'avec un canal de fuite ( $h_t$ ). L'activité neuronale est ainsi modélisée par sa dynamique intrinsèque avec le système

$$\begin{cases} du_t &= I_t dt - F(u_t, n_t, m_t, h_t) dt \\ dn_t &= (\alpha_n(u_t)(1 - n_t) - \beta_n(u_t)n_t) dt \\ dm_t &= (\alpha_m(u_t)(1 - m_t) - \beta_m(u_t)m_t) dt \\ dh_t &= (\alpha_h(u_t)(1 - h_t) - \beta_h(u_t)h_t) dt, \end{cases}$$

où  $F(v, n, m, h) = g_n n^4 (v - v_n) + g_m m^3 h (v - v_m) + g_h (v - v_h)$  et les paramètres  $g_i, v_i, \alpha_i$  et  $\beta_i$  sont des conductances ou des taux associés à chaque canal  $i$ . Le comportement de ce système dépend fortement du stimulus extérieur  $I_t$ , et a beaucoup été étudié depuis, par exemple dans [Rinzel & Miller \(1980\)](#); [Aihara et al. \(1984\)](#), et également dans [Baladron et al. \(2012\)](#); [Bossy et al. \(2015\)](#); [Höpfner et al. \(2017\)](#) avec des versions stochastiques. Cependant, en raison de la complexité de ce modèle (à 4 équations pour un neurone), d'autres modèles ont été introduits pour en simplifier l'analyse.

Le modèle de FitzHugh-Nagumo, introduit dans [FitzHugh \(1961\)](#); [Nagumo et al. \(1962\)](#), simplifie celui de Hodgkin-Huxley en ne considérant plus que deux équations par neurone, l'une pour le potentiel membranaire ( $u_t$ ) et l'autre auxiliaire ( $C_t$ ) :

$$\begin{cases} du_t &= \left( u_t - \frac{u_t^3}{3} - C_t + I_t \right) dt \\ \tau dC_t &= (u_t + a - bC_t) dt, \end{cases}$$

avec  $a, b$  et  $\tau$  des paramètres et  $I$  le stimulus. Des versions stochastiques de ce modèle existent également, on peut les retrouver par exemple dans [Baladron et al. \(2012\)](#); [Luçon & Poquet \(2021b\)](#); [Colombani & Bris \(2022\)](#).

### Une grande population de neurones (échelle mésoscopique)

Si les modèles microscopiques sont nécessaires pour représenter l'activité neuronale à l'échelle d'un neurone, nous pouvons être intéressés par des modèles à plus grande échelle, portant sur une grande partie du cortex par exemple. Mais les neurones ne sont alors plus identiques : il est possible alors de leur attribuer des étiquettes. Mentionnons deux principaux exemples que nous utiliserons dans notre travail. Les neurones sont localisés dans le cerveau, et leur activité ou leur fonction peut être liée à cette position spatiale. Nous pouvons donc étiqueter chaque neurone d'une (grande) population de neurones en fonction de sa position. Un autre modèle très étudié est la *ring model*, qui se concentre plus spécifiquement sur le cortex visuel. Au lieu d'assigner à chaque neurone une position spatiale, les neurones sont différenciés en fonction de leur rôle : chaque neurone possède une *préférence d'orientation* qui stimule son activité (voir [Georgopoulos et al. \(1982\)](#); [Bosking et al. \(1997\)](#)). Nous ne considérons pas alors une population finie de neurones, mais une distribution de neurones en fonction de leur étiquette.

D'après les travaux de [Wilson & Cowan \(1972\)](#); [Amari \(1977\)](#) sur un modèle basé sur la conductance d'un réseau de neurones, [Bressloff \(2012\)](#) établit un modèle purement phénoménologique en introduisant le champ scalaire neuronal (*Neural Field Equations* que nous noterons NFE)

$$\frac{\partial u_t(x)}{\partial t} = -\alpha u_t(x) + \int_S w(x, y) f(u_t(y)) dy + I_t(x). \quad (1.2.1)$$

Dans ce modèle,  $u_t(x)$  représente le potentiel membranaire du neurone portant l'étiquette  $x$ , qui diminue à taux  $\alpha$ . Le noyau  $w$  représente la distribution spatiale des connexions synaptiques neuronales, la fonction  $f$  représente les effets de filtrage

du traitement synaptique et  $I_t(x)$  est le stimulus que le neurone reçoit au temps  $t$ . Cette équation modélise à grande échelle la dynamique de réseaux neuronaux structurés en espace. De nombreux phénomènes peuvent apparaître, comme des oscillations, des ondes (progressives), des *bumps*... ce qui explique le large intérêt que la NFE suscite : Ermentrout & McLeod (1993); Veltz & Faugeras (2010); Bressloff & Webber (2012); Kilpatrick & Ermentrout (2013); Bressloff (2014); Touboul (2014); Faugeras & Inglis (2015); Inglis & MacLaurin (2016); Lang & Stannat (2017); MacLaurin & Bressloff (2020). Ces structures apparaissent notamment grâce à l'invariance de la mesure de Lebesgue (présente dans l'intégration  $dy$  dans (1.2.1)).

### Quelques commentaires sur la stochasticité

En ce qui concerne l'aléa, si les modèles précédents sont déterministes, des versions stochastiques (déjà mentionnées) ont également été développées en ajoutant un bruit aux EDO. Plusieurs raisons expliquent ce choix. L'une des grandes questions des neurosciences mathématiques est de comprendre comment la complexité des phénomènes microscopiques (tels que ceux impliqués dans l'activité neuronale, à l'échelle du neurone) se traduit à l'échelle mésoscopique (pour une portion du cortex, par exemple) ou macroscopique. L'introduction de bruit permet de prendre en compte tous les micro-phénomènes que l'on ne comprend pas, ou qui sont trop difficiles à modéliser à l'échelle considérée.

De plus, il existe de la stochasticité biologique tout au long de la chaîne de réaction. Les arguments suivants sont issus de Pouzat (2020).

- A l'échelle d'une population de neurones, il peut y avoir une variabilité au sein de la population (sur le seuil qui permet les spikes, sur la libération de messagers chimiques...).
- A l'échelle du neurone, certains spikes sont spontanés (Fatt & Katz (1952)). En raison de l'état binaire des canaux ioniques (ouvert/fermé), ces changements d'état entraînent des fluctuations (Verveen & Derksen (1968)). De plus, certaines spikes ne sont pas correctement propagés le long de l'axone et ces échecs de propagation peuvent être aléatoires (Smith (1980)).
- A l'échelle de la synapse, la libération des neurotransmetteurs est très fluctuante : la quantité de messagers chimiques libérés par le neurone présynaptique peut varier, les récepteurs du neurone postsynaptique peuvent être plus ou moins occupés et la réponse des récepteurs se fait également par des portes binaires.

Dans la suite et plus particulièrement dans notre modèle, tout cet aléa est considéré à travers l'aléa associé à l'émission d'un spike, considéré alors comme un phénomène aléatoire.

### 1.2.2 Processus ponctuel

Il est possible de modéliser l'activité neuronale par des processus ponctuels — des processus stochastiques où chaque réalisation consiste en un ensemble de points

isolés. Comme le montre la Figure 1.2, un spike est un phénomène stéréotypé qui est bref dans le temps : l'information est codée dans la distribution inter-spikes. On peut simplifier la modélisation de l'activité neuronale d'un neurone en représentant un spike par un instant — le pic du potentiel d'action — et donc en considérant qu'un spike est un événement d'un processus ponctuel sur la droite réelle. La Figure 1.3 illustre la simplification de l'activité d'un neurone cortical par un processus ponctuel sur la droite réelle.

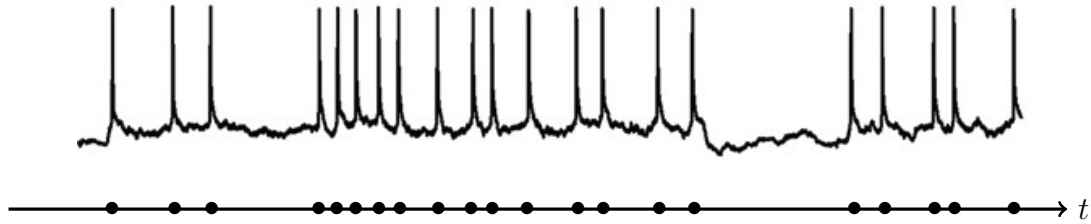


FIGURE 1.3: Réponse d'un neurone cortical à un stimulus (d'après Rossant *et al.* (2011)) et le processus ponctuel associé

Les trains de spike étant modélisés par des processus ponctuels, il convient maintenant de décrire comment ces spikes apparaissent. On utilise pour cela la notion *d'intensité stochastique*, qui caractérise la loi des arrivées des spikes.

### 1.2.2.1 Processus ponctuel et intensité stochastique

Rappelons d'abord quelques notions théoriques sur les processus ponctuels (Daley & Vere-Jones, 2006). Un processus ponctuel  $Z$  sur  $\mathbb{R}_+$  peut être représenté de manière équivalente par

- une suite strictement croissante de points  $(T_n)_{n \geq 1}$  sans point d'accumulation lorsque  $n \rightarrow \infty$ ,
- une mesure de comptage sur  $\mathbb{R}_+$  définie par  $Z = \sum_{n \geq 1} \delta_{T_n}$  — où  $\delta$  est la mesure de Dirac,
- un processus de comptage  $t \mapsto Z(t) = Z([0, t]) = \sum_{n \geq 1} \mathbf{1}_{t \leq T_n}$ ,

et on fixe  $Z_0 = 0$ . A chaque nouveau point  $T_n$ , on dit que le processus *saute*, et les  $(T_n)$  sont les sauts du processus. Pour toute fonction  $f$  on note

$$\int_{[0, t]} f(s) dZ(s) = \sum_{n \geq 1} f(T_n) \mathbf{1}_{T_n \leq t}.$$

Le processus ponctuel le plus commun est le processus de Poisson homogène de paramètre  $\lambda > 0$ , lorsque les temps d'attente entre deux événements consécutifs  $(T_{n+1} - T_n)_{n \geq 1}$  sont indépendants et suivent une loi exponentielle de paramètre  $\lambda$ . Dans ce cas, pour tout  $0 < a < b$ ,  $Z([a, b])$  suit une loi de Poisson de paramètre  $\lambda(b - a)$ , et la constante  $\lambda$  est appelée *l'intensité* du processus. Lorsque  $\lambda$  est une fonction déterministe  $\lambda : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ , on peut définir de manière similaire un



processus de Poisson inhomogène d'intensité  $\lambda$  lorsque pour tout  $0 < a < b$ ,  $Z([a, b])$  suit une loi de Poisson de paramètre  $\int_{[a,b]} \lambda(s)ds$ . Dans ces deux cas, les temps inter-arrivées sont indépendants, il n'y a pas d'auto-excitation du processus, ce que l'on cherche à avoir en neuroscience. Une manière d'introduire cette interaction est de considérer des *intensités stochastiques*, de sorte que la probabilité d'un nouveau saut dépende du passé du processus.

**Définition 1.2.1.** Soient  $(\mathcal{F}_t)$  une filtration sur  $\mathbb{R}_+$ ,  $Z$  un processus ponctuel  $(\mathcal{F}_t)$ -adapté sur  $\mathbb{R}_+$  tel que  $\mathbf{E}[Z(t)] < +\infty$  pour tout  $t \geq 0$ , et  $\lambda$  un processus aléatoire progressif par rapport à  $(\mathcal{F}_t)$ . On dit que  $Z$  admet la  $(\mathcal{F}_t)$ -intensité  $\lambda$  si  $(Z(t) - \int_0^t \lambda_s ds)$  est une  $(\mathcal{F}_t)$ -martingale.

Dans la majeure partie des cas, on prend  $\mathcal{F}_t = \sigma(Z(s), s \leq t)$ , l'histoire du processus. Il n'y a pas d'unicité pour une telle intensité stochastique (on peut considérer  $\lambda(t-)$ ), mais si on se restreint aux intensités prévisibles alors il y a unicité. Sauf mention du contraire dans la suite nous parlerons toujours de l'intensité prévisible. L'intérêt d'introduire l'intensité stochastique est que si  $\lambda$  est l'intensité de  $Z$ , alors informellement

$$\mathbf{P}(Z \text{ saute pendant le laps de temps } (t, t + dt) | \mathcal{F}_t) = \lambda(t)dt.$$

L'intensité est donc vue comme le *taux de saut instantané* du processus : plus elle est grande, plus le processus aura tendance à produire un nouveau saut. Comme ce taux peut dépendre du passé du processus, cela nous permet d'inclure de l'auto-excitation et des interactions dans notre modèle.

### 1.2.2.2 Elaguer des mesures aléatoires de Poisson

Pour construire des processus ponctuels admettant des intensités stochastiques, on peut utiliser des mesures aléatoires de Poisson.

**Définition 1.2.2.**  $\pi$  est une *mesure aléatoire de Poisson* sur  $\mathbb{R}_+ \times \mathbb{R}_+$  d'intensité  $dtdx$  si

1. pour tout  $C$  borélien de  $\mathbb{R}_+ \times \mathbb{R}_+$ ,  $\pi(C)$  suit une loi de Poisson qui a pour paramètre la mesure de Lebesgue de  $C$ ,
2. pour tous boréliens  $C_1, \dots, C_n$  mutuellement disjoints,  $(\pi(C_1), \dots, \pi(C_n))$  est une famille de variables aléatoires indépendantes.

Les processus ponctuels définis par leur intensité stochastique peuvent être construits en élaguant une mesure aléatoire de Poisson — on parle de *thinning*, méthode utilisée par [Ogata \(1988\)](#).

**Proposition 1.2.3.** Soient  $(\mathcal{F}_t)$  une filtration sur  $\mathbb{R}_+$ ,  $\pi$  une mesure aléatoire de Poisson sur  $\mathbb{R}_+ \times \mathbb{R}_+$  d'intensité  $dtdx$  et  $\lambda$  un processus positif progressif par rapport à  $(\mathcal{F}_t)$ . On définit le processus  $Z$  pour tout  $t \geq 0$  par

$$Z(t) = \int_0^t \int_0^\infty \mathbf{1}_{z \leq \lambda(s)} \pi(ds, dz).$$

Alors  $Z$  admet la  $(\mathcal{F}_t)$ -intensité  $\lambda$ .

Cette construction a été fréquemment employée, par exemple dans les travaux [Brémaud & Massoulié \(1996\)](#); [Chevallier \(2017\)](#); [Chevallier et al. \(2019\)](#). Cela signifie que pour construire un processus  $Z$  d'intensité stochastique  $\lambda$ , on peut se donner une réalisation d'une mesure aléatoire de Poisson  $\pi$  et regarder chaque point  $(s, z)$  de  $\pi$  : si  $z \leq \lambda(s)$  alors  $s$  devient un saut du processus  $Z$ . Nous donnons un exemple jouet d'une telle construction par élagage dans la [Figure 1.4](#) pour un processus auto-excitant  $Z(t)$  d'intensité stochastique  $\lambda(t) = 1 + \int_{]0,t[} \mathbf{1}_{[0,1]}(t-s)dZ(s)$  : dès que  $Z$  a un nouveau saut, l'intensité est augmentée de 1 pendant une unité de temps. Cette augmentation représente l'auto-excitation, dans notre exemple le troisième spike  $T_3$  se produit grâce à l'excitation résiduelle du deuxième spike  $T_2$ . Dans cet exemple,  $Z$  est un processus ponctuel particulier puisqu'il appartient à la classe des *processus de Hawkes linéaires*.

### 1.2.2.3 Processus de Hawkes

Les processus de Hawkes ont été introduits par [Hawkes \(1971\)](#) en tant que processus ponctuels auto-excitants. Un processus de Hawkes  $Z$  se définit de manière générale comme un processus ponctuel admettant une intensité stochastique de la forme

$$\lambda(t) = f \left( v(t) + \int_0^{t-} h(t-s)dZ(s) \right), \quad (1.2.2)$$

où  $f : \mathbb{R} \rightarrow \mathbb{R}_+$ ,  $v : \mathbb{R}_+ \rightarrow \mathbb{R}$  et  $h : \mathbb{R}_+ \rightarrow \mathbb{R}$ . La fonction  $h$  est le noyau de mémoire du processus, qui décrit comment un point du passé affecte le taux de saut instantané du processus. Les processus de Hawkes ont été introduits pour décrire les tremblements de terre et leurs répliques dans [Hawkes & Oakes \(1974\)](#); [Ogata \(1988\)](#), mais depuis les domaines d'application se sont diversifiés avec notamment la finance et les neurosciences. Le caractère bien posé de ces processus se montre classiquement en utilisant une itération de Picard ([Delattre et al. \(2016a\)](#); [Chevallier et al. \(2019\)](#)).

Dans le cadre spécifique des neurosciences (et plus généralement pour décrire un système de particules en interaction comme [Delattre et al. \(2016a\)](#)), on étudie une grande population de neurones. Aussi on considère un processus de Hawkes multivarié, soit une collection de processus de Hawkes évoluant conjointement. Chaque neurone est associé à un processus ponctuel (où chaque point représente un potentiel d'action du neurone) et l'intensité de chaque processus ponctuel prend en compte les spikes passés du neurone lui-même mais aussi ceux de la population. Ainsi, une modélisation de  $N$  neurones par un processus de Hawkes multivarié est un processus  $(Z_{N,1}, \dots, Z_{N,N})$  où chaque  $Z_{N,i}$  est un processus ponctuel d'intensité

$$\lambda_{N,i}(t) = f_i \left( v_i(t) + \frac{1}{N} \sum_{j=1}^N \int_0^{t-} w_{ij}^{(N)} h_{j \rightarrow i}(t-s)dZ_{N,j}(s) \right), \quad 1 \leq i \leq N, \quad (1.2.3)$$

les fonctions  $f$ ,  $v$  et  $h$  de (1.2.2) pouvant dépendre du neurone considéré, et avec  $w_{ij}^{(N)}$  l'influence du  $j^e$  neurone sur le  $i^e$ . Nous donnons un exemple pour deux neurones (notés  $A$  et  $B$ ) modélisé par deux processus de Hawkes  $Z_A$  et  $Z_B$  dans

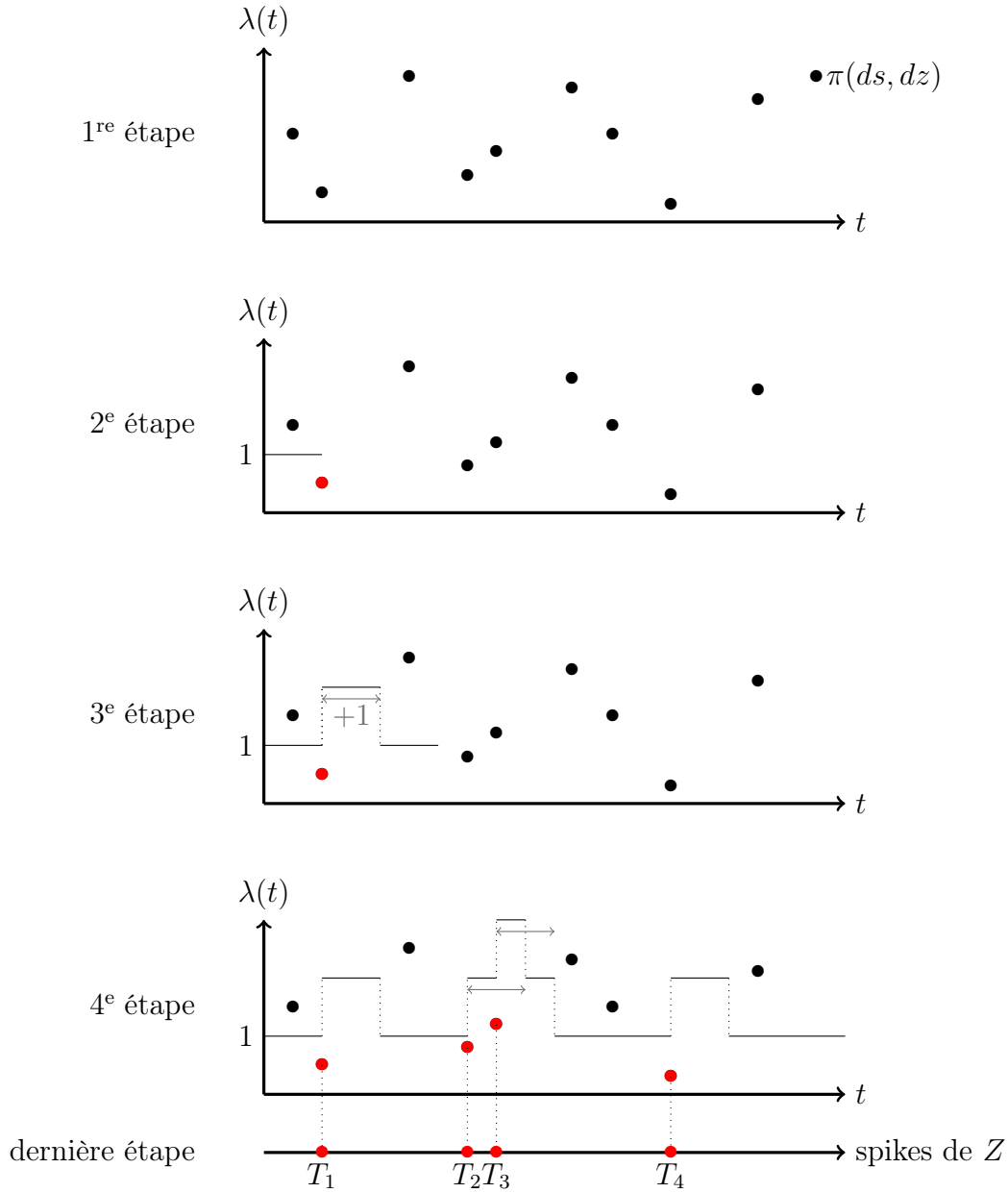


FIGURE 1.4: Méthode par *thinning* pour construire un processus  $Z$  d'intensité stochastique  $\lambda(t) = 1 + \int_{]0,t[} \mathbf{1}_{[0,1]}(t-s) dZ(s)$

1<sup>re</sup> étape : on tire une réalisation  $\pi$  de la mesure aléatoire de Poisson, certains de ces points projetés sur l'axe des abscisses formeront les sauts du processus  $Z$ ,  
 2<sup>e</sup> étape : sans saut, l'intensité reste à 1 jusqu'à ce qu'elle soit au dessus d'un point de  $\pi$ ,  
 3<sup>e</sup> étape : le point rouge (issu de la réalisation de  $\pi$ ) est sous l'intensité, donc devient un point du processus  $Z$  et affecte l'intensité du processus qui augmente de 1 pour une unité de temps,  
 4<sup>e</sup> étape : on itère, lorsque l'intensité se trouve au dessus d'un point de  $\pi$ , elle augmente de 1 pour une unité de temps car le processus ponctuel gagne un nouveau saut,  
 dernière étape : tous les points de  $\pi$  situés sous l'intensité  $\lambda(t)$  sont projetés sur l'axe des abscisses et forment la suite  $(T_n)$  des instants de saut du processus  $Z$ .

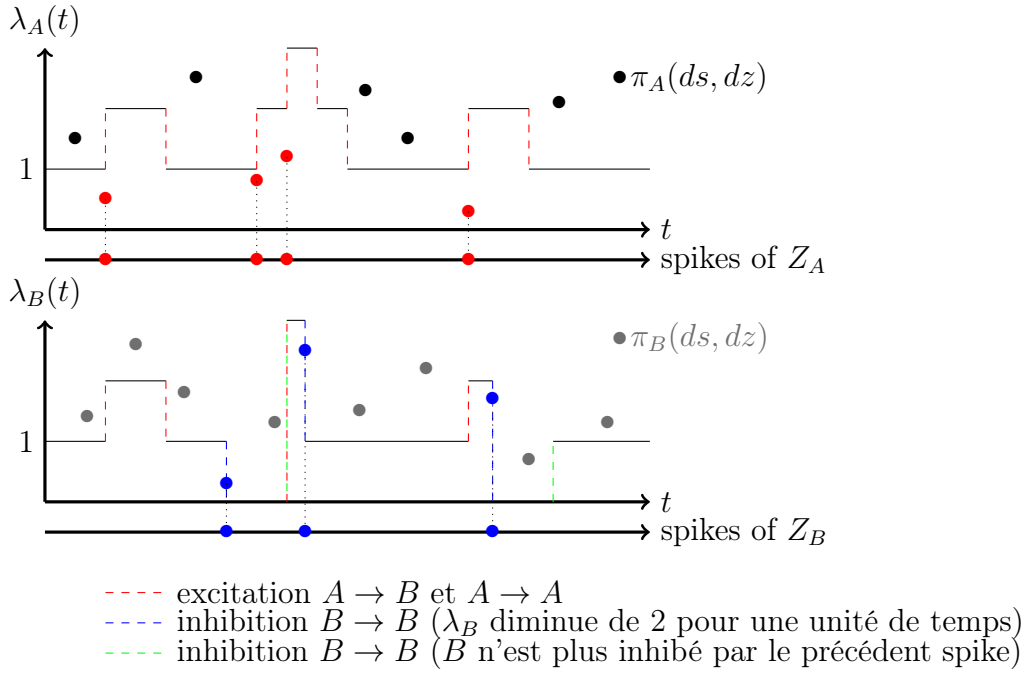


FIGURE 1.5: Deux neurones  $A$  et  $B$  en interaction modélisés par un processus de Hawkes, d'intensités  $\lambda_A$  et  $\lambda_B$  définies en (1.2.4) et (1.2.5)

On associe indépendamment à chaque neurone la réalisation d'une mesure aléatoire de Poisson ( $\pi_A$  et  $\pi_B$ ). Lorsque le neurone  $A$  émet un potentiel d'action (c'est-à-dire lorsqu'un point  $(s, z)$  de  $\pi_A$  est sous l'intensité  $\lambda_A(s)$ ), cela a pour effet d'augmenter les intensités des deux neurones pendant une unité de temps. Lorsque le neurone  $B$  émet un spike (c'est-à-dire lorsqu'un point  $(s, z)$  de  $\pi_B$  est sous l'intensité  $\lambda_B(s)$ ), le neurone s'inhibe lui-même et son intensité diminue de 2 pendant une unité de temps, sans descendre en dessous de 0.

la Figure 1.5, avec  $A$  un neurone auto-excitant et qui excite le neurone  $B$  et  $B$  un neurone auto-inhibant : on prend

$$\lambda_A(t) = 1 + \int_{]0,t[} \mathbf{1}_{[0,1]}(t-s) dZ_A(s), \quad (1.2.4)$$

$$\lambda_B(t) = \max \left\{ 0, 1 + \int_{]0,t[} \mathbf{1}_{[0,1]}(t-s) dZ_A(s) - 2 \int_{]0,t[} \mathbf{1}_{[0,1]}(t-s) dZ_B(s) \right\}. \quad (1.2.5)$$

La fonction  $f$  choisie comme  $f(u) = \max\{0, u\}$  garantit que, malgré la présence d'inhibition, l'intensité du processus ponctuel  $Z_B$  reste positive.

### 1.3 Le modèle général

On modélise l'activité d'une population de  $N$  neurones. L'activité du  $i^e$  neurone est décrite par un processus ponctuel :  $Z_{N,i}(t)$  compte le nombre de spikes que le  $i^e$  neurone a émis pendant l'intervalle  $[0, t]$ , chaque saut du processus correspond à un spike du neurone. Ce processus est caractérisé par son intensité stochastique notée  $\lambda_{N,i}(t)$  : plus cette intensité est importante, plus le neurone est enclin à émettre

un spike. Pour le choix de cette intensité, nous prenons en compte la dépendance de l'activité du neurone sur son interaction avec les autres neurones du système en considérant

$$\lambda_{N,i}(t) = \mu_t(x_i) + f \left( v_t(x_i) + \frac{1}{N} \sum_{j=1}^N w_{ij}^{(N)} \int_0^{t-} h(t-s) dZ_{N,j}(s) \right). \quad (1.3.1)$$

Ici,  $x_i \in I \subset \mathbb{R}^d$  représente une étiquette associée au  $i^e$  neurone, on considérera que  $I$  est le domaine spatial,  $f : \mathbb{R} \rightarrow \mathbb{R}_+$  représente l'intégration synaptique (où comment le modèle traduit le potentiel membranaire en probabilité de saut),  $\mu_t(x_i) \in \mathbb{R}_+$  représente l'activité spontanée et  $v_t(x_i)$  un stimulus passé du neurone au temps  $t$ ,  $h : \mathbb{R}_+ \rightarrow \mathbb{R}$  est la fonction de mémoire par laquelle le neurone se rappelle du passé et enfin le terme  $w_{ij}^{(N)}$  représente l'interaction entre les neurones, plus précisément comment le  $j^e$  neurone agit sur le  $i^e$ . Ce modèle est une continuation des articles [Delattre \*et al.\* \(2016a\)](#); [Chevallier \*et al.\* \(2019\)](#). L'originalité de cette thèse est de considérer que l'interaction entre les neurones n'est non pas déterministe comme dans les précédents articles cités, mais aléatoire : on suppose que l'interaction entre deux neurones d'étiquettes  $x_i$  et  $x_j$  se fait selon la réalisation d'un graphe aléatoire, dans le sens où la probabilité que l'interaction ait lieu dépend des étiquettes  $x_i$  et  $x_j$  à travers un noyau spatial  $W$ .

L'étude du comportement macroscopique de grandes populations de neurones en interaction est un problème mathématique de longue date. Une difficulté réside dans la compréhension de la dynamique des neurones à des échelles différentes : à une échelle microscopique (celle d'un neurone) et à une échelle méso-macroscopique, i.e. au niveau de la dynamique d'une région du cortex cérébral. Dans ce contexte, des questions importantes sont de comprendre l'influence du bruit sur la dynamique à ces différentes échelles ainsi que l'influence de la topologie du graphe d'interaction entre les neurones sur leur comportement en temps long.

### 1.3.1 Chapitre 3 : Processus de Hawkes multivariés sur des graphes aléatoires inhomogènes

Un premier résultat consiste à regarder la limite en grande population (lorsque  $N \rightarrow \infty$ ) du processus multivarié  $(Z_{N,1}(t), \dots, Z_{N,N}(t))_{t \geq 0}$  défini comme suit.

**Définition 1.3.1.** Soit  $(\pi_i(ds, dz))_{1 \leq i \leq N}$  une suite i.i.d. de mesures de Poisson sur  $\mathbb{R}_+ \times \mathbb{R}_+$  d'intensité  $dsdz$ . Le processus de comptage  $(Z_{N,1}(t), \dots, Z_{N,N}(t))_{t \geq 0}$  défini pour tout  $t \geq 0$  et  $1 \leq i \leq N$  par

$$Z_{N,i}(t) = \int_0^t \int_0^\infty \mathbf{1}_{\{z \leq \lambda_{N,i}(s)\}} \pi_i(ds, dz), \quad (1.3.2)$$

où  $\lambda_{N,i}$  est définie en (1.3.1), est appelé *un processus de Hawkes multivarié*.

Nous prouvons dans ce chapitre que la limite macroscopique du système (1.3.1) sur des intervalles de temps bornés s'exprime sous la forme d'une famille de processus

de Poisson inhomogènes spatialement étendus, d'intensités  $(\lambda_t(x))_{x \in I}$  solutions de l'équation de convolution suivante

$$\lambda_t(x) = \mu_t(x) + f \left( v_t(x) + \int_I W(x, y) \int_0^t h(t-s) \lambda_s(y) ds \nu(dy) \right), \quad (1.3.3)$$

où  $\nu$  représente la distribution spatiale macroscopique des neurones. Ce résultat s'obtient en utilisant, grâce à la famille  $(\pi_i)_{1 \leq i \leq N}$ , un couplage adéquat entre le processus  $(Z_{N,1}(t), \dots, Z_{N,N}(t))_{t \geq 0}$  et des processus de Poisson indépendants  $(\bar{Z}_1(t), \dots, \bar{Z}_N(t))_{t \geq 0}$  d'intensités  $(\lambda_t(x_i))_{1 \leq i \leq N}$ .

Un autre intérêt de ce chapitre est de comprendre le comportement en temps long de cette limite macroscopique (1.3.3) et d'observer les effets de l'inhomogénéité du graphe d'interaction. Dans le cadre champ moyen classique (où  $w_{ij}^{(N)} = 1$  pour tout  $i, j$ ), il n'y a plus de dépendance spatiale et (1.3.3) devient une équation de convolution scalaire  $\lambda_t = \mu_t + f(v_t + \int_0^t h(t-s) \lambda_s ds)$ . Lorsque de plus on choisit le cas linéaire, c'est-à-dire  $f(x) = x$ ,  $\mu_t = \mu$  et  $v_t = 0$  dans (1.3.3) (et donc avec  $h$  fonction à valeurs positives), le comportement de  $\lambda_t$  quand  $t \rightarrow \infty$  est bien connu : il y a un phénomène de transition de phase dans le comportement en temps long (Delattre *et al.*, 2016a, Théorèmes 10,11) selon le choix de la fonction de mémoire  $h$ . Lorsque  $\|h\|_1 = \int_0^\infty h(t) dt < 1$  (cas *sous-critique*),  $\lambda_t \xrightarrow[t \rightarrow \infty]{} \frac{\mu}{1 - \|h\|_1}$ , tandis que lorsque  $\|h\|_1 > 1$  (cas *sur-critique*),  $\lambda_t \xrightarrow[t \rightarrow \infty]{} \infty$ . Nous étendons cette transition de phase lorsque l'interaction est inhomogène : en notant  $r_\infty$  le rayon spectral de l'opérateur d'interaction  $T_W g(x) \mapsto \int_I W(x, y) g(y) \nu(dy)$ , dans le cas sous-critique qui devient  $\|h\|_1 r_\infty < 1$ , pour tout  $x \in I$  l'intensité macroscopique  $\lambda_t(x)$  converge vers  $\ell(x)$  l'unique solution de

$$\ell(x) = \mu(x) + \int_I W(x, y) \|h\|_1 \ell(y) \nu(dy); \quad (1.3.4)$$

et dans le cas sur-critique  $\|h\|_1 r_\infty > 1$ ,  $\|\lambda_t\|_2 \rightarrow \infty$ . Enfin, dans ce premier travail nous caractérisons également l'influence de l'interaction  $W$  dans le cas sous critique avec la proposition suivante.

**Proposition 1.3.2** (Proposition 3.4.5). *Supposons que la fonction  $\mu$  est constante. Alors la fonction  $\ell$  solution de (1.3.4) est constante si et seulement si le degré rentrant du graphe  $x \mapsto \int_I W(x, y) \nu(dy)$  est constant.*

### 1.3.2 Chapitre 4 : Stabilité en temps long des processus de Hawkes en interaction sur des graphes aléatoires

Les comportements asymptotiques ( $N \rightarrow \infty$  et  $t \rightarrow \infty$ ) ne sont pas interchangeables pour des modèles de champ moyen. Ce chapitre porte sur l'étude d'un cas où l'on obtient une convergence en temps long du système microscopique décrit en (1.3.1).

Un premier résultat du chapitre étend la condition de sous criticalité  $\|h\|_1 r_\infty < 1$  du chapitre précédent pour des cas non linéaires (c'est-à-dire où  $f$  n'est plus

l'identité). Dans ce cas, on demande comme condition suffisante de sous-criticalité

$$\|f'\|_\infty \|h\|_1 r_\infty < 1, \quad (1.3.5)$$

et alors l'intensité macroscopique  $\lambda_t$  converge vers une fonction spatiale solution de l'équation de convolution  $\ell(x) = \mu(x) + f\left(v(x) + \|h\|_1 \int_I W(x,y)\ell(y)\nu(dy)\right)$ , où  $\mu = \lim_{t \rightarrow \infty} \mu_t$  et  $v = \lim_{t \rightarrow \infty} v_t$ .

Un second résultat du chapitre consiste à se placer dans le cas sous critique (1.3.5) avec une fonction de mémoire spécifique de la forme  $h(t) = e^{-\alpha t}$  pour  $\alpha > 0$  fixé et regarder le comportement du système microscopique (1.3.2) sur des échelles de temps polynomiales en la taille de la population  $N$ . Le domaine spatial est ici réduit au segment  $I = [0, 1]$ , et les  $N$  neurones de la population sont disposés de manière régulière le long de  $I$ . L'interaction entre deux neurones  $i$  et  $j$  est fixée par le terme  $w_{ij}^{(N)} = \frac{\xi_{ij}}{\rho_N}$ , où  $\xi_{ij}$  suit une loi de Bernoulli de paramètre  $\rho_N W(x_i, x_j)$  avec  $\rho_N$  un facteur de dilution du graphe d'interaction. Le choix d'un noyau exponentiel permet de faire apparaître que la limite macroscopique du potentiel membranaire  $u_t$  (défini en lisant  $\lambda_t = \mu_t + f(u_t)$  dans (1.3.3)) lorsque  $v_t = \rho e^{-\alpha t}$  est solution de la NFE (1.2.1).

La quantité microscopique qui nous intéresse est le potentiel synaptique, défini par

$$X_{N,i}(t) = \sum_{j=1}^N \frac{w_{ij}^{(N)}}{N} \int_0^t e^{-\alpha(t-s)} dZ_{N,j}(s), \quad X_N(t)(x) := \sum_{i=1}^N \mathbf{1}_{\frac{i-1}{N} < x \leq \frac{i}{N}} X_{N,i}(t).$$

Il admet une limite en grande population donnée par  $X_t(x) = \int_I W(x,y) \int_0^t h(t-s)\lambda_s(y) ds dy$ . Le cas sous-critique (1.3.5) permet de garantir l'existence d'une unique solution stationnaire à la NFE (1.2.1) et la convergence de la limite déterministe  $X_t$  vers  $X_\infty := \|h\|_1 T_W \ell$ . Le résultat principal, de stabilité de  $X_t$  en temps polynomiaux, s'écrit sous la forme du théorème suivant :

**Théorème 1.3.3** (Théorème 4.2.12). *Fixons  $t_f > 0$  et  $m \geq 1$ . Alors sous des hypothèses de régularité des paramètres  $(f, \mu, v)$  et de dilution du graphe, dans le cas sous-critique (1.3.5),  $\mathbb{P}$  p.s. par rapport à la réalisation du graphe on a pour tout  $\varepsilon > 0$*

$$\mathbf{P} \left( \sup_{t \in [t_\varepsilon, (N\rho_N)^m t_f]} \|X_N(t) - X_\infty\|_2 \geq \varepsilon \right) \xrightarrow{N \rightarrow \infty} 0,$$

où  $t_\varepsilon > 0$  est indépendant de  $N$ .

Le temps  $t_\varepsilon$  correspond au temps nécessaire pour que la dynamique déterministe  $X_t$  soit suffisamment proche de la limite en temps long  $X_\infty$ . Les résultats du chapitre précédent impliquent que  $X_N$  est lui même proche de  $X_t$  avec grande probabilité sur des intervalles de temps finis. Enfin, pour lier  $X_N$  et  $X_\infty$ , on utilise un raisonnement par récurrence : sur un intervalle de temps fini, la dynamique déterministe prend le dessus sur le bruit donc si  $X_N$  et  $X_\infty$  sont proches, ils le restent

pour toute la durée de cet intervalle. En utilisant que le processus  $(X_{N,i}(t))_{1 \leq i \leq N, t \geq 0}$  est Markovien grâce au choix  $h(t) = e^{-\alpha t}$ , on obtient le résultat en étudiant le comportement infinitésimal du processus et en contrôlant le bruit provenant des sauts et des connections.

### 1.3.3 Chapitre 5 : Stabilité des *bumps* pour les processus de Hawkes en interaction sur le cercle

Ce dernier chapitre propose un cadre où l'on peut regarder ensemble les limites  $N \rightarrow \infty$  et  $t \rightarrow \infty$  lorsque la condition (1.3.5) n'est pas vérifiée. Les deux premiers chapitres de ma thèse étudient le comportement du modèle présenté en (1.3.2) dans un cadre sous-critique, dans lequel la limite obtenue est *unique*, car la NFE (1.2.1) admet alors une unique solution stationnaire. Une autre situation intéressante consiste à perturber le système pour obtenir des comportements asymptotiques plus variés, c'est à dire par exemple que la NFE (1.2.1) admette plusieurs solutions stationnaires. C'est l'objet de ce dernier travail.

On considère maintenant une interaction avec de l'inhibition, non plus aléatoire mais déterministe avec  $w_{ij}^{(N)} = 2\pi \cos(x_i - x_j)$  dans (1.3.1), et les neurones sont placés sur le cercle  $S = (-\pi, \pi]$  de manière régulière. La fonction  $f$  n'est plus générale puisqu'on prend une fonction sigmoïde de paramètre  $(\kappa, \varrho)$  :

$$f(u) = f_{\kappa, \varrho}(u) := (1 + e^{-(u-\varrho)/\kappa})^{-1}. \quad (1.3.6)$$

On reste dans le cadre exponentiel avec  $h(t) = e^{-t}$ , on prend de plus  $\mu_t = 0$  et  $v_t(x) = \rho(x)e^{-t}$  avec  $\rho$  une distribution initiale des potentiels membranaires. Ce choix spécifique des paramètres de (1.3.1) vient du *ring model*, qui modélise l'activité de neurones dans le cortex visuel à une échelle mésoscopique : chaque position  $x \in S$  représente une orientation préférentielle du neurone sur le cercle, les neurones ayant des orientations préférentielles similaires ont tendance à s'exciter mutuellement, alors que ceux ayant des orientations préférentielles opposées s'inhibent, d'où le choix du cosinus comme noyau d'interaction. Lorsque  $\kappa \rightarrow 0$ , la sigmoïde (1.3.6) converge vers la fonction de Heaviside  $H_\varrho(u) = \mathbf{1}_{u \geq \varrho}$  de seuil  $\varrho$ . Ainsi, lorsque  $\kappa$  est assez petit, lorsque son potentiel membranaire est élevé un neurone émet un spike à taux 1, et ce taux est proche de 0 si le potentiel n'est pas assez haut.

La quantité microscopique qui nous intéresse ici est

$$U_{N,i}(t) = \rho(x_i)e^{-t} + \frac{2\pi}{N} \sum_{j=1}^N \cos(x_i - x_j) \int_0^t e^{-(t-s)} dZ_{N,j}(s) =: \rho(x_i)e^{-t} + X_{N,i}(t), \quad (1.3.7)$$

et on définit similairement  $U_N(t)(x) := \sum_{i=1}^N U_{N,i}(t) \mathbf{1}_{x \in (x_{i-1}, x_i]}$  (avec  $x_0 = -\pi$  et  $x_i = \frac{2\pi}{N} (i - \frac{N}{2})$ ). Avec ce choix de paramètres, pour  $\kappa$  suffisamment petit la NFE (1.2.1) admet un cercle de solutions stationnaires stables

$$\mathcal{U} := (A \cos(\cdot + \phi))_{\phi \in S} =: (u_\phi)_{\phi \in S}, \quad (1.3.8)$$

où  $A$  est la plus grande solution de  $A = \int_S \cos(y) f_{\kappa, \varrho}(A \cos(y)) dy$ . La variété  $\mathcal{U}$  est ainsi indexée par le cercle  $S$ , chaque fonction  $u_\phi$  est un cosinus translaté



appelé *bump*. L'objet du chapitre est d'analyser le comportement en temps long du processus (1.3.7) le long de  $\mathcal{U}$  grâce au théorème suivant.

**Théorème 1.3.4** (Théorème 5.2.11). *Soit  $\tau_f > 0$ . Si  $\rho$  est suffisamment proche de  $\mathcal{U}$ , il existe  $\theta_0 \in S$ , un certain temps  $\tau_0(N) \propto \frac{\log(N)}{N}$  et un processus càdlàg  $(W_N(t))_{t \in (\tau_0(N), \tau_f)}$  convergeant faiblement dans  $\mathbb{D}([0, \tau_f], S)$  vers un mouvement brownien tels que pour tout  $\varepsilon > 0$ ,*

$$\lim_{N \rightarrow \infty} \mathbf{P} \left( \sup_{\tau \in (\tau_0(N), \tau_f)} \|U_N(N\tau) - u_{\theta_0 + \sigma W_N(\tau)}\|_2 \leq \varepsilon \right) = 1, \quad (1.3.9)$$

$$\text{où } \sigma := \left( 2\pi \int_S \sin^2(x) f(A \cos(x)) dx \right)^{\frac{1}{2}}.$$

L'idée principale derrière ce résultat est que, si l'on part d'une condition initiale proche de  $\mathcal{U}$ , la dynamique déterministe ramène le potentiel proche du cercle de solutions stationnaires  $\mathcal{U}$ , et le bruit du système induit un déphasage aléatoire sur le cercle. De plus, ce déphasage correctement renormalisé est quantifié par un mouvement brownien. Le temps  $N\tau_0(N)$  correspond au temps nécessaire pour que  $U_N(t)$  soit suffisamment proche de  $\mathcal{U}$ .

# Chapter 2

## Introduction

*Sections 2.1 and 2.2 consists in an English translation of Sections 1.1 and 1.2. Section 2.3 develops the mathematical background and our main model. Section 2.4 contains a detailed summary of the results obtained for each chapter.*

The goal of this chapter is to present the motivation and neuroscience context behind the model studied in this thesis. A brief and non-exhaustive history of mathematical models derived from neuroscience is presented, as well as some mathematical tools that will be useful in explaining our model. We then introduce our model and give the main results of the thesis.

### Contents

---

2.1	Biological context . . . . .	<b>18</b>
2.1.1	Neural activity . . . . .	18
2.1.2	Basics about neurons . . . . .	18
2.2	How to model neural activity? . . . . .	<b>20</b>
2.2.1	Dynamical systems and continuous models . . . . .	20
2.2.2	Point processes . . . . .	22
2.3	Our general model . . . . .	<b>26</b>
2.3.1	Mathematical context and motivation . . . . .	26
2.3.2	The main model . . . . .	31
2.4	Chapter organisation and contents . . . . .	<b>33</b>
2.4.1	Chapter 3: Multivariate Hawkes processes on inhomogeneous random graphs . . . . .	33
2.4.2	Chapter 4: Long-term stability of interacting Hawkes processes on random graphs . . . . .	40
2.4.3	Chapter 5: Stability of wandering bumps for Hawkes processes interacting on the circle . . . . .	42
2.4.4	Conclusion . . . . .	47

---

## 2.1 Biological context

### 2.1.1 Neural activity

Neurons are the primary components of the nervous system, as cells specialized in the reception, integration and transmission of information. While the human brain has an average of 86 billion neurons, there is huge diversity in the animal kingdom: insects have around a million neurons, and one of the world's smallest brain (a microscopic worm) contains around 300 neurons.

Neuronal dynamics are highly complex, and the study of neuronal mechanisms is fairly recent. [Galvani](#) discovered that nerve propagation is linked to electrical activity in the 18th century, while experimenting on frogs. Around 1900, [Bernstein](#) discovered the key element in information coding, the action potential. [Golgi](#) and [Ramón y Cajal](#) described neurons anatomically during the same period, observing complex and highly heterogeneous networks. [Broca](#) was interested in the spatial organization of the cortex. Advances in monitoring techniques used in science from the 1950s onwards such as EEG (*Electroencephalography*) and MRI (*Magnetic Resonance Imaging*) have contributed to the understanding of the physiology of a neuron. The first mathematical models describing neuronal activity followed, and are discussed below.

Nevertheless, our understanding of the brain is far from complete, and is still under study. We present here a brief summary of the mechanisms behind neural activity, starting with the ways in which one neuron communicates.

### 2.1.2 Basics about neurons

The emission of spikes encodes the transfer of information in the brain. In a few words, a neuron emits an action potential according to the information it receives from other neurons, and in return this action potential influences the dynamics of other neurons. These processes are complex, involving many chemical and electrical elements.

There are several families of neurons (sensory, motor, and others; see [Luo \(2020\)](#)), but they all share a number of common features: their membrane is polarized and they have distinct morphological regions that one can see in [Figure 2.1](#), each one with specific functions. The dendritic tree, made up of *dendrites* that are cellular extensions with many branches, is the zone where the neuron receives connections. The *soma*, the body of the neuron, contains the *nucleus*. The *axon* is a long extension that carries nerve signals away from the soma to an extensive set of branches called the terminal regions. Also called the *axon terminals*, they enable communication with other neurons through *synapses*, where chemical messengers are released.

The plasma membrane of the neuron is permeable: ion channels allow the exchange of ions between the inside and outside of the cell. The ions involved in

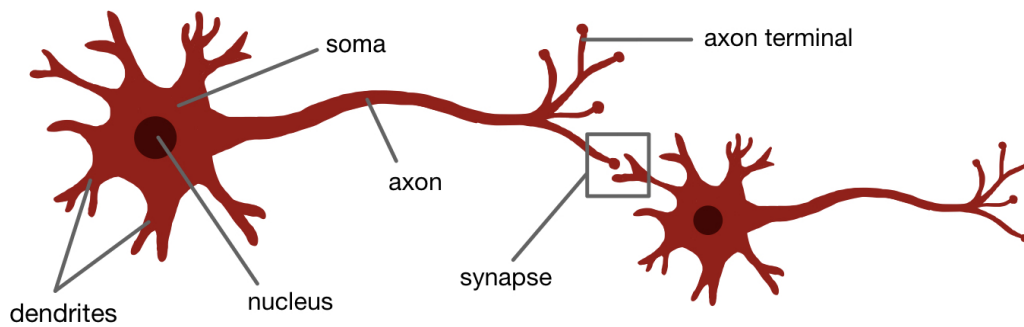


Figure 2.1: Diagram of a neuron

these exchanges are sodium  $Na^+$ , potassium  $K^+$ , calcium  $Ca^{2+}$  and chloride  $Cl^-$ . In the absence of interaction, differences in intra-membrane ionic concentration induce a potential difference, called the *resting state*. During synapses, the receptors of the dendritic tree disturb the permeability of ion channels in the membrane and thus induce a change in the membrane potential due to the transmembrane current of  $K^+$  and  $Na^+$ , as can be seen in Figure 2.2. If this change is large enough over a short time interval, that is if it reaches the voltage threshold, the neuron generates an all-or-nothing electric signal called an *action potential* or *spike*. We also say that the neuron *fires*. The diagram of a typical spike can be found in Figure 2.2. It consists in a depolarization (the entry of sodium ions into the cell) and a repolarization (the exit of potassium ions from the cell), followed by a refractory period in which the neuron cannot fire another spike (as the sodium channels recover), or with unusual difficulty with a higher threshold to reach.

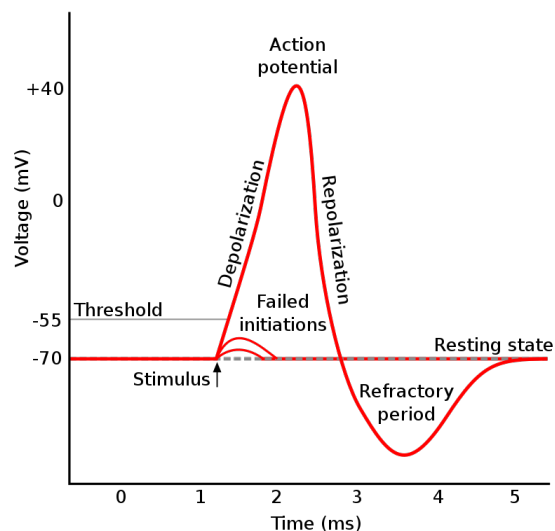


Figure 2.2: A typical action potential (from [https://commons.wikimedia.org/wiki/File:Action\\_potential.svg](https://commons.wikimedia.org/wiki/File:Action_potential.svg) Wikimedia Commons, CC BY-SA 3.0)

The duration of one spike is short, around 1 ms, and the shape of the action potential remains the same for a given neuron. Once emitted, the spike is propagated along the axon to the synapses, where the process of synaptic transmission occurs.

## 2.2 How to model neural activity?

Some simplifying mathematical models have been introduced to study neural activity. Their diversity reflects the large number of questions biologists have, and the scale we want to focus on: are we more interested in the evolution of the potential current? in the release of neurotransmitters? or in the spike trains of one or several neurons? We present in the following some of these models that are commonly found in mathematical neuroscience.

### 2.2.1 Dynamical systems and continuous models

Let us start with the so-called continuous models, that focus on the evolution of the membrane potential modeled by ordinary differential equations (ODE) or stochastic differential equations (SDE). They are not the focus of the thesis, but we present them here because they are the first historical models.

#### Single neurons and finite population of neurons (microscopic scale)

One way to simplify the biological process is to consider Integrate-and-fire models. Let  $u$  represent the membrane potential of a neuron, and let it follow some ODE  $\dot{u}_t = F(u_t)$ . When  $u$  reaches a given value, a fixed threshold  $U_{\text{th}}$ , we say that the neuron has emitted a spike and we let  $u$  take its resting value  $U_{\text{rest}}$ . This model has been introduced by [Lapicque \(1907\)](#) (see [Lapicque \(2007\)](#) for a recent translation) and extensively studied since, with variations integrating stochasticity (see [Galves & Löcherbach \(2013\)](#); [Delarue \*et al.\* \(2015\)](#); see also the review by [Sacerdote & Giraudo \(2013\)](#)). Note that the last articles consider a finite population of neurons, where each neuron is modeled by an ODE or a SDE with possible interaction inside the population.

Other models avoid fixing the spiking threshold, as it can exhibit variability in nature. The [Hodgkin-Huxley](#) model is probably the most famous one concerning the deterministic modelling of a neuron. Based on the study of a squid giant axon for which they received the 1963 Nobel Prize in Physiology or Medicine, [Hodgkin & Huxley \(1952\)](#) introduced a conductance-based dynamic system of 4 differential equations, representing interactions between the membrane potential ( $u_t$ ), potassium ( $n_t$ ) and sodium ( $m_t$ ) ion channels, and a leakage ( $h_t$ ) channel. The neuronal activity is thus modeled via its intrinsic dynamics

$$\begin{cases} du_t &= I_t dt - F(u_t, n_t, m_t, h_t) dt \\ dn_t &= (\alpha_n(u_t)(1 - n_t) - \beta_n(u_t)n_t) dt \\ dm_t &= (\alpha_m(u_t)(1 - m_t) - \beta_m(u_t)m_t) dt \\ dh_t &= (\alpha_h(u_t)(1 - h_t) - \beta_h(u_t)h_t) dt, \end{cases}$$

where  $F(v, n, m, h) = g_n n^4 (v - v_n) + g_m m^3 h (v - v_m) + g_h (v - v_h)$  with  $g_i$ ,  $v_i$ ,  $\alpha_i$  and  $\beta_i$  being conductance and rates associated with channel  $i$ . The behavior of this system depends heavily on the exterior stimulus  $I_t$  and have been thoroughly studied, see amongst others [Rinzel & Miller \(1980\)](#); [Aihara \*et al.\* \(1984\)](#), for stochastic versions

see Baladron *et al.* (2012); Bossy *et al.* (2015); Höpfner *et al.* (2017). However, due to its complexity, other models have been introduced to simplify the analysis.

The FitzHugh-Nagumo model (from FitzHugh (1961); Nagumo *et al.* (1962)) simplifies the Hodgkin-Huxley model by considering only two equations, one for the membrane potential ( $u_t$ ) and one for a recovery variable ( $C_t$ ):

$$\begin{cases} du_t &= \left( u_t - \frac{u_t^3}{3} - C_t + I_t \right) dt \\ \tau dC_t &= (u_t + a - bC_t) dt, \end{cases}$$

where  $a$ ,  $b$  and  $\tau$  are constant parameters and  $I$  the stimulus. Stochastic versions of this system also exist; see Baladron *et al.* (2012); Luçon & Poquet (2021b); Colombani & Bris (2022).

### Large population of neurons (mesoscopic scale)

While microscopic models are necessary to represent neuronal activity at the scale of one neuron, we may be interested in models on a larger scale, looking at a large portion of the cortex for example. However, at this scale, the assumption of identity of neurons is not realistic. We can therefore assign them labels. Let us mention two main examples that we will use in our work. Neurons are spatially located in the brain, and their activity or role can be linked to this spatial position. We can then label each neuron of a (large) population by its position. Another much studied model is the *ring model*, which focuses more specifically on the visual cortex. Instead of assigning each neuron to a spatial position, neurons are differentiated according to their role: each neuron possesses an *orientation preference* that stimulates its activity (see Georgopoulos *et al.* (1982); Bosking *et al.* (1997)). Then, in all generality, we do not consider finite population of neurons, but a distribution of neurons according to their label.

This led to the establishment of a purely phenomenological model: from the original works of Wilson & Cowan (1972); Amari (1977) about a conductance-based model of a network of synaptically coupled spiking neurons, Bressloff (2012) derives continuum *Neural Field Equations* (NFE) that can be written

$$\frac{\partial u_t(x)}{\partial t} = -\alpha u_t(x) + \int_S w(x, y) f(u_t(y)) dy + I_t(x). \quad (2.2.1)$$

Here,  $u_t(x)$  stands for the membrane potential of a neuron with label  $x$  and leaking rate  $\alpha$ . The kernel  $w$  represents the spatial distribution of neuronal synaptic connections, the function  $f$  the temporal filtering effects of the synaptic processing and  $I_t(x)$  the stimulus neuron with label  $x$  receives at time  $t$ . This equation models the large-scale dynamics of spatially structured biological neural networks. It has been extensively studied for the diversity of nonlinear dynamics patterns that can be obtained, such as oscillations, (traveling) waves and bumps, as evidenced by the wide range of articles on the subject, see Ermentrout & McLeod (1993); Veltz & Faugeras (2010); Bressloff & Webber (2012); Kilpatrick & Ermentrout (2013); Bressloff (2014); Touboul (2014); Faugeras & Inglis (2015); Inglis & MacLaurin

(2016); Lang & Stannat (2017); MacLaurin & Bressloff (2020). These structures appear thanks to the invariance of the Lebesgue measure (the integration  $dy$  in (2.2.1)).

### Comments on stochasticity

Let's take a slight digression to talk about stochasticity. One of the primary issues in mathematical neuroscience is the understanding of how the complexity of microscopic phenomena (such as those involved in neural activity, on the scale of the neuron) translates on a mesoscopic (for a portion of the cortex, for example) or macroscopic scale. While previous models were deterministic, stochastic versions (already mentioned) have been developed by adding noise to the ODEs. Introducing some noise allows us to take into account all the microscopic phenomena that we do not understand, or that are too difficult to model at the scale of interest.

Moreover, there is in particular biological stochasticity throughout spike emission. The following points can be found in Pouzat (2020):

- At the scale of a population of neurons, there might be variability within the population (of the threshold that allows spikes, of the release of chemical messengers, etc.).
- At the scale of the neuron, some spikes are spontaneous (Fatt & Katz (1952)). Due to the binary state of ion channels, there are some fluctuations from these changes of state (Verveen & Derksen (1968)). Moreover, some spikes are not correctly propagated along the axon and this physical failure may be random (Smith (1980)).
- At the scale of the synapse, the release of neurotransmitters is subject to fluctuations: the quantity of chemical messenger liberated by the presynaptic neuron may vary, the receptors of the postsynaptic neuron may be more or less occupied and the response of the receptors is also controlled by binary gates.

In the following and in our model, all this randomness is incorporated through the randomness associated with the emission of a spike, seen as a random phenomenon.

### 2.2.2 Point processes

A different point of view to model neural activity from SDEs or PDEs is to use point processes, which are collections of randomly located points. As seen in Figure 2.2, a spike is a stereotyped phenomenon that lasts only a short time, and information is coded in the inter-spike distribution. A way to simplify the modelling of spike trains is to represent a spike by one instant in time (the peak of the spike) and thus consider a spike as one event of a point process on the real line. See the following Figure 2.3 where a spike train from a cortical neuron is simplified into a point process on the real line.

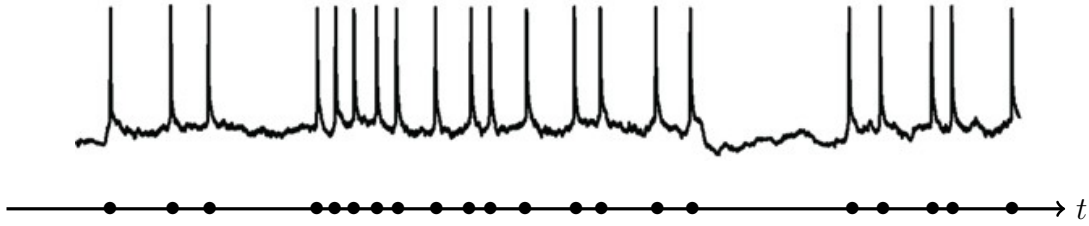


Figure 2.3: Response of a cortical neuron to a stimulus (from Rossant *et al.* (2011)) and the associated point process

Now that spike trains are represented by point processes, the question is how to characterise the arrival of the spikes. To do so, we introduce the notion of *stochastic intensity*, which explains the law of the arrival of the spikes.

### 2.2.2.1 Point processes and stochastic intensity

First, let us recall some elements of point process theory (Daley & Vere-Jones, 2006). A point process  $Z$  on  $\mathbb{R}_+$  can be equivalently represented by a strictly increasing sequence of jumping times  $(T_n)_{n \geq 1}$  with no accumulation point as  $n \rightarrow \infty$ , by a counting measure on  $\mathbb{R}_+$  defined by  $Z = \sum_{n \geq 1} \delta_{T_n}$  where  $\delta$  denotes the Dirac measure, or by the counting process  $t \mapsto Z(t) = \bar{Z}([0, t]) = \sum_{n \geq 1} \mathbf{1}_{t \leq T_n}$ . By convention, we set  $Z_0 = 0$ . For any function  $f$ , write

$$\int_{[0, t]} f(s) dZ(s) = \sum_{n \geq 1} f(T_n) \mathbf{1}_{T_n \leq t}.$$

The most famous point process is the homogeneous Poisson point process of parameter  $\lambda > 0$ , where the lengths of time between two jumps  $(T_{n+1} - T_n)_{n \geq 1}$  are independent and follow an exponential law of parameter  $\lambda$ . Then for any  $0 < a < b$ ,  $Z([a, b])$  follows a Poisson distribution with parameter  $\lambda(b - a)$ , and the constant  $\lambda$  is called the intensity of the process. When  $\lambda$  is a deterministic function  $\lambda : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ , we can define similarly an inhomogeneous Poisson point process of intensity  $\lambda$  where for any  $0 < a < b$ ,  $Z([a, b])$  follows a Poisson distribution with parameter  $\int_{[a, b]} \lambda(s) ds$ . However, in these two cases, the inter-spike intervals are independent. We do not have self-excitation nor interaction as we are looking for in the neuroscience setting. A way to introduce interaction it is to consider *stochastic intensities*, where the probability of a new jump depends on the history of the process.

**Definition 2.2.1.** Let  $(\mathcal{F}_t)$  be a filtration on  $\mathbb{R}_+$ , let  $Z$  be a point process  $(\mathcal{F}_t)$ -adapted on  $\mathbb{R}_+$  satisfying  $\mathbf{E}[Z(t)] < +\infty$  for all  $t \geq 0$ , and let  $\lambda$  be a progressive nonnegative process with respect to  $(\mathcal{F}_t)$ . We say that  $Z$  admits the  $(\mathcal{F}_t)$ -stochastic intensity  $\lambda$  if  $(Z(t) - \int_0^t \lambda_s ds)$  is a  $(\mathcal{F}_t)$ -martingale.

Most of the time, the filtration is chosen as  $\mathcal{F}_t = \sigma(Z(s), s \leq t)$  (the history of the process). The main point of introducing stochastic intensity is that if  $\lambda$  is the



intensity of  $Z$ , we have informally

$$\mathbf{P}(Z \text{ jumps between } (t, t + dt) | \mathcal{F}_t) = \lambda(t)dt.$$

The intensity is then seen as the *instantaneous jump rate* of the process: the larger the intensity is, the more likely is the process to jump. Having this rate depend on the past on the system allow us to include self-excitation and interaction in the model.

### 2.2.2.2 Thinning of Poisson Random Measures

An easy (and visual) way to construct point processes with stochastic intensities is to use Poisson random measures.

**Definition 2.2.2.**  $\pi$  is a *Poisson random measure* (PRM) on  $\mathbb{R}_+ \times \mathbb{R}_+$  with intensity  $dtdx$  if

1. for any Borel set  $C$  of  $\mathbb{R}_+ \times \mathbb{R}_+$ ,  $\pi(C)$  follows a Poisson distribution with parameter the Lebesgue measure of the set  $C$ ,
2. for all Borel sets  $C_1, \dots, C_n$  that are mutually disjoint,  $(\pi(C_1), \dots, \pi(C_n))$  is an independent family of random variables.

Point processes defined by means of their stochastic intensity can be represented by a *thinning* of a PRM, as e.g. done by Ogata (1988).

**Proposition 2.2.3.** Let  $(\mathcal{F}_t)$  be a filtration on  $\mathbb{R}_+$ , let  $\lambda$  be a progressive nonnegative process with respect to  $(\mathcal{F}_t)$  and let  $\pi$  be a PRM on  $\mathbb{R}_+ \times \mathbb{R}_+$  with intensity  $dtdx$ . Let, for any  $t \geq 0$ ,

$$Z(t) = \int_0^t \int_0^\infty \mathbf{1}_{z \leq \lambda(s)} \pi(ds, dz).$$

Then  $Z$  admits the  $(\mathcal{F})_t$ -stochastic intensity  $\lambda$ .

This construction has been widely used, see Brémaud & Massoulié (1996); Chevallier (2017); Chevallier *et al.* (2019). It means that, to construct  $Z$  which admits  $\lambda$  as stochastic intensity, we need to look at each point  $(s, z)$  of the PRM  $\pi$ , and if  $z \leq \lambda(s)$ ,  $s$  becomes a jump of the point process  $Z$ . Let us give a toy example in Figure 2.4 of the thinning method for a self-exciting process  $Z(t)$  with stochastic intensity  $\lambda(t) = 1 + \int_{]0, t[} \mathbf{1}_{[0, 1]}(t - s) dZ(s)$ : it means that as soon as there is a new jump, the intensity increases by 1 for 1 time unit. This temporary increase in the intensity implies self-excitation, in our example the third spike  $T_3$  occurs thanks to the excitation due to the second spike  $T_2$ . In fact,  $Z$  can be described as a *linear Hawkes process*.

### 2.2.2.3 Hawkes Processes

Hawkes Processes have been introduced by Hawkes (1971) as self-exciting and mutually exciting point processes. A general way to define a Hawkes process  $Z$  is to have a point process which admits a stochastic intensity of the form

$$\lambda(t) = f \left( v(t) + \int_0^{t-} h(t - s) dZ(s) \right), \quad (2.2.2)$$

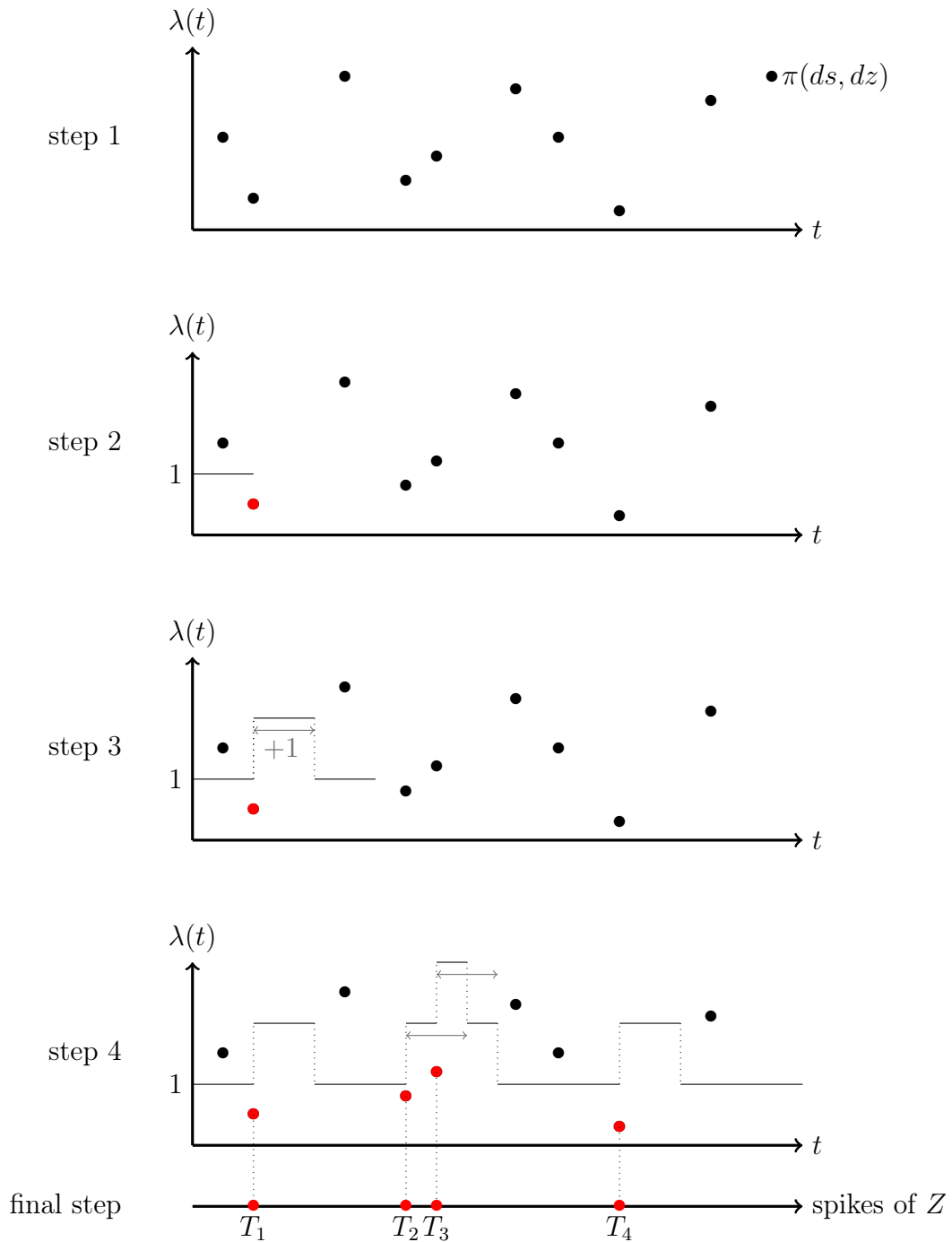


Figure 2.4: Thinning method to construct a point process  $Z$  with stochastic intensity  $\lambda(t) = 1 + \int_{0,t[} \mathbf{1}_{[0,1]}(t-s)dZ(s)$

Step 1: Draw a PRM  $\pi$ . Some of these points projected on the x-axis will be the spikes of the process  $Z$ .

Step 2: If there is no jump, the intensity remains at one until it is above a point of  $\pi$ .

Step 3: The red point under the intensity becomes a point of the point process  $Z$  and induces a change in the intensity as it rises by 1 for one time unit.

Step 4: Iterate, when the intensity is above a point of  $\pi$  it rises by 1 for one time unit and the point process  $Z$  gains a new point

final step: We put together all the red points projected on the x-axis to form the sequence  $(T_n)$  of jumps of  $Z$

where  $f : \mathbb{R} \rightarrow \mathbb{R}_+$ ,  $v : \mathbb{R}_+ \rightarrow \mathbb{R}$  and  $h : \mathbb{R}_+ \rightarrow \mathbb{R}$ . The function  $h$  is the memory kernel and represents how a point in the past of the process affects the current intensity. Hawkes processes were originally introduced to study earthquakes and their aftershocks in [Hawkes & Oakes \(1974\)](#); [Ogata \(1988\)](#), but have since been used in many other fields, such as finance and neuroscience. Note that the well posedness of such processes is classically established by a Picard argument as in [Delattre \*et al.\* \(2016a\)](#); [Chevallier \*et al.\* \(2019\)](#).

In the specific context of neuroscience (and more broadly to describe systems of interacting particles, see [Delattre \*et al.\* \(2016a\)](#)), we study large population of neurons so we consider multivariate Hawkes processes, which are Hawkes point processes evolving jointly. Each neuron is assigned to a point process (recall that each point represents a spike of the neuron), and the intensity of each neuron takes into account the past spikes of not only itself but also the past spikes of the population. A common writing for  $N$  neurons modeled by a multivariate Hawkes process is the process  $(Z_{N,1}, \dots, Z_{N,N})$  with respective intensities

$$\lambda_{N,i}(t) = f_i \left( v_i(t) + \frac{1}{N} \sum_{j=1}^N \int_0^{t-} w_{ij}^{(N)} h_{j \rightarrow i}(t-s) dZ_{N,j}(s) \right), \quad 1 \leq i \leq N, \quad (2.2.3)$$

where the functions  $f$ ,  $v$  and  $h$  in (2.2.2) may depend on the neuron, and  $w_{ij}^{(N)}$  represents the influence of the  $j$ th neuron on the  $i$ th neuron. We give an example of two neurons  $A$  and  $B$  (that is  $N = 2$ ) modeled with Hawkes processes  $Z_A$  and  $Z_B$  in [Figure 2.5](#), where neuron  $A$  is self-exciting and excites neuron  $B$ , and neuron  $B$  is self-inhibiting:

$$\lambda_A(t) = 1 + \int_{]0,t[} \mathbf{1}_{[0,1]}(t-s) dZ_A(s), \quad (2.2.4)$$

$$\lambda_B(t) = \max \left\{ 0, 1 + \int_{]0,t[} \mathbf{1}_{[0,1]}(t-s) dZ_A(s) - 2 \int_{]0,t[} \mathbf{1}_{[0,1]}(t-s) dZ_B(s) \right\}. \quad (2.2.5)$$

The function  $f(u) = \max\{0, u\}$  is used here with inhibition to ensure that the intensity of the point process  $Z_B$  stays nonnegative.

## 2.3 Our general model

In this thesis, we study a multivariate Hawkes process to model a population of  $N$  neurons interacting with each other. Before presenting formally our main model, we present the mathematical background and the main questions of interest.

### 2.3.1 Mathematical context and motivation

By modelling a population of  $N$  neurons interacting as a Hawkes process given by (2.3.7), we give a microscopic description of the neural activity for this population. In the neuroscience context, studying large particle systems makes sense due to the large amount of neurons in the brain. But as  $N \rightarrow \infty$ , one can derive simpler

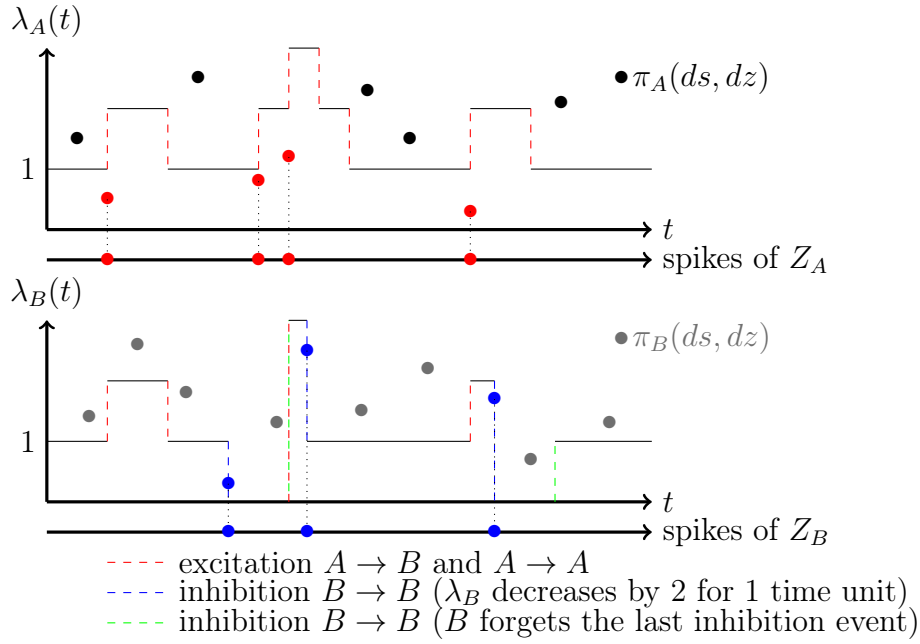


Figure 2.5: Two neurons  $A$  and  $B$  in interaction modeled by Hawkes processes, with respective stochastic intensities  $\lambda_A$  and  $\lambda_B$  given in (2.2.4) and (2.2.5).

Each neuron is assigned with a PRM ( $\pi_A$  and  $\pi_B$ , here the points were arbitrarily and pedagogically chosen). When neuron  $A$  fires a spike, it increases the intensity of both neurons by 1 for one time unit. When neuron  $B$  fires a spike, there is self inhibition and  $\lambda_B$  decreases by 2 for one time unit.

models from this microscopic point of view to obtain mesoscopic models of the neural activity.

Let us first introduce the work of [Delattre et al. \(2016a\)](#). The authors generalise the construction of multivariate Hawkes processes  $(Z_i)$  interacting on a countable directed graph. The main focus is made on the *mean-field* situation, where the interaction between a particle and all the others is approximated by the mean of interactions, and the  $N$  particles behave similarly. The process of interest is  $(Z_{N,1}, \dots, Z_{N,N})$  with respective intensities (2.2.3) for the choice  $v_i \equiv 0$ ,  $f_i \equiv f$ ,  $h_{j \rightarrow i} \equiv h$  and  $w_{ij}^{(N)} \equiv 1$ . In this case, the particles are exchangeable and [Delattre et al. \(2016a\)](#) show that the Hawkes process can be approximated by a suitable coupling with an i.i.d family  $(\bar{Z}_1, \dots, \bar{Z}_N)$  of inhomogenous Poisson processes of intensity  $\lambda_t$  solving

$$\lambda_t = f \left( \int_0^t h(t-s) \lambda_s ds \right). \quad (2.3.1)$$

Theorem 8 of [Delattre et al. \(2016a\)](#) states a *propagation of chaos* result, in the sense that for any  $T > 0$ , there exists  $C = C_{h,f,T}$  such that

$$\mathbf{E} \left[ \sup_{t \in [0, T]} |Z_{N,i}(t) - \bar{Z}_i(t)| \right] \leq \frac{C}{\sqrt{N}}. \quad (2.3.2)$$

The *large time* behavior of (2.3.1) is also treated in this paper for the linear case (where  $f(u) = \mu + u$  and  $h \geq 0$ ). There is a phase transition around the critical value  $\|h\|_1 = 1$ : in the subcritical case  $\|h\|_1 < 1$  the intensity  $\lambda_t$  (2.3.1) converges as  $t \rightarrow \infty$  towards a constant  $\frac{\mu}{1 - \|h\|_1}$  (Delattre *et al.*, 2016a, Theorem 10) whereas in the supercritical case  $\|h\|_1 > 1$ ,  $\lambda_t$  explodes exponentially (Delattre *et al.*, 2016a, Theorem 11).

Mean-field analysis goes back to McKean (1967); Sznitman (1991), originally for diffusion models where the process of interest is the empirical measure  $\mu_{N,t} := \frac{1}{N} \sum_{i=1}^N \delta_{\theta_{N,i}(t)}$  of  $N$  particles  $(\theta_{N,i}(t))_{1 \leq i \leq N, t \geq 0}$  diffusing according to the system

$$d\theta_{N,i}(t) = f(\theta_{N,i}(t))dt + \frac{1}{N} \sum_{j=1}^N g(\theta_{N,i}(t), \theta_{N,j}(t))dt + \sigma dB_{i,t} \quad (2.3.3)$$

with  $(B_{i,t})_{1 \leq i \leq N, t \geq 0}$  Brownian motions,  $g$  an interaction kernel and  $\sigma$  a diffusion coefficient. The mean-field limit to this particle system is described by the nonlinear process  $(\bar{\theta}(t))_{t \geq 0}$  solution to

$$d\bar{\theta}(t) = f(\bar{\theta}(t)) dt + \int g(\bar{\theta}(t), \theta) \mu_t(d\theta) + \sigma dB_t, \quad (2.3.4)$$

where  $\mu_t$  is the law of  $\bar{\theta}(t)$ . The classic propagation of chaos result states that for any  $T \geq 0$ , there exists  $C = C_{f,g,\sigma,T}$  such that

$$\mathbf{E} \left[ \sup_{t \in [0, T]} |\theta_{N,i}(t) - \bar{\theta}_i(t)| \right] \leq \frac{C}{\sqrt{N}}, \quad (2.3.5)$$

where  $(\theta_i(t))_{1 \leq i \leq N, t \geq 0}$  is an i.i.d. family of variables following (2.3.4), suitably coupled to  $(\theta_{N,i}(t))_{1 \leq i \leq N, t \geq 0}$  with the same Brownian motions and initial conditions (Sznitman (1991)). Equivalently on bounded time intervals  $\mu_{N,t}$  converges as  $N \rightarrow \infty$  towards  $\mu_t$ , the solution to some nonlinear Fokker-Planck equation. This is to be compared to our present situation of Hawkes processes where, as seen with Delattre *et al.* (2016a), in the mean-field analysis the large population limit is given in terms of an inhomogeneous Poisson process whose intensity solves a convolution equation, and similarly (2.3.5) is to be compared with (2.3.2).

One of the purposes of this thesis is to study the links between the microscopic scale, represented by a particular instance of (2.2.3), and its mesoscopic limit.

Interaction between neurons is a key element in information transfer mechanisms. There are about 10,000 billion nerve connections in 1  $cm^3$  of the human brain. However, assuming a complete interaction graph is not relevant in this context: neurons in the brain are not interchangeable, while they share a similar mechanism, their function, form and connections may differ. Some neurons are excitatory (they tend to promote the neuronal activity of their neighbors), while others are inhibitory (they tend to extinguish or diminish it). Some neurons are densely connected to the population, while others are not. Certain areas of the cortex are

also spatially differentiated: for example, the visual cortex has a columnar spatial organization, linked to the roles of the neurons, see [Bosking \*et al.\* \(1997\)](#). It is therefore important to examine the so-called *interaction graph* between neurons, and its influence on neuronal dynamics.

A first step to enrich the model of [Delattre \*et al.\* \(2016a\)](#) is to add labels to the population of particles. Let us mention the work of [Ditlevsen & Löcherbach \(2017\)](#): there, a population of  $N$  particles, divided into a finite number of classes, is also described by a multivariate Hawkes process. The interaction is inhomogeneous at the scale of the population as the classes interact differently with each other, but homogeneous at the scale of each class as the particles are exchangeable inside the same class. The paper studies the large population limit (as  $N \rightarrow \infty$ ) and also describes a propagation of chaos property: each class behavior can be approximated by a Poisson process. The large time behavior of this limit system is also considered in particular cases and exhibits oscillatory behaviors.

Other extensions of multivariate Hawkes processes exist, in particular the works of [Chevallier \(2017\)](#) consider an additional feature with age dependence, and [Duval \*et al.\* \(2022\)](#) consider two interacting classes of neurons (one exciting and one inhibiting). In [Chevallier \*et al.\* \(2019\)](#), the authors extend the mean-field framework to take into account the presence of a macroscopic spatial structure in the interaction. A position  $x_i$  is attributed to each particle  $i$ , and the intensities of the multivariate Hawkes process studied are given by (2.2.3) for the choice  $v_i(t) = e^{-\alpha t}v(x_i)$ ,  $f_i \equiv f$ ,  $h_{j \rightarrow i}(t) \equiv e^{-\alpha t}$  and  $w_{ij}^{(N)} = w(x_i, x_j)$ , where  $w$  is a matrix of synaptic lengths,  $v$  an initial input and  $\alpha > 0$  a leakage rate. The authors study the associated mean-field limit given by an inhomogeneous Poisson process of intensity  $\lambda_t(x) = f(u_t(x))$  where  $u$  solves the NFE (to compare with (2.2.1))

$$\partial_t u_t(x) = -\alpha u_t(x) + \int w(x, y) f(u_t(y)) \nu(dy) \quad (2.3.6)$$

with initial condition  $u_0(x) = v_0(x)$  and with  $\nu$  the distribution of the spatial positions. It was the first rigorous derivation of the NFE as a mean-field limit of spatially structured Hawkes processes. Compared to the first NFE introduced in (2.2.1), here there is no external stimulus ( $I_t = 0$ ) and the integration is done with respect to a probability measure  $\nu$  (of the position) instead of the Lebesgue measure. Note that this change disrupts the existence of traveling waves obtained with the classic NFE (2.2.1).

A motivation of our work is to further extend this inhomogeneity in the graph of interaction, and give a microscopic interpretation of the spatial structure in terms of random graphs. We will assume that the interaction between neurons is given by a possibly inhomogeneous and diluted graph, where the probability of presence of an edge (that is when  $w_{ij}^{(N)} \neq 0$  in (2.2.3)) depends on the positions of its vertices.

The main example that we have in mind concerns the class of  $W$ -random graph (see [Diaconis & Janson \(2008\)](#); [Lovász \(2012\)](#); [Janson \(2013\)](#); [Borgs \*et al.\* \(2018, 2019\)](#)). The previous works on particle systems with similar interaction address the case of diffusions as (2.3.3), see the wide range of articles on the

subject: Medvedev (2013); Delattre *et al.* (2016b); Coppini *et al.* (2019); Oliveira & Reis (2019); Bet *et al.* (2020); Luçon (2020); Bayraktar *et al.* (2023) among others. A question usually addressed is the dilution of the graph: to obtain a similar propagation of chaos as the classic mean-field framework (on a complete graph of interaction with exchangeability), one needs to make assumptions on the connectivity of the graph. For instance, in case of homogeneous Erdős-Rényi graph where two particles interact with each other with probability  $\rho_N$ , the classic threshold for the connectedness is given by  $\rho_N = \frac{\log(N)}{N}$  (Durrett (2006)). Another question is what sense to give to the notion of convergence of interactions when  $N \rightarrow \infty$ . In this direction, the works of Lovász (2012); Borgs *et al.* (2019, 2018) guided us to use the notion of *graphon*, that can be seen as an infinite graph. In this thesis, a large focus is made on the adaptation of similar propagation of chaos results for Hawkes processes on *inhomogeneous* random graphs. A common motivation between Hawkes processes and diffusions is to understand how the inhomogeneity of the underlying graph may or may not influence the long time dynamics of the system.

The behavior of the dynamics introduced (diffusions and Hawkes processes) on a time scale no longer bounded (but that may depend on the size of the population) is a current research problem: an issue common to all mean-field models (and their perturbations) is that there is, in general, no possibility to interchange the limits  $N \rightarrow \infty$  and  $t \rightarrow \infty$ . The usual propagation of chaos result seen in (2.3.2) is valid only on bounded time intervals, with  $C$  of the form  $\exp(CT)$ , so that it remains relevant only up to  $T \propto \log N$  in the case of Hawkes processes. It has been improved for the so-called subcritical case (Delattre *et al.*, 2016a, Remark 9) to  $C$  linear in  $T$ , so that it remains relevant up to  $T = o(\sqrt{N})$ . Others works have complemented this propagation of chaos result: Central Limit Theorems (CLT) have been obtained in Delattre *et al.* (2016a); Ditlevsen & Löcherbach (2017); Heesen & Stannat (2021) for homogeneous mean-field Hawkes processes (when both time and  $N$  go to infinity), with age-dependence in Chevallier (2017) and with spatial extension in Chevallier & Ost (2020). In particular, in the latter second order approximations of the microscopic system are given in terms of stochastic versions of the NFE. All of these works provide approximation results that are either valid on a bounded time interval  $[0, T]$  or under a strict growth condition on  $T$  (see in particular the condition  $T/N \rightarrow 0$  for the CLT in Ditlevsen & Löcherbach (2017)). Similar results have been considered in the context of diffusions, notably for phase oscillators (e.g. the Kuramoto model), see e.g. Bertini *et al.* (2014); Giacomin & Poquet (2015); Luçon & Poquet (2017); Giacomin *et al.* (2018); Luçon & Poquet (2021a); Coppini (2022) and references therein. In particular, Coppini (2022) goes even further by showing that the propagation of chaos is valid up to  $T \propto \exp(o(N))$ . In our context of Poissonian noise, we are concerned with time-scales that grow polynomially with  $N$ , and how these long time-scales influence the microscopic dynamics.

### 2.3.2 The main model

Let us now formally introduce the general model that we study throughout this thesis. We consider a population of  $N$  neurons,  $N \geq 1$ . Each one has a label  $x_i \in I \subset \mathbb{R}^d$  and its activity is represented by a point process  $Z_{N,i}$ , and  $Z_{N,i}(t)$  counts the number of spikes the  $i$ th neuron has emitted during the time interval  $[0, t]$ . The domain  $I$  represents the labelling of neurons, we have in mind for instance a spatial domain where each neuron is located in one position, or a role domain where each neuron is associated with one feature. Let  $\mathcal{F}_t := \sigma(Z_{N,i}(s), s \leq t, 1 \leq i \leq N)$ . The process  $(Z_{N,1}(t), \dots, Z_{N,N}(t))_{t \geq 0}$  is defined as a multivariate Hawkes process where for any  $1 \leq i \leq N$ ,  $Z_{N,i}(t)$  admits the  $(\mathcal{F}_t)$ -intensity

$$\lambda_{N,i}(t) = \mu_t(x_i) + f \left( v_t(x_i) + \frac{1}{N} \sum_{j=1}^N w_{ij}^{(N)} \int_0^{t-} h(t-s) dZ_{N,j}(s) \right). \quad (2.3.7)$$

At time  $t \geq 0$ , the different parameters stand for

- $\mu_t : I \rightarrow \mathbb{R}_+$  the spontaneous activity of the neuron,
- $f : \mathbb{R} \rightarrow \mathbb{R}_+$  the synaptic integration, that is the temporal filtering effects of the synaptic processing,
- $v_t : I \rightarrow \mathbb{R}_+$  the past activity or stimulus on the membrane of the neuron that may depend on the label,
- $w_{ij}^{(N)} \in \mathbb{R}$  the influence of the  $j$ th neuron on the  $i$ th neuron,
- $h : \mathbb{R}_+ \rightarrow \mathbb{R}$  the way the neuron remembers a past spike.

We can rewrite the intensity as

$$\lambda_{N,i}(t) = \mu_t(x_i) + f(v_t(x_i) + X_{N,i}(t-)) = \mu_t(x_i) + f(U_{N,i}(t-)), \quad (2.3.8)$$

and then  $U_{N,i}$  represents the membrane potential of the  $i$ th neuron, and  $X_{N,i}$  gathers the stochastic terms. The function  $f$  is thus the link between the membrane potential and the activity of the neuron.

Let us discuss possible choices for the parameters that appear in our model (2.3.7).

#### About the interaction

The main novelty and originality of this work, as already mentioned, is the choice of the interaction term  $w_{ij}^{(N)}$ . The classic mean-field framework takes  $w_{ij}^{(N)} = 1$  for all  $1 \leq i, j \leq N$  (Delattre *et al.* (2016a)), but as we saw before it is not biologically realistic in neuroscience. In this work and more especially in both Chapters 3 and 4, we go further as we provide a microscopic interpretation of the spatial structure in terms of random graphs. For  $N \geq 1$ ,  $w_{ij}^{(N)}$  is taken proportional to  $\xi_{ij}^{(N)}$  which follows a Bernoulli distribution with parameter  $W_N(x_i, x_j)$ :  $\xi_{ij}^{(N)} = 1$  means that the  $j$ th neuron excites the  $i$ th neuron whereas  $\xi_{ij}^{(N)} = 0$  means there is no edge  $j \rightarrow i$



in the graph of interaction, and the  $\left(\xi_{ij}^{(N)}\right)$  are mutually independent. The spatial kernel  $W_N$  stands for a microscopic interaction structure, depending on the labels of the two neurons considered. We assume we know a macroscopic limit of these synaptic interactions given by a graphon  $W$ . All the asymptotic results given in this work are *quenched*, it means that when we consider  $N \rightarrow \infty$ , we fix the sequence of the interaction terms  $\mathcal{G}^{(N)} := \left(\xi_{ij}^{(N)}\right)_{1 \leq i, j \leq N}$ . The idea behind is that while brains share the same interaction laws, the realization and connections among different brains may vary. Thus, we have two different sources of randomness: one from the interaction graph — in the following defined on a probability space  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mathbb{P})$ , and one from the spikes arrival — in the following defined on a filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbf{P})$ . Our main results are given  $\mathbb{P}$ -a.s.

This quenched setting, while new in the case of Hawkes processes, has been studied in the case of diffusions given by (2.3.3). Law of Large Numbers (LLN) and Large Deviations results on homogeneous Erdős-Rényi graphs (that is when  $W$  is a constant function) have been considered in Delattre *et al.* (2016b); Coppini *et al.* (2019); Oliveira & Reis (2019), and further extended to the inhomogeneous case in Bet *et al.* (2020); Luçon (2020); Bayraktar *et al.* (2023); Medvedev (2013) on a bounded time interval.

In the last Chapter 5 however, we do not consider the interaction to be random anymore as we fix  $w_{ij}$  proportional to a cosine kernel  $\cos(x_i - x_j)$  — depending on the preferential orientation of neurons labelled  $x_i$  and  $x_j$ , cf page 21. This choice is made to focus more on the particular asymptotic behavior in the framework introduced in the chapter.

### About the synaptic integration

A usual assumption for  $f$  is to be Lipschitz continuous (Brémaud & Massoulié (1996); Delattre *et al.* (2016a); Chevallier *et al.* (2019)). Common choices are to have  $f$  linear or affine, then the process  $(Z_{N,1}(t), \dots, Z_{N,N}(t))_{t \geq 0}$  is linear. In case of inhibition (that is, we allow  $U_{N,i}$  in (2.3.8) to take negative values), we have to take  $f \geq 0$ .

In this work, in both Chapters 3 and 4, the results are valid for general  $f$  Lipschitz continuous. In Chapter 4 we are interested in a particular case and we take  $f$  a sigmoid  $f_{\kappa, \varrho}(u) := \left(1 + e^{-(u-\varrho)/\kappa}\right)^{-1}$  with  $\kappa, \varrho > 0$ . This choice offers the advantage of having a smooth synaptic integration function  $f$  close to the Heaviside function  $H_\varrho = \mathbf{1}_{[\varrho, \infty)}$  when  $\kappa$  is small, thus allowing neurons to spike only if they have potential membrane high enough.

### About the memory kernel

A usual assumption for  $h$  is to be locally integrable or integrable (Brémaud & Massoulié (1996); Delattre *et al.* (2016a); Chevallier *et al.* (2019)). This memory kernel is central to the model because it governs how the system remembers the past: we can model long memory (if the support of  $h$  is large), short memory (for instance with our previous example in Figure 2.4 with  $h = \mathbf{1}_{[0,1]}$ )... It is also possible

to include inhibition by making  $h$  take negative values: in [Cattiaux \*et al.\* \(2022\)](#), the Hawkes processes are chosen with signed memory kernel allowing excitation and inhibition, and exhibiting a renewal structure when  $h$  is compactly supported.

The choice commonly found in the literature is that of the exponential kernel, with  $h(t) = \exp(-\alpha t)$  ([Zhu \(2015\)](#); [Chevallier \*et al.\* \(2019\)](#)): it means that the particle forgets the past with a leakage rate  $\alpha > 0$ . With this kernel, the Hawkes process is Markovian. Note that the Markov property remains valid in the more general setting of Erlang kernels of the type  $h(t) = \frac{t^n}{n!} e^{-\alpha t}$ , see e.g. [Duarte \*et al.\* \(2019\)](#). In the particular context of neuroscience, the exponential kernel presents also the advantage of coming back to the NFE (2.2.1), as shown in [Chevallier \*et al.\* \(2019\)](#): as said before, the macroscopic intensity is written  $\lambda_t(x) = f(u_t(x))$  where  $u$  solves the NFE (2.3.6).

In this thesis, the results of Chapter 3 and the first part of Chapter 4 are given for a general locally integrable memory kernel  $h$ . For the long time behavior studied in the second part of Chapter 4 and in Chapter 5, the results highly depend on this exponential kernel as the Markovian properties and the stability of the underlying large population limit dynamics resulting from it play a central role in the proofs.

## 2.4 Chapter organisation and contents

In this section, we present the main results of each chapter of this thesis, all of which being published or submitted articles. With each Proposition or Theorem, we indicate where it can be found in the manuscript. These results will be reintroduced in their respective chapters with their relationship to existing literature. Here the point is to give an exhaustive summary of the results demonstrated over the thesis.

### 2.4.1 Chapter 3: Multivariate Hawkes processes on inhomogeneous random graphs

In this thesis I focus on the study of the process  $(Z_{N,1}(t), \dots, Z_{N,N}(t))_{t>0}$  with intensity (2.3.7). In Chapter 3, we consider  $\mu_t = 0$  and  $w_{ij}^{(N)} = \kappa_i \xi_{ij}^{(N)}$ , where  $\kappa_i$  is a dilution parameter (so that the interaction term remains of order 1 as  $N \rightarrow \infty$ ) and where  $\xi_{ij}^{(N)}$  follows a Bernoulli distribution with parameter  $W_N(x_i, x_j)$ ,  $W_N$  being the microscopic interaction kernel, on a common probability space  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mathbb{P})$ , and the  $(\xi_{ij}^{(N)})$  are mutually independent.

The main motivation is to understand how the inhomogeneity in the interactions plays a role in the large population limit, and its long time behavior.

We fix the sequences of positions  $(\underline{x}^{(N)})_{N \geq 1} = \left( (x_1^{(N)}, \dots, x_N^{(N)}) \right)_{N \geq 1}$  and interactions  $(\xi^{(N)})_{N \geq 1} = \left( (\xi_{ij}^{(N)})_{i,j \in [1,N]} \right)_{N \geq 1}$ . We work on a filtered probability

space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbf{P})$  as usual, rich enough for our process of interest to be defined thanks to a family of Poisson random measures as seen in Definition 2.2.2.

**Definition 2.4.1.** Let  $(\pi_i(ds, dz))_{1 \leq i \leq N}$  be a sequence of i.i.d. Poisson random measures on  $\mathbb{R}_+ \times \mathbb{R}_+$  with intensity measure  $dsdz$ . A  $(\mathcal{F}_t)$ -adapted multivariate counting process  $(Z_{N,1}(t), \dots, Z_{N,N}(t))_{t \geq 0}$  defined on  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbf{P})$  is called a *multivariate Hawkes process* with the set of parameters  $(N, f, \xi^{(N)}, W_N, v, h, \underline{x}^{(N)})$  if  $\mathbf{P}$ -almost surely, for all  $t \geq 0$  and  $i \in \llbracket 1, N \rrbracket$ :

$$Z_{N,i}(t) = \int_0^t \int_0^\infty \mathbf{1}_{\{z \leq \lambda_{N,i}(s)\}} \pi_i(ds, dz) \quad (2.4.1)$$

with  $\lambda_{N,i}$  defined by

$$\lambda_{N,i}(t) = f \left( v_t(x_i^{(N)}) + \frac{\kappa_i^{(N)}}{N} \sum_{j=1}^N \xi_{ij}^{(N)} \int_0^{t-} h(t-s) dZ_{N,j}(s) \right). \quad (2.4.2)$$

One preliminary result concerns the well-posedness of  $(Z_{N,1}, \dots, Z_{N,N})$  given by (2.4.1).

**Proposition 2.4.2** (Proposition 3.2.5). *We suppose that  $f$  is Lipschitz continuous,  $h$  is locally square integrable on  $[0, +\infty)$ , that  $(t, x) \mapsto v_t(x)$  is continuous in  $t$  and Lipschitz continuous in  $x$  (uniformly in  $t$ ) and that  $v$  is bounded uniformly on  $\mathbb{R}_+ \times I$ . Then, for a fixed realization of the family  $(\pi_i)_{1 \leq i \leq N}$ , there exists a pathwise unique multivariate Hawkes process such that for any  $T < \infty$ ,*

$$\sup_{t \in [0, T]} \sup_{1 \leq i \leq N} \mathbf{E}[Z_{N,i}(t)] < \infty.$$

The existence and pathwise uniqueness of this process is standard (see Delattre *et al.* (2016a); Chevallier *et al.* (2019)) and is ensured by the use of the total variation distance and a Picard iteration argument.

### 2.4.1.1 Large population limit

We are interested in the large population behavior of  $(Z_{N,1}, \dots, Z_{N,N})$ . The first main result of Chapter 3 is the description of this limit in terms of an inhomogeneous Poisson process whose intensity involves the macroscopic spatial structure of the graph.

Recall (2.4.2), and let us look at its behavior when  $N \rightarrow \infty$  (large population). As we consider  $N$  positions  $(x_i)_{1 \leq i \leq N}$ , a minimal requirement is that the empirical distribution of the positions  $\frac{1}{N} \sum_{i=1}^N \delta_{x_i^{(N)}}$  has itself a macroscopic limit  $\nu$ . Regarding the interaction terms  $w_{ij}^{(N)}$  (that may depend on the labels), another minimal requirement is that the graph  $\mathcal{G}^{(N)} = (\{1, \dots, N\}, \xi^{(N)})$  converges in some way. The way to define this graph convergence uses the formalism of graph convergence introduced in Lovász (2012); Diaconis & Janson (2008) and further developed in Borgs *et al.* (2011, 2018, 2019) (and references therein). The idea is that, for every

$N \geq 1$ , one can divide the spatial domain  $I$  into  $N$  parts  $(B_{N,i})_{1 \leq i \leq N}$  of measure  $1/N$  with respect to the macroscopic distribution of the positions  $\nu$ . It is possible to represent the annealed graph  $\mathcal{G}_m^{(N)} = \left( \{1, \dots, N\}, (\kappa_i W_N(x_i, x_j))_{1 \leq i, j \leq N} \right)$  (here,  $m$  stands for *mean*) by a step function

$$W^{\mathcal{G}_m^{(N)}}(u, v) := \sum_{i, j=1}^N \kappa_i W_N(x_i, x_j) \mathbf{1}_{B_{N,i} \times B_{N,j}}(u, v)$$

defined on  $I \times I$ . The graphon convergence is then expressed in terms of the convergence of  $W^{\mathcal{G}_m^{(N)}}$  towards  $W : I \times I \rightarrow \mathbb{R}_+$  the macroscopic interaction graphon using the *cut-distance* norm defined by

$$\|W\|_{\square} := \sup_{S, T \subset I^2} \left| \int_{S \times T} W(x, y) \nu(dx) \nu(dy) \right|.$$

The analogy between this distance and graphs is detailed in the added Section 3.B. The simple case to keep in mind is when  $I = [0, 1]$ ,  $x_i = 1/N$  and  $B_{N,i} = \left( \frac{i-1}{N}, \frac{i}{N} \right]$  for  $1 \leq i \leq N$  (so that  $\nu(dx) = dx$ ): then the interaction graph  $\mathcal{G}^{(N)}$  (respectively its mean  $\mathcal{G}_m^{(N)}$ ) can be represented by a step function on the square  $I \times I$  that takes the value  $\xi_{ij}^{(N)}$  on  $B_{N,i} \times B_{N,j}$  (respectively  $\kappa_i W_N(x_i, x_j)$ ). For instance in Figure 2.6, we can see the representation of the step functions associated to the simulations of  $\mathcal{G}^{(100)}$  and  $\mathcal{G}^{(1000)}$ , and the graphon limit for the choice  $W_N(x, y) = W(x, y) = xy$  and  $\kappa_i \equiv 1$ .

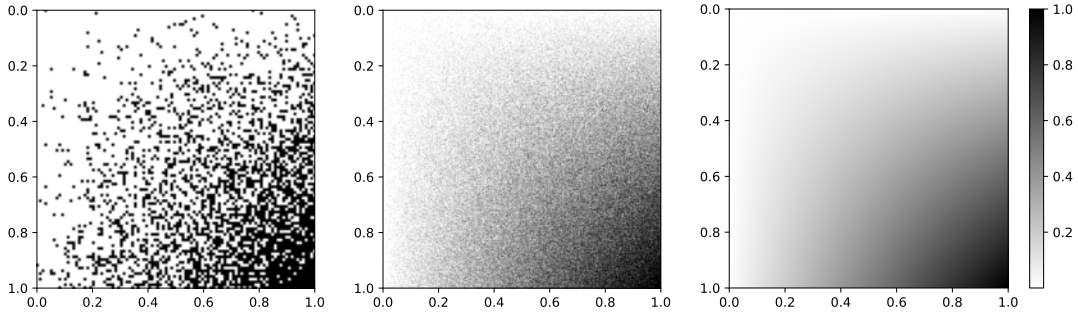


Figure 2.6: Graph convergence

On the figures on the left and the center, we represent the step function  $(x, y) \mapsto \sum_{i, j=1}^N \xi_{ij}^{(N)} \mathbf{1}_{\left( \frac{i-1}{N}, \frac{i}{N} \right]}(x) \mathbf{1}_{\left( \frac{j-1}{N}, \frac{j}{N} \right]}(y)$  on  $[0, 1]^2$ , where  $\xi_{ij}^{(N)} \sim \mathcal{B}\left(\frac{ij}{N^2}\right)$  for  $N = 100$  (left) and  $N = 1000$  (center). On the figure on the right, we represent the graphon limit  $W : (x, y) \mapsto xy$ .

Then, as  $N \rightarrow \infty$ , an informal Law of Large Numbers argument shows that the empirical mean in (2.4.2) becomes an expectation with respect to both the candidate limit for  $Z_{N,i}$  and the macroscopic law  $\nu$  of the positions: the sum in (2.4.2) becomes an integral over  $\nu$  on  $I$  and  $y$  describes the positions of the neurons, the microscopic interaction term  $w_{ij}^{(N)}$  becomes the macroscopic term  $W(x, y)$ , and the past activity of the neuron  $dZ_{N,j}(s)$  is replaced by its intensity in large population. Hence, the macroscopic description of a neuron at position  $x \in I$

should be described in terms of its intensity  $\lambda_t(x)$  solving the convolution equation

$$\lambda_t(x) = f \left( v_t(x) + \int_I W(x, y) \int_0^t h(t-s) \lambda_s(y) ds \nu(dy) \right). \quad (2.4.3)$$

Thus the limit process at position  $x$  is defined as an inhomogeneous Poisson point process with deterministic intensity  $(\lambda_t)_{t \geq 0}$  satisfying (2.4.3). This result is a spatial extension to [Delattre \*et al.\* \(2016a\)](#); [Chevallier \*et al.\* \(2019\)](#), with a microscopic interpretation of the interactions.

The well-posedness and uniqueness of the macroscopic intensity  $(\lambda_t)_{t \in [0, T]}$  for any  $T > 0$  is ensured under some regularity assumptions on  $W$  with respect to  $\nu$ . Let us mention in particular the macroscopic indegree boundedness hypothesis: we ask that  $x \mapsto D(x) = \int_I W(x, y) \nu(dy)$  has a Hölder regularity and is uniformly bounded on  $I$ . Then we can introduce a suitable coupling between the Hawkes process  $(Z_{N,1}, \dots, Z_{N,N})$  and independant Poisson processes  $(\bar{Z}_1(t), \dots, \bar{Z}_N(t))_{t \geq 0}$  of respective intensities  $(\lambda_t(x_i))_{1 \leq i \leq N}$  given by a thinning on the same family of PRM  $(\pi_i)_{1 \leq i \leq N}$  :

$$\bar{Z}_i(t) = \int_0^t \int_0^\infty \mathbf{1}_{\{z \leq \lambda_s(x_i)\}} \pi_i(ds, dz).$$

This coupling enables us to prove a propagation of chaos result on finite time intervals.

**Theorem 2.4.3** (Theorem 3.3.10). *Let  $T > 0$ . Under suitable regularity assumptions on the dilution of the graph  $\mathcal{G}_m^{(N)}$ , if the indegrees and outdegrees of the annealed graph are bounded, that is if there exists  $C$  such that*

$$\max \left\{ \sup_{1 \leq i \leq N} \frac{1}{N} \sum_{j=1}^N \kappa_j W_N(x_i, x_j), \sup_{1 \leq j \leq N} \frac{1}{N} \sum_{i=1}^N \kappa_i W_N(x_i, x_j) \right\} \leq C,$$

then for  $\mathbb{P}$ -almost realizations of the connectivity sequence and positions,

$$\frac{1}{N} \sum_{i=1}^N \mathbf{E} \left[ \sup_{t \in [0, T]} |Z_{N,i}(t) - \bar{Z}_i(t)| \right] \xrightarrow{N \rightarrow \infty} 0.$$

A direct consequence of these large population results concerns the behavior as  $N \rightarrow \infty$  of the empirical distribution on the space  $\mathbb{D}([0, T], \mathbb{N}) \times I$  of trajectories and positions. Define the probability measures:

$$\begin{aligned} \mu_N(d\eta, dx) &:= \frac{1}{N} \sum_{i=1}^N \delta_{(Z_{N,i}([0, T]), x_i^{(N)})}(d\eta, dx), \\ \mu_\infty(d\eta, dx) &:= P_{[0, T], \infty}(d\eta | x) \nu(dx), \end{aligned}$$

where  $P_{[0, T], \infty}(\cdot | x)$  is the law of an inhomogeneous Poisson point process with intensity  $(\lambda_t(x))_{0 \leq t \leq T}$  (solution of (2.4.3)). Note that  $\mu_N$  is random. Then, under

the same assumptions, Theorem 3.3.15 gives that  $\mathbf{E} [d_{BL}(\mu_N, \mu_\infty)] \xrightarrow[N \rightarrow \infty]{} 0$ , where  $d_{BL}$  is the bounded Lipschitz distance defined by

$$d_{BL}(\mu_1, \mu_2) := \sup_{g, \|g\|_L + \|g\|_\infty \leq 1} \left| \int_E g (d\mu_1 - d\mu_2) \right|.$$

Another consequence relates to the spatial profile

$$U_N(t)(x) := \sum_{i=1}^N U_{N,i}(t) \mathbf{1}_{B_{N,i}}(x), \quad x \in [0, 1],$$

where  $U_{N,i}(t)$  is defined in (2.3.8) in the case  $I = [0, 1]$ ,  $x_i = i/N$  and  $\nu(dx) = dx$ . In this case the membrane potential  $U_N(t)$  has a large population limit  $u_t$  defined by

$$u_t(x) := v_t(x) + \int_0^1 W(x, y) \int_0^t h(t-s) \lambda_s(y) dy$$

in the following sense:

**Proposition 2.4.4** (Proposition 3.3.17). *Under suitable regularity assumptions on the parameters, for  $\mathbb{P}$ -almost realizations of the connectivity sequence and positions,*

$$\mathbf{E} \left[ \int_0^T \int_0^1 |U_N(t, x) - u_t(x)| dx dt \right] \xrightarrow[N \rightarrow \infty]{} 0.$$

### 2.4.1.2 Asymptotics of $\lambda_t$ in the linear case: phase transition

The second part Chapter 3 focuses on the long time behavior of the macroscopic intensity  $\lambda_t$  solving (2.4.3), in the linear case, that is when  $f(u) = u$  in (2.4.2) with  $h, v_t, W \geq 0$ . We assume in this part that  $\lim_{t \rightarrow \infty} v_t = v_\infty$  with  $v_\infty$  sufficiently smooth. The aim is to understand how the inhomogeneity of the graph affects the phase transition described for mean-field linear Hawkes processes in Delattre *et al.* (2016a) previously mentioned on page 27.

Let us first consider an explicit example, the exponential case  $h(t) = \exp(-\alpha t)$  with the kernel interaction  $W(x, y) = f(x)g(y)$  for  $f, g \geq 0$  (and assuming without loss of generality  $\int_I g d\nu = 1$ ). It is possible to compute exactly the solution  $\lambda$  of (2.4.3), and show that there is also a phase transition in terms of the sign of  $\langle f, g \rangle - \alpha$ , where  $\langle f, g \rangle = \int_I f(x)g(x)\nu(dx)$ :

$$\begin{aligned} \langle f, g \rangle > \alpha &\Rightarrow \forall x \in I, \quad \lambda_t(x) \xrightarrow[t \rightarrow \infty]{} +\infty \text{ and} \\ \langle f, g \rangle < \alpha &\Rightarrow \forall x \in I, \quad \lambda_t(x) \xrightarrow[t \rightarrow \infty]{} v_\infty(x) + \frac{\langle g, v \rangle}{\alpha - \langle f, g \rangle} f(x), \end{aligned}$$

see Proposition 3.4.2 and Example 3.4.3 for precise computations.

Adding the spatial inhomogeneity in a more general context (that is, for general  $h$  and  $W$ ) leads also to a new phase transition defined in terms of the linear operator

$$\begin{aligned} T_W : L^\infty(I) &\longrightarrow L^\infty(I) \\ g &\longmapsto (T_W g : x \longmapsto \int_I W(x, y)g(y)\nu(dy)) \end{aligned}$$

and more specifically its spectral radius  $r_\infty$ . The phase transition is now given in terms of  $\|h\|_1 r_\infty < 1$  (subcritical case) and  $\|h\|_1 r_\infty > 1$  (supercritical case).

In the subcritical case we find again that  $\lambda_t$  converges when  $t \rightarrow \infty$ .

**Theorem 2.4.5** (Theorem 3.4.4). *Assume  $\|h\|_1 r_\infty < 1$ . Under suitable regularity assumptions on the parameters,*

- *there exists a unique function  $\ell : I \mapsto \mathbb{R}^+$  solution of  $\ell(x) = v_\infty(x) + \|h\|_1 \int_I W(x, y) \ell(y) \nu(dy)$  for all  $x \in I$ , continuous and bounded on  $I$ ,*
- *for any  $x \in I$ ,  $\lambda_t(x) \xrightarrow[t \rightarrow \infty]{} \ell(x)$ .*

This asymptotic limit can be explicitly defined thanks to the linear operator by  $\ell = \sum_{k=0}^{\infty} \|h\|_1^k T_W^k v_\infty$ . When  $v_\infty$  is constant, we can even write for all  $x \in I$

$$\ell(x) = \left( \sum_{k=0}^{\infty} \|h\|_1^k D^{(k)}(x) \right) v_\infty,$$

with  $D^{(0)} = 1$ ,  $D^{(1)} = D(x)$  and  $D^{(k+1)} = T_W D^{(k)}$ . We can see how the inhomogeneity of the underlying graph influences the macroscopic dynamics. We obtain also with Proposition 3.4.5 an important characterization: if  $(t, x) \mapsto v_t(x)$  is constant, then  $\ell$  is uniform (i.e.  $x \mapsto \ell(x) = \ell$ ) if and only if the indegree is uniform (i.e.  $x \mapsto D(x) = \int_I W(x, y) \nu(dy) = D$ ) and in such a case,  $r_\infty = D$ . A consequence is that we can have a homogeneous asymptotic limit with an inhomogeneous graph interaction if the indegree of the graph is constant (we can say that the graph is not "inhomogeneous enough").

With uniform interaction (i.e.  $W = 1$ ), the subcritical threshold  $\|h\|_1 r_\infty > 1$  reduces to  $\|h\|_1 > 1$  and it can be shown (see Delattre *et al.* (2016a), Theorem 11) that  $\lambda_t \sim \alpha e^{\beta t} \rightarrow \infty$  for some  $\alpha, \beta > 0$ . In our context with nontrivial  $W$ , different behaviors inside the population are possible: some parts of the population can have a converging intensity and others can explode. Let us give an example of this situation for  $I = [0, 1]$ : consider  $W(x, y) = \alpha \mathbf{1}_{[0, \frac{1}{2}]^2}(x, y) + \beta \mathbf{1}_{[\frac{1}{2}, 1]^2}(x, y)$  for  $\alpha > \beta$ , then  $r_\infty = \frac{\alpha}{2}$ . This corresponds to two disconnected mean-field components  $A$  (for neurons with positions in  $I_A = [0, \frac{1}{2})$ ) and  $B$  (for neurons with positions in  $I_B = [\frac{1}{2}, 1]$ ). The critical parameter for population  $A$  (respectively  $B$ ) is hence  $\alpha_c = \frac{2}{\|h\|_1}$  (respectively  $\beta_c = \frac{2}{\|h\|_1}$ ). With  $\alpha > \alpha_c$  and  $\beta < \beta_c$ , we have indeed  $\|h\|_1 r_\infty > 1$  but one does not have  $\lambda_t \xrightarrow[t \rightarrow \infty]{} \infty$  uniformly on  $I$  as the population  $B$  is subcritical, we only have  $\lambda_t(x) \xrightarrow[t \rightarrow \infty]{} \infty$  for  $x \in I_A$ . This possibility of having different asymptotic behaviors (i.e. converging towards a constant or  $\infty$ ) along the spatial domain as  $t \rightarrow \infty$  led us to add hypotheses on the connectivity of the graphon.

**Proposition 2.4.6** (Proposition 3.4.8). *For  $k \geq 1$ , let  $W^{(k)}$  be the kernel of the operator  $T_W^k$ , that is  $W^{(k)}(x, y) := \int_{I \times \dots \times I} W(x, x_1) \cdots W(x_{k-1}, y) dx_1 \cdots dx_{k-1}$ . Assume there exists  $k$  such that  $W^{(k)} > 0$  and assume that  $W$  is symmetric and smooth enough, then in the supercritical case  $\|h\|_1 r_\infty > 1$ ,  $\int_I \lambda_t(x)^2 \nu(dx) \xrightarrow[t \rightarrow \infty]{} \infty$ .*

The added hypotheses ask the graphon  $W$  to be sufficiently connected, so that having two different asymptotic behaviors on  $I$  for  $\lambda_t$  is not possible anymore.

### 2.4.1.3 Mid-conclusion and next steps

The first part of this chapter sets out the formal framework of the model (2.3.7) and demonstrates its convergence in large population towards inhomogeneous Poisson processes of intensities  $\lambda_t$  on bounded time intervals. In the second part of this chapter, the focus is on the special case of the linear framework and the asymptotic behavior of the limit  $\lambda_t$  by exhibiting a phase transition. We can sum up these results by writing the different convergences of the intensities defining the processes

$$\begin{aligned} \lambda_N(t) &\xrightarrow[N \rightarrow \infty]{} \lambda_t \xrightarrow[t \rightarrow \infty]{\text{linear case}} \ell \quad \text{for the subcritical case,} \\ \lambda_N(t) &\xrightarrow[N \rightarrow \infty]{} \lambda_t \xrightarrow[t \rightarrow \infty]{\text{linear case}} +\infty \quad \text{for the supercritical case.} \end{aligned}$$

There is, in general, no possibility to interchange the limits  $N \rightarrow \infty$  and  $t \rightarrow \infty$ . The natural question is then: can we find some frameworks for which the approximation of  $\lambda_N$  remains valid for unbounded time scales?

The next step is to look at the long-term behavior of the process, and find cases where the limits  $t \rightarrow \infty$  and  $N \rightarrow \infty$  could be treated jointly. As the analysis of mean-field interacting processes on long time scales has a significant history in the case of interacting diffusions, we were inspired by the Kuramoto model that focus on the empirical measure  $\mu_{N,t} := \frac{1}{N} \sum_{i=1}^N \delta_{\theta_{i,t}}$  of the system of interacting diffusions

$$(\theta_1, \dots, \theta_N) \text{ solving the system of coupled SDEs } d\theta_{i,t} = -\frac{K}{N} \sum_{j=1}^N \sin(\theta_{i,t} - \theta_{j,t}) dt +$$

$dB_{i,t}$ . Standard propagation of chaos techniques show the weak convergence of  $\mu_N$  on a bounded time interval  $[0, T]$  to the solution  $\mu_t$  to the nonlinear Fokker-Planck (NFP) equation  $\partial_t \mu_t = \frac{1}{2} \partial_\theta^2 \mu_t + K \partial_\theta \left( \mu_t (\sin * \mu_t) \right)$ , to be compared to our  $U_N(t)$  that converges towards  $u_t$  solving the NFE. The Kuramoto model presents a phase transition: when  $K \leq 1$ ,  $\mu \equiv \frac{1}{2\pi}$  is the only (stable) stationary point of the previous NFP (subcritical case), whereas it coexists with a stable circle of synchronised profiles when  $K > 1$  (supercritical case), we refer to Bertini *et al.* (2014); Luçon & Poquet (2017); Giacomini *et al.* (2012); Coppini (2022) for references on the Kuramoto model. Our two following articles (Chapters 4 and 5) propose hypotheses on the parameters for recovering similar behaviors within the framework of Hawkes processes. Chapter 4 recovers a similar subcriticality condition that can be written  $\|f\|_\infty \|h\|_1 r_\infty < 1$  under which there is a unique stationary limit for the NFE. Chapter 5 presents a framework where this subcriticality condition is not satisfied and with particular choices for  $f$ ,  $h$  and  $W$ , so that the NFE admits a manifold of stable stationary solutions.

One main difficulty for these works was to deal with the different noise (a jumping Poissonian noise rather than a continuous Brownian one) in the microscopic dynamics.



## 2.4.2 Chapter 4: Long-term stability of interacting Hawkes processes on random graphs

In this Chapter, we carry on with the study of our main model (2.3.7). We restrict ourselves to the case  $I = [0, 1]$  and  $x_i = \frac{i}{N}$  for  $1 \leq i \leq N$ . We take  $w_{ij}^{(N)} = \frac{\xi_{ij}^{(N)}}{\rho_N}$ , where  $\rho_N$  is a dilution parameter,  $\xi_{ij}^{(N)}$  follows a Bernoulli distribution with parameter  $\rho_N W(x_i, x_j)$  and  $W$  is a macroscopic interaction kernel. This is a simplification of the interactions introduced in Chapter 3 by taking  $\kappa_i \equiv \rho_N^{-1}$  and  $W_N = \rho_N W$ . We give another writing of the intensity (2.3.7) as we consider

$$\lambda_{N,i}(t) = F(X_{N,i}(t-), \eta_t(x_i)),$$

where  $\eta$  represents deterministic contributions to the intensity and  $X_{N,i}$  is given in (2.3.8). It is straightforward to go back to (2.3.7) by considering  $\eta = (\mu, v)$  and  $F(X, \eta) = \mu + f(v + X)$ .

### 2.4.2.1 Extension of the asymptotic behavior of $\lambda_t$ in the general case

The first part of Chapter 4 concerns the extension of Theorem 2.4.5 when  $f$  is no longer restricted to be linear. It affects the subcritical case definition in the sense that we ask now  $\|\partial_X F\|_\infty \|h\|_1 r_\infty < 1$  instead of  $\|h\|_1 r_\infty < 1$ . We state a general convergence result as  $t \rightarrow \infty$  of  $X_t$ , the large population limit of  $X_N$  defined as

$$X_t(x) = \int_I W(x, y) \int_0^t h(t-s) \lambda_s(y) ds dy.$$

**Theorem 2.4.7** (Theorem 4.2.7). *Assume there exists  $\eta_\infty$  such that  $\eta_t \rightarrow \eta_\infty$  uniformly on  $I$ . In the subcritical case and under suitable regularity assumptions on the parameters,*

- (i) *there exists a unique continuous function  $X_\infty : I \mapsto \mathbb{R}^+$  solution of  $X_\infty = \|h\|_1 T_W F(X_\infty, \eta_\infty)$ ,*
- (ii)  *$X_t$  converges uniformly on  $I$  when  $t \rightarrow \infty$  towards  $X_\infty$ .*

This Theorem directly implies that  $\lambda_t$  converges uniformly on  $I$  towards  $\ell$  the solution to  $\ell = F(\|h\|_1 T_W \ell, \eta_\infty)$ , the correspondance being  $X_\infty = \|h\|_1 T_W \ell$ . This result comes from the classic study and convergence of an ODE.

In the particular case of the exponential memory function ( $h(t) = \exp(-\alpha t)$ ), we show that  $X_\infty$  is a stable stationary point to

$$\partial_t Y_t = -\alpha Y_t + T_W (\partial_X F(X_\infty, \eta_\infty) Y_t), \quad (2.4.4)$$

which is the linearised system associated with the dynamics of  $Y_t := (X_t - X_\infty)$ . The subcritical condition, which becomes  $\|\partial_X F\|_\infty r_\infty < \alpha$ , translates into the existence of a spectral gap for the operator associated with (2.4.4).

**Proposition 2.4.8** (Proposition 4.2.10). *Assume that the memory kernel is  $h(t) = e^{-\alpha t}$  and define the linear operator*

$$\begin{aligned} \mathcal{L} : L^2(I) &\longrightarrow L^2(I) \\ g &\longmapsto \mathcal{L}(g) = -\alpha g + T_W(\partial_X F(X_\infty, \eta_\infty)g). \end{aligned}$$

*Under suitable regularity assumptions on the parameters,  $\mathcal{L}$  generates a contraction semi-group  $(e^{t\mathcal{L}})_{t \geq 0}$  on  $L^2(I)$  such that for any  $g \in L^2(I)$ ,  $\|e^{t\mathcal{L}}g\|_2 \leq e^{-t\gamma}\|g\|_2$  where  $\gamma := \alpha - r_\infty \|\partial_X \bar{F}\|_\infty > 0$  is the spectral gap.*

### 2.4.2.2 Long-term stability of the microscopic profile $X_N$ for the exponential case

The main contribution of the article is centered on the long-term stability of the microscopic profile  $X_N$  for the particular choice of the exponential kernel  $h(t) = e^{-\alpha t}$ . To study the dynamics of  $X_N$ , we assume some additional dilution controls on the graph: in the case of  $F$  bounded for instance, we ask  $N\rho_N^2 \rightarrow \infty$ .

Recall the large population convergence of the membrane potential  $U_N$  seen in Proposition 2.4.4. We show first uniform convergence of  $X_N(t)$  in  $L^2(I)$  instead of  $L^1(I)$ .

**Proposition 2.4.9.** *Let  $T > 0$ . In the subcritical case and with the exponential memory kernel, for any  $\varepsilon > 0$ ,  $\mathbb{P}$ -a.s.*

$$\mathbf{P} \left( \sup_{t \in [0, T]} \|X_N(t) - X_t\|_2 \geq \varepsilon \right) \xrightarrow{N \rightarrow \infty} 0.$$

Hence  $X_N(t)$  is close to  $X_t$  on any bounded time interval. The asymptotic dynamics of  $(X_t)$  is known, it converges towards  $X_\infty$  which is stable as seen in Proposition 2.4.8. It is then natural to study the proximity between  $X_N(t)$  and this limit  $X_\infty$ . We thus state the main result of the paper, the proximity between  $X_N(t)$  and  $X_\infty$  valid up to arbitrary polynomial times in  $N\rho_N$ .

**Theorem 2.4.10** (Theorem 4.2.12). *Choose some  $t_f > 0$  and  $m \geq 1$ . Then, under regularity assumptions on the parameters  $(F, \eta)$ , under dilution controls of the graph, in the subcritical case and with the exponential memory kernel,  $\mathbb{P}$ -a.s. for any  $\varepsilon > 0$ ,*

$$\mathbf{P} \left( \sup_{t \in [t_\varepsilon, (N\rho_N)^m t_f]} \|X_N(t) - X_\infty\|_2 \geq \varepsilon \right) \xrightarrow{N \rightarrow \infty} 0,$$

for some  $t_\varepsilon > 0$  independent of  $N$ .

Let us give a short sketch of the proof, divided into several steps.

- We give a semimartingale decomposition of  $Y_N := X_N - X_\infty$ . The point is to decompose the dynamics of  $Y_N$  in terms of the linear dynamics (2.4.4), with

drift terms coming from the graph and its mean-field approximation, a noise term and a quadratic remaining error coming from the nonlinearity of  $F$ . We do a careful control on each of these terms in the semimartingale expansion on a bounded time interval.

- The variable  $t_\varepsilon$  corresponds to the time the deterministic dynamics of  $X_t$  needs to reach a neighborhood of  $X_\infty$  of size  $\varepsilon$ . With the spectral gap seen in Proposition 2.4.8, one can show that  $t_\varepsilon$  is of order  $-\log \varepsilon/\gamma$ . Then, using Proposition 2.4.9, the microscopic process  $X_N$  is itself  $\varepsilon$ -close to  $X_\infty$  with high probability.
- The previous argument is the starting point of an iterative procedure that works as follows: the point is to see that, provided  $X_N$  is initially close to  $X_\infty$ , it will remain close to  $X_\infty$  on some  $[0, T]$  for some sufficiently large deterministic  $T > 0$ . The key argument is that on a bounded time interval, the deterministic linear dynamics dominates over the contribution of the noise, so that one only has to wait some sufficiently large  $T$  so that the deterministic dynamics prevails against the other contributions. The rest of the proof consists in an iterative procedure from the previous argument, taking advantage of the Markovian structure of the dynamics of  $X_N$  (because of the exponential memory kernel). The time horizon at which one can pursue this recursion, polynomial in  $N\rho_N$ , is controlled by moment estimates on the noise.

### 2.4.2.3 Mid-conclusion and comments

This Chapter presents thus a case for which the mean-field limit of our microscopic system is stable up to polynomial times: we have

$$X_N(t) \xrightarrow[N \rightarrow \infty]{} X_t \xrightarrow[t \rightarrow \infty]{\text{exponentially}} X_\infty,$$

and for the general subcritical case and an exponential memory kernel, we can write

$$X_N(t) \xrightarrow[N \rightarrow \infty, t \text{ polynomial in } (N\rho_N)]{\mathbb{P}} X_\infty.$$

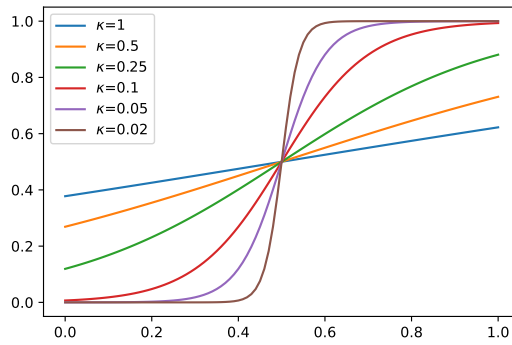
However the Chapter raises new questions.

The subcriticality condition  $\|\partial_x F\|_\infty r_\infty < \alpha$  is used several times: it ensures the existence of a unique fixed point to  $X_\infty = \|h\|_1 T_W F(X_\infty, \eta_\infty)$ , it guarantees the convergence of  $X_t$  towards  $X_\infty$  and the long-term stability of  $X_N$  around  $X_\infty$ . What happens if this subcriticality condition is not met anymore?

The different proofs rely on the fact that in the framework of the Chapter, the dynamics of the membrane potential  $X_t$  admit a unique stationary solution  $X_\infty$ . What happens when we do not have this uniqueness anymore?

## 2.4.3 Chapter 5: Stability of wandering bumps for Hawkes processes interacting on the circle

In this Chapter, we study our main model (2.3.7) in a particular context, when the subcriticality condition  $\|f\|_\infty r_\infty \|h\|_1 < 1$  is not met. To do so, we focus on a case


 Figure 2.7: Graph of  $f_{\kappa, \varrho}$  for  $\varrho = 0.5$  and different choices of  $\kappa$ 

where the NFE no longer admits a unique stationary solution.

We consider the ring model that focuses on the visual cortex (Kilpatrick & Ermentrout (2013)): each neuron possesses an orientation preference  $x_i$  in  $I = (-\pi, \pi] =: S$ , and we assume that the labels  $(x_i)_{1 \leq i \leq N}$  are regularly distributed on  $I$ . The interaction term is no longer random and we consider inhibition as we take a cosine kernel:  $w_{ij} = 2\pi \cos(x_i - x_j)$ , so that neurons with similar preferential orientations tend to excite each other, while those with opposite preferential orientations inhibit each other. The synaptic function  $f$  is fixed as a sigmoid of parameter  $(\kappa, \varrho)$ , where  $\kappa$  is related to the slope of the sigmoid at its threshold point  $\varrho$ :

$$f_{\kappa, \varrho}(u) := (1 + e^{-(u-\varrho)/\kappa})^{-1}.$$

We stay in the exponential case as we take  $h(t) = e^{-t}$  (the leakage rate is fixed as  $\alpha = 1$ ), and moreover we take  $\mu_t = 0$  and  $v_t(x) = \rho(x)e^{-t}$  with  $\rho$  an initial distribution of the membrane potential.

We thus study the multivariate Hawkes processes  $(Z_{N,1}(t), \dots, Z_{N,N}(t))_{t \geq 0}$  with intensities

$$\lambda_{N,i}(t) = f_{\kappa, \varrho} \left( \rho(x_i)e^{-t} + \frac{2\pi}{N} \sum_{j=1}^N \cos(x_i - x_j) \int_0^{t-} e^{-(t-s)} dZ_{N,j}(s) \right), \quad i = 1, \dots, N.$$

When  $\kappa \rightarrow 0$ , the sigmoid  $f_{\kappa, \varrho}$  converges to the Heaviside function  $H_\varrho(u) = \mathbf{1}_{u \geq \varrho}$  with threshold  $\varrho$ , see in Figure 2.7 the illustration of the influence of the choice of  $\kappa$  on  $f_{\kappa, \varrho}$ . Moreover, the slope of  $f_{\kappa, \varrho}$  as  $\kappa \rightarrow 0$  is steeper, so that we are not in the subcritical case as  $\|f_{\kappa, \varrho}\|_\infty \gg 1$ . Thus, for  $\kappa$  small enough, when its membrane potential is high a neuron emits a spike at rate close to 1, and this rate is close to 0 if the potential is not high enough.

This model is linked to a particular instance of the NFE introduced in (2.2.1): the limit of the membrane potential  $U_N$  is given by  $u_t$  which solves the NFE

$$\frac{\partial u_t(x)}{\partial t} = -u_t(x) + \int_S \cos(x - y) f_{\kappa, \varrho}(u_t(y)) dy, \quad t \geq 0. \quad (2.4.5)$$

Recall our previous comment on page 21 on the NFE (2.2.1), where thanks to the translation invariance of the Lebesgue measure on  $\mathbb{R}$ , one can show the existence

of traveling waves solutions (see also [Lang & Stannat \(2016\)](#) for stability estimates of these traveling waves), and the NFE (2.3.6) derived by [Chevallier et al. \(2019\)](#), which requires a probability measure  $\nu$  on  $I$  but where such traveling waves are lost. The present choice of  $I = S$  and  $\nu(dy) = \frac{1_{[-\pi, \pi]}}{2\pi} dy$  combines the two previous advantages:  $\nu$  is a probability measure (hence the large population analysis when  $N \rightarrow \infty$  applies) and translation invariance is preserved in the present periodic case. In particular as explained below, the NFE (2.4.5) exhibits localized patterns, called *bumps*, that are stationary pulse solutions.

### 2.4.3.1 A manifold of stationary solutions to the NFE

We first study the asymptotic behavior of  $u_t$  satisfying (2.4.5). From a standard Grönwall estimate, we can show that for any  $T > 0$ , there exists a unique solution  $(u_t)_{t \in [0, T]}$  in  $\mathcal{C}_b(S, \mathbb{R})$  to (2.4.5) with initial condition  $u_0 = g$  smooth enough. We can then define the flow of (2.4.5) by  $(t, g) \mapsto \psi_t(g)$ , that is the solution at time  $t$  of (2.4.5) starting from  $g$  at  $t = 0$ :

$$\psi_t(g)(x) = e^{-t}g(x) + \int_0^t e^{-(t-s)} \int_S \cos(x-y) f_{\kappa, \varrho}(\psi_s(g)(x)) ds. \quad (2.4.6)$$

We show that, unlike in the previous chapter, there is no uniqueness in the limit of  $u_t$  when  $t \rightarrow \infty$ : there exists a manifold of (stable) stationary solutions to (2.4.5).

**Proposition 2.4.11** (Proposition 5.2.2 and Corollary 5.2.6). *If  $\kappa$  is small enough, there exists  $A > 0$  solving  $A = \int_S \cos(y) f_{\kappa, \varrho}(A \cos(y)) dy$  such that the following manifold of functions  $S \rightarrow \mathbb{R}$  indexed by  $S$*

$$\mathcal{U} := (A \cos(\cdot + \phi))_{\phi \in S} =: (u_\phi)_{\phi \in S}$$

*is a set of stationary solutions to (2.4.5). Moreover,  $\mathcal{U}$  is locally stable under the flow (2.4.5): there exists  $\varepsilon_0 > 0$  such that, for any  $g \in L^2(S)$  satisfying  $\text{dist}_{L^2}(g, \mathcal{U}) \leq \varepsilon_0$ , we have  $\lim_{t \rightarrow \infty} \text{dist}_{L^2}(\psi_t(g), \mathcal{U}) = 0$ .*

The manifold  $\mathcal{U}$  is then a set of stationary solution  $(u_\phi)_{\phi \in S}$  parameterized by  $S$ . Each function is a translation of a cosine, a *bump* on  $S$ . We can define a proper phase reduction of functions along  $\mathcal{U}$ : as the manifold  $\mathcal{U}$  is stable and attractive, a solution to the NFE (2.4.5) from a neighborhood of  $\mathcal{U}$  is attracted to  $\mathcal{U}$ . As  $t \rightarrow \infty$ , it identifies with one stationary solution of the manifold, its isochron.

**Proposition 2.4.12** (Isochronal phase, Proposition 5.2.8). *For any  $g$  close enough to  $\mathcal{U}$ , there exists a unique  $\theta(g) \in S$  called the isochronal phase of  $g$  such that*

$$\|\psi_t(g) - u_{\theta(g)}\|_2 \xrightarrow[t \rightarrow \infty]{} 0. \quad (2.4.7)$$

Isochrons have been introduced by [Winfree \(1974\)](#) and studied by [Guckenheimer \(1975\)](#), in a larger context than here with periodic orbits. The main idea is that, given a metric space  $(E, d)$  and a dynamics evolving with a flow map  $(t, x) \in \mathbb{R} \times M \mapsto \psi_t(x)$  with  $M \subset E$  a smooth manifold, one can define the orbit of  $\psi$  through  $x \in M$  by the set  $\{\psi_t(x), t \in \mathbb{R}\}$ . The orbit is periodic if we can define

$\tau = \inf \{t > 0, \psi_t(x) = x\}$ . If for all  $t \geq 0$ ,  $\psi_t(x) = x$ ,  $x$  is a stationary point to the dynamics of  $\psi$  — with a period of zero. If  $\Gamma$  is a stable limit cycle of  $\psi$ , that is a periodic orbit stable in the sense that there exists a neighborhood  $V$  of  $\Gamma$  such that for all  $y \in V$ ,  $d(\psi_t(y), \Gamma) \xrightarrow[t \rightarrow \infty]{} 0$ , one can define the isochron of  $x \in \Gamma$  as expressed in [Giacomin \*et al.\* \(2018\)](#) by

$$W(x) = \left\{ y \in V : \lim_{t \rightarrow \infty} d(\psi_t(x), \psi_t(y)) = 0 \right\}.$$

The isochron is then constituted of points of  $V$  that have a similar behavior to the dynamics starting at  $x$ , meaning the points that are *synchronized* with  $x$ .

In our context, we consider stationary solutions to the dynamics that constitute a manifold  $\mathcal{U}$  parameterized by  $S$ . Each  $u_\phi \in \mathcal{U}$  can be seen as a stable limit point of  $\psi$  (that is considering  $\Gamma_\phi := \{u_\phi\}$ ), and one can similarly define the isochron by

$$W(u_\phi) = \left\{ g \in V(u_\phi) : \lim_{t \rightarrow \infty} d(\psi_t(g), u_\phi) = 0 \right\},$$

as  $\psi_t(u_\phi) = u_\phi$  for all  $t \geq 0$  (the period is zero). We can reformulate this by saying that the isochron of  $u_\phi$  is constituted of the functions  $g$  close enough to  $\mathcal{U}$  such that the flow  $\psi$  brings  $g$  on  $u_\phi$ . One can show that  $\phi \in S \mapsto W(u_\phi)$  is injective, so that we can define the *isochronal phase* on a well-chosen neighborhood of  $\mathcal{U}$  by  $\theta : g \mapsto \phi \in S$  such that  $g \in W(u_\phi)$ , which is exactly (2.4.7). Here we do not have *one* stable limit cycle  $\Gamma$ , but a *manifold* of stable stationary solutions (which we can denote informally by  $\mathcal{U} = \cup_{\phi \in S} \Gamma_\phi$ ), so that the isochronal phase  $\theta$  is a way to define a projection of a neighborhood of  $\mathcal{U}$  to  $S$ , and thus we can reduce the dimensional complexity: the system dynamics can be seen as phase dynamics.

### 2.4.3.2 Long time behavior: the emergence of a Brownian motion

The rest of the Chapter is about our microscopic process  $(Z_{N,1}(t), \dots, Z_{N,N}(t))_{t>0}$  and its proximity with  $\mathcal{U}$  provided the initial condition is close to the manifold.

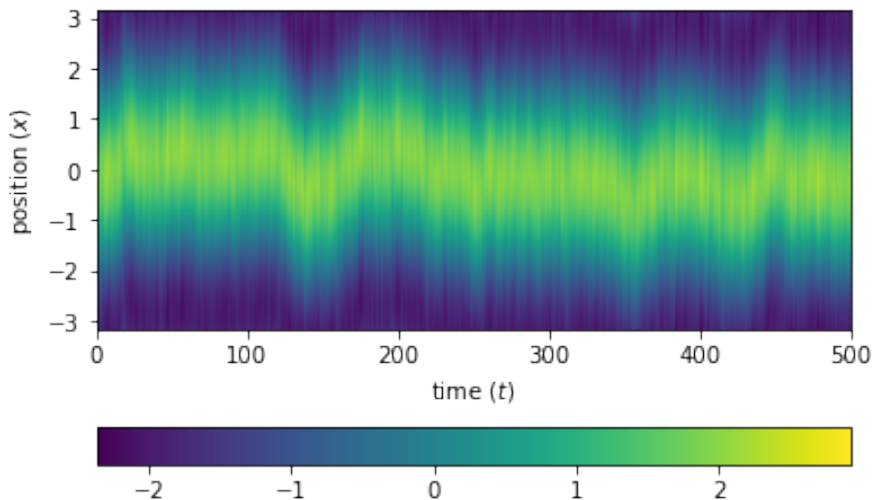
The first result ensures that  $U_N(t)$  reaches a neighborhood of  $\mathcal{U}$  in time of order  $\log(N)$  and stays inside it for arbitrary polynomial times in  $N$ .

**Theorem 2.4.13** (Theorem 5.2.9). *Suppose that  $\rho$  is smooth enough and close enough to  $\mathcal{U}$ . Let  $\alpha, \tau_f > 0$ . There exists some  $C > 0$  such that, defining for any  $N \geq 1$ ,  $T_0(N) := C \log(N)$ , for any  $\varepsilon > 0$ ,*

$$\mathbf{P} \left( \sup_{t \in [T_0(N), N^\alpha \tau_f]} \text{dist}_{L^2}(U_N(t), \mathcal{U}) \leq \varepsilon \right) \xrightarrow[N \rightarrow \infty]{} 1.$$

For this result, we use the same strategy as the one explained for Theorem 2.4.10 in page 41. Here, the time  $T_0(N) \propto \log(N)$  plays the same role as  $t_\varepsilon$  introduced for Theorem 2.4.10, it is the time we have to wait for which  $\text{dist}_{L^2}(U_N(T_0(N)), \mathcal{U}) \leq \frac{N^{2\eta}}{\sqrt{N}}$ , with  $0 < \eta < \frac{1}{4}$ . This essentially boils down to following the predominant deterministic dynamics of the NFE.

Now that we know that  $U_N(t)$  is close to  $\mathcal{U}$ , we can use the isochronal phase and an adapted rescaling to characterize the behavior of  $U_N(t)$  along the manifold.

Figure 2.8: Evolution of  $U_N(t)(x)$ 

We run a simulation for  $N = 500$  and represent the evolution of  $U_N(t)(x)$ . The time  $t \in [0, 500]$  is on the x-axis and the position  $x \in S$  is on the y-axis. The system is initialised close to  $\mathcal{U}$  with  $\rho(x) = A(\kappa) \cos(x) + \cos(2x)$ , for the choice  $(\kappa, \varrho) = (\frac{1}{20}, \frac{1}{2})$ .

**Theorem 2.4.14** (Theorem 5.2.11). *Under the same hypotheses with  $\alpha = 1$ , let  $\tau_f > 0$ . There exist a deterministic  $\theta_0 \in S$  and for every  $N$  some  $\tau_0(N) \propto \frac{\log(N)}{N}$  and a càdlàg process  $(W_N(t))_{t \in (\tau_0(N), \tau_f)}$  that converges weakly in  $\mathbb{D}([0, \tau_f], S)$  towards a standard Brownian such that for every  $\varepsilon > 0$ ,*

$$\lim_{N \rightarrow \infty} \mathbf{P} \left( \sup_{\tau \in (\tau_0(N), \tau_f)} \|U_N(N\tau) - u_{\theta_0 + \sigma W_N(\tau)}\|_2 \leq \varepsilon \right) = 1,$$

where  $\sigma := \left( 2\pi \int_S \sin^2(x) f(A \cos(x)) dx \right)^{\frac{1}{2}}$ .

The main idea behind this result is that, provided the initial condition is close enough to  $\mathcal{U}$ , the deterministic dynamics brings the potential close to the manifold of stationary solutions  $\mathcal{U}$ , and the noise induces a random phase shift on the circle. Moreover, this phase shift correctly renormalized is quantified by a Brownian motion. On Figure 2.8, one can see an illustration of this theorem: the maximum intensity seen by the population (in yellow), initially in 0 spatially wanders through time.

Let us give a short sketch of the proof. The time  $N\tau_0(N)$  corresponds to the required time for  $U_N(t)$  to be sufficiently close to  $\mathcal{U}$ : it is related to  $T_0(N)$  given by Theorem 2.4.13. We use the semimartingale decomposition of  $U_N$  given by  $dU_N(t) = B_N(t)dt + dM_N(t)$ , with  $B_N$  some drift and  $M_N$  a martingale, and Itô formula to write the semimartingale decomposition of  $\theta(U_N(t))$  on the interval  $[T_0(N), N\tau_f]$ , for which one can show a careful control on each of the terms appearing. The difficulty here is to show rigorously that there is no macroscopic drift appearing on this time scale (this point is essentially due to the invariance by rotation of the whole problem). After rescaling the time by  $N$ , we identify the

noise with a Brownian motion thanks to Aldous' tightness criterion and Lévy's characterization.

### 2.4.4 Conclusion

The first article (Chapter 3) deals with the process (2.3.7) in all its generality, and lays a solid foundation for more detailed study of particular cases. The "random graph" vision required a lot of work and adaptation to find the right metric and the least restrictive assumptions for our model. The two following articles (Chapter 4 and 5) explore the long-time stability of the microscopic system introduced in Chapter 3, give two examples of long-time behavior, one where the system is attracted to a single stationary solution to the NFE and the other where the system is attracted to a manifold of stationary solutions and wanders on it.

### Perspectives

#### Scaling

Let us briefly comment on the scaling choice  $1/N$  made in (2.3.7). It is a usual scaling found in the mean-field framework and the propagation of chaos: in the large population limit, as  $N$  goes to infinity a particle is connected to more and more particles and to counterbalance this, the strengths of the interactions reduce. However, different scaling can be found in the litterature such as  $1/\sqrt{N}$ , the diffusive regime. See the LLN results in the case of exchangeable particles in Erny *et al.* (2022), the LLN and CLT results where there is a random environment in Erny (2023) or with inhibition in Pfaffelhuber *et al.* (2022).

A question could be how to adapt our propagation of chaos result of Chapter 3 (on finite time intervals) with this diffusive scaling while keeping inhomogeneity between the interactions. The convergence of graph we used here is not relevant anymore and should potentially be changed to adapt to this different scaling.

#### Dilution

Our results of Chapter 3 work on possibly diluted graph, with a controlled dilution: for the Erdős-Rényi random graph for instance where  $\rho_N w_{ij}^{(N)} \sim \mathcal{B}(\rho_N)$ , we essentially ask that  $\rho_N \gg \log(N)/N$  (see the details on page 62).

Can we go beyond this dilution condition? Several works with diffusions as Oliveira *et al.* (2020); Lacker *et al.* (2023) study limit theorems with the interaction made on sparse graphs. In particular Oliveira *et al.* (2020) shows that for the stochastic Kuramoto model, the propagation of chaos result (on finite time intervals) is valid when  $N\rho_N \rightarrow c \in \mathbb{R}_+$  and in such a case the limiting object is described in terms of a dynamics evolving on a Galton-Watson tree.

Similar results might be applied to Hawkes processes. The first step is to formally define an (infinite) Hawkes process on a Galton-Watson tree, in this direction one could adapt the results of Delattre *et al.* (2016a) for the well posedness of Hawkes process on infinite graph. The main difficulty would be to adapt the formalism developped by Oliveira *et al.* (2020) for a similar large population convergence. The other interesting point would be the long time dynamics of such



particle system living on a Galton-Watson tree, see similarly the case of nearest neighbors interactions on  $\mathbb{Z}^d$  in [Delattre \*et al.\* \(2016a\)](#).

### Stable limit cycle: towards periodicity

On page 45, we introduced the notion of isochrons for stable limit cycle  $\Gamma$  with period  $\tau$  for dynamics with flow  $\psi$ . In Chapter 5, we only need the case where  $\tau = 0$  as we focus on stationary solutions. A natural extension could be to consider (deterministic) dynamics allowing a stable limit cycle  $\Gamma = \{q_t, t \in [0, \tau)\}$  of the flow, with period  $\tau > 0$ .

It is the case in the article [Ditlevsen & Löcherbach \(2017\)](#), presented on page 29: the population is divided into several classes, and the large population limit is described for each class by a nonlinear differential equation driven by a Poisson random measure. They consider also a particular case where the system interact through a monotone feedback loop, and prove the existence of a non trivial attractive periodic orbit for the large population limit system. To adapt these results to our situation (and obtain large time behavior of the microscopic system around the limit cycle), there are several difficulties. First the memory function is not the same: in [Ditlevsen & Löcherbach \(2017\)](#)  $h$  is chosen as an Erlang kernel whereas we use a decreasing exponential in both Chapters 4 and 5. Erlang kernels offer the same advantage of Markovian structure as our exponential kernel, but the computations are harder as they follow a cascade structure as seen in [Ditlevsen & Löcherbach \(2017\)](#). The exponential kernel also enables us to find a spectral gap  $\gamma$  in the dynamics of the limit process, thus having stability and attraction to the manifold as stated in Proposition 2.4.8. It is not clear whether this spectral gap persists in case of Erlang kernels. Another difficulty relies on the analysis of the behavior on the limit cycle: in our case, we prove that the Brownian motion that emerges does not have any drift, but with a periodic stable limit, one can expect that the deterministic dynamics would drive a drift. Characterizing and computing this drift remains open.

Note that periodicity seems to be extremely linked to inhibition, see for example [Duval \*et al.\* \(2022\)](#), even if the existence of stable cycles is not yet proven.

### Random graph in the ring model

In Chapter 5, the interaction between neurons is deterministic, unlike the two previous chapters. This choice was made to focus mainly on the large time dynamics and to study how the Brownian motion that the isochronal phase follows appears, which is a similar result as what was done for diffusions in [Bertini \*et al.\* \(2014\)](#); [Giacomin \*et al.\* \(2018\)](#); [Luçon & Poquet \(2020, 2021b\)](#). A natural question is then, what happens when we add the random graph framework to Chapter 5?

In this direction, [Coppini \(2022\)](#) in the case of diffusions on random graphs shows the long time stability of the microscopic process around the limit manifold. Similar results in our situation can be expected, the main issue is how to characterize the behavior on the manifold: the random interactions add a new term in the Itô's decomposition used in the proof of Theorem 2.4.14, and it is not clear how it affects the dynamics.

**Time dependent graph**

Another direction could be to consider time dependent graphs with  $\xi_{ij} = \xi_{ij}(t)$ , see for instance [Bhamidi \*et al.\* \(2019\)](#); [Nguyen \*et al.\* \(2020\)](#). With the neuroscience context, the idea behind is to model learning behavior of the brain, where the interactions between neurons can evolve through time to represent new learnt paths. One difficulty to deal with is the different time scales involved: one can imagine to have a slow dynamics for the graph interactions evolution (representing the learning process), and a fast one for the neural activity. How to correctly define Hawkes processes on such interaction dynamics and the large population and long time behaviors resulting remain open issues.



# Chapter 3

## Multivariate Hawkes processes on inhomogeneous random graphs

*This chapter consists in a modified version of my article [Agathe-Nerine \(2022\)](#) (expanded with some computations), published in *Stochastic Processes and their Applications*, with minor wording and notation changes for harmonization purposes.*

In this Chapter, we study our general model (2.3.7) for the choice  $\mu_t = 0$  and  $w_{ij}^{(N)} = \kappa_i \xi_{ij}^{(N)}$ , where  $\kappa_i$  is a dilution parameter (so that the interaction term remains of order 1 as  $N \rightarrow \infty$ ) and where  $\xi_{ij}^{(N)}$  follows a Bernoulli distribution with parameter  $W_N(x_i, x_j)$ ,  $W_N$  being the microscopic interaction kernel.

### Abstract

We consider a population of  $N$  interacting neurons, represented by a multivariate Hawkes process: the firing rate of each neuron depends on the history of the connected neurons. Contrary to the mean-field framework where the interaction occurs on the complete graph, the connectivity between particles is given by a random possibly diluted and inhomogeneous graph where the probability of presence of each edge depends on the spatial position of its vertices. We address the well-posedness of this system and Law of Large Numbers results as  $N \rightarrow \infty$ . A crucial issue will be to understand how spatial inhomogeneity influences the large time behavior of the system.

### Contents

---

3.1	Introduction . . . . .	<b>53</b>
3.1.1	Biological and mathematical context . . . . .	53
3.1.2	Our model . . . . .	55
3.1.3	Organisation of the paper . . . . .	55
3.2	A system of $N$ interacting particles on a graph and its limit . .	<b>56</b>
3.2.1	Notation . . . . .	56

3.2.2	The model . . . . .	57
3.2.3	Large population limit process . . . . .	58
3.3	Convergence of the model in large population . . . . .	<b>59</b>
3.3.1	Coupling . . . . .	59
3.3.2	Hypotheses . . . . .	60
3.3.3	Convergence . . . . .	62
3.3.4	Consequence on the empirical measure . . . . .	63
3.3.5	Spatial profile . . . . .	64
3.4	Large time behavior of the limit process in the linear case . . .	<b>65</b>
3.4.1	The exponential case . . . . .	66
3.4.2	Subcritical case . . . . .	67
3.4.3	Supercritical case . . . . .	68
3.5	Applications . . . . .	<b>70</b>
3.5.1	A general class of examples . . . . .	70
3.5.2	Example: Erdős-Rényi graph . . . . .	71
3.5.3	Example: P-nearest neighbor model (Omelchenko <i>et al.</i> , 2012) . . . . .	72
3.5.4	Example: Inhomogeneous graph with EDD (Chung & Lu, 2002) . . . . .	72
3.5.5	Example: Multi-class interaction populations . . . . .	76
3.6	Possible extensions . . . . .	<b>76</b>
3.7	Existence and uniqueness of the model and its limit (proofs) .	<b>78</b>
3.7.1	Proof of Proposition 3.2.5 . . . . .	78
3.7.2	Proof of Theorem 3.2.7 . . . . .	81
3.8	Convergence of the mean-field process (proofs) . . . . .	<b>85</b>
3.8.1	Toolbox . . . . .	85
3.8.2	Proof of Theorem 3.3.10 . . . . .	87
3.8.3	Proof of Theorem 3.3.12 . . . . .	92
3.8.4	Proofs: Application to the Scenarios of Definition 3.3.8	94
3.9	The empirical measure and the spatial profile (proofs) . . . . .	<b>107</b>
3.9.1	Proof of Theorem 3.3.15 . . . . .	107
3.9.2	Proof of Proposition 3.3.17 . . . . .	110
3.10	Behavior in large time limit - Linear case (proofs) . . . . .	<b>111</b>
3.10.1	Proof of Theorem 3.4.4 . . . . .	111
3.10.2	Proof of Propositions 3.4.7 and 3.4.8 . . . . .	114
3.A	Useful results . . . . .	<b>116</b>
3.A.1	Various technical lemmas . . . . .	116
3.A.2	Spectral theory . . . . .	118

---

3.A.3	Jentzsch/Krein–Rutman Theorem . . . . .	119
3.A.4	Renewal theory . . . . .	119
3.B	About matrices, graphs and graphon . . . . .	119

---

## 3.1 Introduction

### 3.1.1 Biological and mathematical context

Neurons are cells specialised in the reception, integration and transfer of information in the brain. A propagating electrical signal is transmitted from a neuron to the others in terms of all-or-none emission of action potential also called *spike* which is a stereotyped phenomenon. More precisely, neurons possess a permeable membrane which allows ion exchanges. Without stimulus, the difference of respective ion concentrations induces a voltage gradient called resting potential. This potential evolves depending on the information received from other neurons: a presynaptic neuron emitting a spike leads to the release of neurotransmitters, and induces a change in the ions distribution around the membrane of post-synaptic neurons. If the stimulus reaches a sufficient threshold, the neuron generates an action potential, the *synaptic integration*.

The progress of monitoring methods as MRI (Magnetic Resonance Imaging) and ECG (Electrocardiography) since the 50's led to a better understanding of the physiology of a neuron. As a result, the implementation of mathematical models started with the Hodgkin-Huxley model (Hodgkin & Huxley, 1952) describing the evolution of the membrane potential in terms of a system of four ODEs, further simplified in two equations by FitzHugh (1961) and Nagumo *et al.* (1962).

Stochasticity is intrinsic to the neuronal activity: noise in neuronal systems may come from different sources. To name a few, randomness accounts for the emergence of spontaneous spikes (Fatt & Katz, 1952), failed propagation (Smith, 1980), and the stochastic opening and closing of the ion channels (the probability of the channel being open or closed depends on the membrane potential). Stochasticity is also present at the scale of a whole population in the large variability of synaptic connections between neurons. From a mathematical perspective, this naturally led to diffusion models: mean-field Hodgkin-Huxley and FitzHugh-Nagumo's models in Baladron *et al.* (2012), mean-field Piecewise Deterministic Markov Processes (PDMP) in Masi *et al.* (2014); Cormier *et al.* (2020). Another popular model is the integrate-and-fire dynamics, first introduced in the seminal work of Lapicque (1907), and still studied mathematically, as e.g. in Delarue *et al.* (2015).

The previous type of modeling of the membrane potential typically leads to non-linear Fokker–Planck equations whose large time behavior is often hard to determine analytically. A usual approach in this context (that we follow here) yields more tractable and explicitly solvable models: as spikes are stereotyped, all the information is coded in the duration of time between the spikes. Hence we model the activity of a neuron by a point process where each point represents the time of a spike. In this context, the framework of *Hawkes processes* is particularly relevant

since it can account for the dependence of the activity of a neuron on the past of the whole population: the spike of one neuron can trigger others spikes. Hawkes processes have been first introduced in Hawkes (1971) to model earthquakes, and have been thoroughly studied since (with applications for instance to seismology in Ogata (1988)). It is not possible to quote the vast mathematical literature on Hawkes processes since the seminal works of Hawkes (1971); Hawkes & Oakes (1974); Brémaud & Massoulié (1996), we refer nonetheless to Delattre *et al.* (2016a); Hodara & Löcherbach (2017); Chevallier (2017) and references therein.

In this paper, the main issue we concentrate on is the structure of interaction between neurons. There is indeed experimental evidences that neurons are spatially organized (Bosking *et al.*, 1997; Mountcastle, 1997). The first approach, where this spatial structure is missing, assumes a complete graph of interaction (mean-field framework). Mean-field analysis goes back to McKean (1967); Sznitman (1991), originally for diffusion models as in Baladron *et al.* (2012). The literature on mean-field analysis is huge and does not restrict to neuroscience applications (see the following references as far as neurosciences are concerned: integrate and fire models Delarue *et al.* (2015), PDMP Masi *et al.* (2014); Cormier *et al.* (2020)). As for mean-field Hawkes processes, similar models have been considered in Delattre *et al.* (2016a); Heesen & Stannat (2021); Hodara & Löcherbach (2017) and expanded with additional features (age dependence in Chevallier (2017); Raad *et al.* (2020), inhibition in Costa *et al.* (2020); Duval *et al.* (2022); Raad & Löcherbach (2020)). What makes the mean-field analysis for Hawkes processes particularly tractable is that the large population limit is given in terms of an inhomogeneous Poisson process whose intensity solves a convolution equation (Delattre *et al.*, 2016a).

The spatial organization in the brain has been originally analysed mathematically from a phenomenological perspective: we may refer to the celebrated neural field equation (Wilson & Cowan, 1972; Amari, 1977; Bressloff, 2012), which has given a macroscopic description of excitable units with non-local interaction. Several works have extended the mean-field framework to take into account the presence of a macroscopic spatial structure in the interaction (originally for diffusion models (Touboul, 2014; Luçon & Stannat, 2016; Budhiraja & Wu, 2016), as well as for Hawkes processes (Ditlevsen & Löcherbach, 2017; Chevallier *et al.*, 2019)). More specifically, Chevallier *et al.* (2019) has given a mesoscopic interpretation of the neural field equation in terms of the limit of spatially extended Hawkes processes interacting through a mesoscopic spatial kernel.

The main contribution of this paper is to go further and provide a microscopic interpretation of this spatial structure in terms of random graphs. We assume that the interaction between neurons is given by a possibly inhomogeneous and diluted graph, where the probability of presence of an edge depends on the positions of its vertices. The main example that we have in mind concerns the class of  $W$ -random graph (see Diaconis & Janson (2008); Lovász (2012); Janson (2013); Borgs *et al.* (2018, 2019)), that includes homogeneous Erdős-Rényi graphs. The only previous works so far on particle systems with similar interaction address the case of diffusions. Law of Large Numbers (LLN) and Large Deviations results on homogeneous Erdős-Rényi graphs have been considered in Delattre *et al.* (2016b); Coppini *et al.* (2019); Oliveira & Reis (2019), and further extended to the inhomogeneous case in Bet *et al.* (2020); Luçon (2020); Bayraktar *et al.* (2023); Medvedev

(2013) on a bounded time interval. The behavior of such systems on a time scale no longer bounded (but may depend on the size of the population) is more difficult, and remains largely open so far (in this direction, see Coppini (2022)). The present work is, to the best of our knowledge, the first paper to address similar issues to Hawkes processes. We address here quenched LLN results on bounded time interval and large time asymptotics of the limiting process. The behavior of the system on unbounded time scale is a work in progress. Note also that all the existing works consider graphs with interaction of diverging degrees. The case with sparse interaction (see Oliveira *et al.* (2020); Lacker *et al.* (2023) for diffusions) remains open for Hawkes processes and will be the object of future works.

### 3.1.2 Our model

The aim of this paper is to describe the behavior in large population and large time of a network of particles interacting on a spatially structured random graph. Let  $N$  be the size of the population, consider the multivariate Hawkes process  $(Z_{N,1}(t), \dots, Z_{N,N}(t))_{t>0}$ : for  $i = 1 \dots N$ , the  $i$ th neuron is located on  $x_i \in I$  where  $I \subset \mathbb{R}^d$  represents the spatial domain of the neuron (suppose e.g. that  $I = [0, 1]$  or  $I = \mathbb{R}^d$ ),  $Z_{N,i}(t)$  counts the number of spikes during the time interval  $[0, t]$ . Its intensity at time  $t$  conditioned on the past  $[0, t)$  is given by

$$\lambda_{N,i}(t) = f \left( v_t(x_i) + \frac{1}{N} \sum_{j=1}^N w_{ij}^{(N)} \int_0^{t-} h(t-s) dZ_{N,j}(s) \right). \quad (3.1.1)$$

Here,  $f : \mathbb{R} \rightarrow \mathbb{R}_+$  represents the synaptic integration,  $v_t : \times I \rightarrow \mathbb{R}$  a spontaneous activity of the neuron at time  $t$ ,  $h : \mathbb{R}_+ \rightarrow \mathbb{R}$  a memory function which models how a past jump of the system affects the present intensity. The novelty here is  $w_{ij}^{(N)}$ , representing the random inhomogeneous interaction between the neurons  $i$  and  $j$  that depends on their positions  $x_i$  and  $x_j$ . We refer to Section 3.2 for precise definitions.

We study the behavior of the process  $(Z_{N,1}(t), \dots, Z_{N,N}(t))_{t>0}$  as  $N \rightarrow \infty$  and  $t \rightarrow \infty$ . The large population convergence is considered for a fixed realization of the graph (quenched model). Its limit is described in terms of an inhomogeneous Poisson process whose intensity involves the macroscopic spatial structure of the graph. A second aspect of the present work of independent interest will be to analyse the long time dynamics of the macroscopic process. We generalise the phase transition already observed for mean-field linear Hawkes processes in Delattre *et al.* (2016a). An important issue will be to understand how the inhomogeneity of the graph influences the long time dynamics. This will be illustrated by different examples and simulations.

### 3.1.3 Organisation of the paper

After introducing some notation, we start in Section 3.2 by defining formally the process of interest (3.2.2). The well-posedness of such process is treated by Proposition 3.2.5. We study the large population behavior of the Hawkes process  $(Z_{N,1}(t), \dots, Z_{N,N}(t))_{t>0}$  in Section 3.2.3. We show, under suitable hypotheses on



the parameters, that the behavior of a neuron located in  $x \in I$  within an infinite population is described by an intensity  $\lambda_t(x)$  solving

$$\lambda_t(x) = f \left( v_t(x) + \int_I W(x, y) \int_0^t h(t-s) \lambda_s(y) ds \nu(dy) \right). \quad (3.1.2)$$

Here,  $W : I \times I \rightarrow \mathbb{R}_+$  is seen as the limit interaction kernel, and  $\nu$ , probability measure on  $I$  describes the macroscopic distribution of the positions. Well-posedness and regularity of (3.1.2) is considered in Theorem 3.2.7. In Section 3.3, we study the behavior of the process (3.2.2) in large population (Theorems 3.3.10 and 3.3.12). The behavior of the empirical measure and respectively the spatial profile (Definition 3.3.16) is analysed in Section 3.3.4 (resp. Section 3.3.5). In Section 3.4, we study the behavior of (3.1.2) as  $t \rightarrow \infty$  in the linear case, that is when  $f(u) = u$ . We extend the phase transition observed without spatial structure in Delattre *et al.* (2016a) to a general interaction kernel  $W$ . Finally in Section 3.5, we apply our results to concrete cases and present some simulations. The proofs are gathered in the remaining Sections.

## 3.2 A system of $N$ interacting particles on a graph and its limit

### 3.2.1 Notation

For  $n \in \mathbb{N}$ , we write  $\|\cdot\|$  for the usual Euclidian norm in  $\mathbb{R}^n$ ,  $\|(x_1, \dots, x_n)\| = (|x_1|^2 + \dots + |x_n|^2)^{\frac{1}{2}}$ . For  $(E, \mathcal{A}, \mu)$  a measured space, for a function  $h$  in  $L^p(E, \mu)$  with  $p \geq 1$ , we write  $\|h\|_{E, \mu, p} := \left( \int_E |h|^p d\mu \right)^{\frac{1}{p}}$ . When  $p = 2$ , we write as  $\langle f, g \rangle_{E, \mu} = \int_E fg d\mu$  the scalar product. Without ambiguity, we may omit the subscript  $(E, \mu)$  or  $\mu$ . For instance, for  $T > 0$  and  $h$  in  $L^p([0, T])$ , we write  $\|h\|_{[0, T], p} := \left( \int_0^T |h(t)|^p dt \right)^{\frac{1}{p}}$ . When we omit the notation  $[0, T]$ , the integration is on  $\mathbb{R}_+$ . For a real-valued bounded function  $g$  on a space  $E$ , we write  $\|g\|_\infty := \|g\|_{E, \infty} = \sup_{x \in E} |g(x)|$ . If  $d$  is a distance on  $E$ , we denote by  $\|f\|_L = \sup_{x \neq y} |f(x) - f(y)|/d(x, y)$  the Lipchitz seminorm of a real-valued function  $f$  on  $E$ . We also denote by  $\|f\|_{BL} := \|f\|_L + \|f\|_{E, \infty}$  the bounded Lipschitz norm of  $f$ . For  $\mu$  and  $\nu$  measures on  $E$ , we define

$$d_{BL}(\mu, \nu) := \sup_{g, \|g\|_{BL} \leq 1} \left| \int_E g(d\mu - d\nu) \right|. \quad (3.2.1)$$

We denote by  $\mathbb{D}([0, T], \mathbb{N})$  the space of càdlàg (right continuous with left limits) functions defined on  $[0, T]$  and taking values in  $\mathbb{N}$ . For any integer  $N \geq 1$ , we denote by  $\llbracket 1, N \rrbracket$  the set  $\{1, \dots, N\}$ . For any distribution  $\nu$ ,  $X \sim \nu$  means that the random variable  $X$  has distribution  $\nu$ . We denote by  $\mathcal{U}(0, 1)$  the uniform distribution on  $[0, 1]$ , and for any  $p \in [0, 1]$ ,  $\mathcal{B}(p)$  denotes the Bernoulli distribution with parameter  $p$ .

### 3.2.2 The model

#### 3.2.2.1 Definitions

The graph of interaction for (3.1.1) is constructed as follows:

**Definition 3.2.1.** On a common probability space  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mathbb{P})$ , we consider a sequence  $\left( \left( x_i^{(N)} \right)_{i \in \llbracket 1, N \rrbracket} \right)_{N \geq 1}$  of (possibly random) positions and a family of random variables  $\xi^{(N)} = \left( \xi_{ij}^{(N)} \right)_{N \geq 1, i, j \in \llbracket 1, N \rrbracket}$  on  $\tilde{\Omega}$  such that under  $\mathbb{P}$ , for any  $N \geq 1$  and  $i, j \in \llbracket 1, N \rrbracket$ , conditioned on the positions  $(x_1^{(N)}, \dots, x_N^{(N)})$ ,  $\xi^{(N)}$  is a collection of mutually independent Bernoulli random variables such that for  $1 \leq i, j \leq N$ ,  $\xi_{ij}^{(N)}$  has parameter  $W_N(x_i, x_j)$ . We assume that the particles in (3.1.1) are connected according to the oriented graph  $\mathcal{G}^{(N)} = (\{1, \dots, N\}, \xi^{(N)})$ . For any  $i$  and  $j$ ,  $\xi_{ij}^{(N)} = 1$  encodes for the presence of the edge  $j \rightarrow i$  and  $\xi_{ij}^{(N)} = 0$  for its absence.

It is possible to construct via a coupling this graph simultaneously for all  $N$ : consider an infinite sequence of fixed positions in  $I (x_1, \dots, x_N, \dots)$  (that is, for each  $N \geq 1$ ,  $x_i^{(N)} = x_i$ ) and i.i.d. random variables  $(U_{i,j})_{i,j \in \mathbb{N}} \sim \mathcal{U}[0, 1]$ . Define  $\xi_{ij}^{(N)} = \mathbf{1}_{\{U_{i,j} \leq W_N(x_i, x_j)\}}$ : conditioned on the positions  $(x_1, \dots, x_N)$ ,  $\xi^{(N)}$  is a collection of independent variables and  $\xi_{ij}^{(N)} \sim \mathcal{B}(W_N(x_i, x_j))$ . We now fix these sequences, and work on a filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbf{P})$  rich enough for all the following processes can be defined. We denote by  $\mathbf{E}$  the expectation under  $\mathbf{P}$  and  $\mathbb{E}$  the expectation with respect to  $\mathbb{P}$ . In the following definitions,  $N$  is fixed and we denote by  $\underline{x}^{(N)} = (x_1^{(N)}, \dots, x_N^{(N)})$  the vector of positions.

**Definition 3.2.2.** Let  $(\pi_i(ds, dz))_{1 \leq i \leq N}$  be a sequence of i.i.d. Poisson random measures on  $\mathbb{R}_+ \times \mathbb{R}_+$  with intensity measure  $dsdz$ . A  $(\mathcal{F}_t)$ -adapted multivariate counting process  $(Z_{N,1}(t), \dots, Z_{N,N}(t))_{t \geq 0}$  defined on  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbf{P})$  is called a *multivariate Hawkes process* with the set of parameters  $(N, f, \xi^{(N)}, W_N, v, h, \underline{x}^{(N)})$  if  $\mathbf{P}$ -almost surely, for all  $t \geq 0$  and  $i \in \llbracket 1, N \rrbracket$ :

$$Z_{N,i}(t) = \int_0^t \int_0^\infty \mathbf{1}_{\{z \leq \lambda_{N,i}(s)\}} \pi_i(ds, dz) \quad (3.2.2)$$

with  $\lambda_{N,i}(t)$  defined by

$$\lambda_{N,i}(t) = f \left( v_t(x_i^{(N)}) + \frac{\kappa_i^{(N)}}{N} \sum_{j=1}^N \xi_{ij}^{(N)} \int_0^{t-} h(t-s) dZ_{N,j}(s) \right). \quad (3.2.3)$$

We denote by  $\kappa_i^{(N)} \geq 0$  a dilution parameter which may depend on  $\underline{x}^{(N)}$ , and  $\xi^{(N)}$ . The idea behind this dilution parameter is that  $\kappa_i^{(N)} \simeq \frac{N}{\mathbb{E}[\deg_N(i)]}$  (where  $\deg_N(i) = \sum_{j=1}^N \xi_{ij}^{(N)}$  is the indegree of the particle  $i$ , that is, the number of edges incident to it), so that the interaction term remains of order 1 as  $N \rightarrow \infty$ . This means that the interaction in (3.1.1) is fixed as  $w_{ij}^{(N)} = \kappa_i^{(N)} \xi_{ij}^{(N)}$ .

**Remark 3.2.3.** By (Delattre *et al.*, 2016a, Proposition 3), the Hawkes process  $(Z_{N,1}(t), \dots, Z_{N,N}(t))_{t \geq 0}$  defined by (3.2.2) is such that  $\mathbf{P}$ -almost surely, for all  $i \neq j$  the processes  $Z_{N,i}$  and  $Z_{N,j}$  do not jump simultaneously and for all  $i \in \llbracket 1, N \rrbracket$ , the compensator of  $Z_{N,i}(t)$  is  $\int_0^t \lambda_{N,i}(s) ds$  (see Jacod & Shiryaev (2013) about compensators of increasing processes).

### 3.2.2.2 Existence

We first provide well-posedness results of  $(Z_{N,1}, \dots, Z_{N,N})$  given by (3.2.2). We require the following assumptions:

**Hypothesis 3.2.4.** We suppose that  $f$  is Lipschitz continuous with Lipschitz constant  $L_f \geq 0$ , and that either  $f$  is nonnegative or that  $f(x) = x$  with  $v_t \geq 0$  for all  $t \geq 0$  and  $h \geq 0$  (linear case). We also suppose that  $h$  is locally square integrable on  $[0, +\infty)$ , that  $(t, x) \mapsto v_t(x)$  is continuous in  $t$  and Lipschitz continuous in  $x$  (uniformly in  $t$ ) with Lipschitz constant  $L_v \geq 0$ . Moreover  $v$  is supposed bounded uniformly in  $(t, x)$  i.e.  $\|v\|_\infty := \sup_{t \geq 0} \|v_t\|_\infty < \infty$ .

**Proposition 3.2.5.** *Under Hypothesis 3.2.4, for a fixed realization of the family  $(\pi_i)_{1 \leq i \leq N}$ , there exists a pathwise unique multivariate Hawkes process (in the sense of Definition 3.2.2) such that for any  $T < \infty$ ,  $\sup_{t \in [0, T]} \sup_{1 \leq i \leq N} \mathbf{E}[Z_{N,i}(t)] < \infty$ .*

The proof of Proposition 3.2.5 will be given in Section 3.7.1.

## 3.2.3 Large population limit process

We want to study the behavior of the process defined in Definition 3.2.2 when  $N \rightarrow \infty$  on bounded time interval. After some heuristics, we show the well-posedness of the limit of the system 3.2.2.

### 3.2.3.1 Heuristics

In this paragraph, we motivate the proper limit for the particle system (3.2.2) as  $N \rightarrow \infty$ . A minimal requirement is that the empirical distribution of the positions  $\nu^{(N)} := \frac{1}{N} \sum_{i=1}^N \delta_{x_i^{(N)}}$  has itself a macroscopic limit  $\nu$ . We will consider below different scenarios under which such LLN holds. Concerning the macroscopic behavior of the graph, another minimal requirement is that in a way to define later on, the graph  $\mathcal{G}^{(N)}$  given in Definition 3.2.2 converges towards a macroscopic interaction kernel  $W : I \times I \rightarrow \mathbb{R}_+$ . We refer to Section 3.3.2 for more precise statements. Then, as  $N \rightarrow \infty$ , an informal LLN argument shows that the empirical mean in (3.2.3) becomes an expectation w.r.t both the candidate limit for  $Z_{N,i}$  and w.r.t the macroscopic law  $\nu$  of the positions: we can replace the sum in (3.1.1) by the integral in (3.1.2), the microscopic interaction term  $w_{ij}^{(N)}$  in (3.1.1) by the macroscopic term  $W(x, y)$  in (3.1.2) (where  $y$  describes the macroscopic distribution of the positions), and the past activity of the neuron  $dZ_{N,j}(s)$  by its intensity in large population. Hence, the macroscopic description of a neuron at position  $x \in I$  should be described in terms of its intensity  $\lambda_t(x)$  solving (3.1.2). This heuristics

gives a limit process at position  $x$  defined as an inhomogeneous Poisson point process with deterministic intensity  $(\lambda_t)_{t \geq 0}$  satisfying (3.1.2).

### 3.2.3.2 Well-posedness of the macroscopic limit

We propose a framework under which (3.1.2) is well-posed, with more hypotheses on the regularity of  $(f, v, W)$ .

**Hypothesis 3.2.6.** Assume that the macroscopic indegree at position  $x$  defined by

$$D(x) = \int_I W(x, y) \nu(dy) \quad (3.2.4)$$

has a Hölder regularity and is uniformly bounded on  $I$ : there exist  $C_w > 0$  and  $\vartheta \in ]0, 1]$  such that

$$\int_I |W(x, y) - W(x', y)| \nu(dy) \leq C_w \|x - x'\|^\vartheta, \quad x, x' \in I \quad \text{and} \quad (3.2.5)$$

$$\sup_{x \in I} D(x) =: C_W^{(1)} < \infty. \quad (3.2.6)$$

**Theorem 3.2.7.** *Let  $T > 0$ . Under Hypotheses 3.2.4 and 3.2.6, there exists a unique solution  $(t, x) \mapsto \lambda_t(x)$  to (3.1.2) that is continuous and bounded on  $[0, T] \times I$  and this solution is nonnegative. Moreover, there exists  $C_\lambda > 0$  depending on  $(f, v, W, h, \nu, T)$  such that for all  $(t, x, z) \in [0, T] \times I \times I$ ,*

$$|\lambda_t(x) - \lambda_t(z)| \leq C_\lambda (\|x - z\| + \|x - z\|^\vartheta) =: C_\lambda \phi(\|x - z\|). \quad (3.2.7)$$

*In the linear case  $f(x) = x$ ,  $h \geq 0$ ,  $v_t \geq 0$  for all  $t \geq 0$ , if  $v$  is continuously differentiable in time and  $(t, x) \mapsto \frac{\partial v_t}{\partial t}(x)$  is bounded on  $[0, T] \times I$ ,  $h$  is continuous and piecewise continuously differentiable, then  $\lambda$  is differentiable in time and*

$$\frac{\partial \lambda_t}{\partial t}(x) = \frac{\partial v_t}{\partial t}(x) + h(t) \int_I W(x, y) \lambda_0(y) \nu(dy) + \int_I \int_0^t h(t-s) W(x, y) \frac{\partial \lambda_s}{\partial s}(y) \nu(dy) ds, \quad (3.2.8)$$

*and  $(t, x) \mapsto \left| \frac{\partial \lambda_t}{\partial t}(x) \right|$  is bounded on  $[0, T] \times I$ .*

Theorem 3.2.7 will be proved in Section 3.7.2. Note that Theorem 3.2.7 provides the existence of a unique solution  $\lambda$  of (3.1.2) that is continuous on  $\mathbb{R}_+ \times I$  and locally bounded.

## 3.3 Convergence of the model in large population

### 3.3.1 Coupling

From now on,  $\lambda$  refers to the unique solution to (3.1.2). To check that our heuristics about the large population behavior is correct, we introduce a suitable coupling between the process defined in (3.2.2) (at positions  $x_i$ ) and a Poisson process with intensity  $\lambda(x_i)$  at the same position  $x_i$ .

**Definition 3.3.1.** For the family  $(\pi_i(ds, dz))_{1 \leq i \leq N}$  of i.i.d. PRM on  $\mathbb{R}_+ \times \mathbb{R}_+$  from Definition 3.2.2, we construct for all  $i$  in  $\llbracket 1, N \rrbracket$ :

$$\bar{Z}_i(t) = \int_0^t \int_0^\infty \mathbf{1}_{\{z \leq \lambda_s(x_i)\}} \pi_i(ds, dz) \quad (3.3.1)$$

with  $\lambda$  satisfying (3.1.2). Each process  $\bar{Z}_i$  is an inhomogenous Poisson process with (deterministic) intensity  $(\lambda_t(x_i))_t$ , and as the family  $(\pi_i)$  is independent, the processes  $(\bar{Z}_i)_{i=1, \dots, N}$  are also independent.

### 3.3.2 Hypotheses

Regarding the behavior of the graph when  $N \rightarrow \infty$ , we use here the formalism of graph convergence developed in Lovász (2012) and introduce different norms on  $I^2$ . The key notion is to represent graphs in term of graphons, that are positive kernels defined on  $I^2$ . Note that we will not necessarily restrict ourselves to the symmetric case and bounded graphons.

**Definition 3.3.2.** Let  $W$  be a  $\mathbb{R}$ -valued function defined on  $I \times I$ , where  $I$  is endowed with some probability measure  $\nu$ . When the following terms are correctly defined, we write:

$$\|W\|_{\square, \nu} := \sup_{S, T \subset I} \left| \int_{S \times T} W(x, y) \nu(dx) \nu(dy) \right|, \quad (3.3.2)$$

$$\|W\|_{\infty \rightarrow 1, \nu} := \sup_{\|g\|_\infty \leq 1} \int_I \left| \int_I W(x, y) g(y) \nu(dy) \right| \nu(dx), \quad (3.3.3)$$

$$\|W\|_{\infty \rightarrow \infty, \nu} := \sup_{\|g\|_\infty \leq 1} \sup_{x \in I} \left| \int_I W(x, y) g(y) \nu(dy) \right|. \quad (3.3.4)$$

These norms go back to the formalism of graph convergence introduced in Lovász (2012); Diaconis & Janson (2008) and further developed in Borgs *et al.* (2011, 2018, 2019) (and references therein). The last two norms can be seen as the norms of the linear operator  $T_W : g \mapsto (x \mapsto \int_I W(x, y) g(y) \nu(dy))$  when considering respectively  $T_W : L^\infty(I, \nu) \rightarrow L^1(I, \nu)$  and  $T_W : L^\infty(I, \nu) \rightarrow L^\infty(I, \nu)$ . We also define the *cut-distance* between two functions by

$$d_{\square, \nu}(W_1, W_2) = \|W_1 - W_2\|_{\square, \nu}. \quad (3.3.5)$$

**Remark 3.3.3.** Lemma 8.11 of Lovász (2012) gives that  $\|\cdot\|_{\square, \nu}$  and  $\|\cdot\|_{\infty \rightarrow 1, \nu}$  are equivalent: if  $W$  is a function defined on  $I^2$  with values in  $\mathbb{R}$ , then

$$\|W\|_{\square, \nu} \leq \|W\|_{\infty \rightarrow 1, \nu} \leq 4\|W\|_{\square, \nu}. \quad (3.3.6)$$

As  $\|W\|_{1, \nu} := \int_{I^2} |W(x, y)| \nu(dx) \nu(dy)$ , we always have  $\|W\|_{\square, \nu} \leq \|W\|_{1, \nu}$ .

Usual representations of graphons consist in taking  $I = [0, 1]$  endowed with Lebesgue measure. We extend this definition to the general case where  $\nu$  is a probability measure on  $I$ . To do this, we require the following assumption for the whole article.

**Hypothesis 3.3.4.** The probability measure  $\nu$  is absolutely continuous with respect to Lebesgue measure on  $\mathbb{R}^d$ .

**Lemma 3.3.5.** *Under Hypothesis 3.3.4, for any  $N \geq 1$ , there exists a partition  $\mathcal{P}_N := (B_{N,i})_{i=1,\dots,N}$  of  $I$  (and we use the notation  $I = \bigsqcup_{i=1}^N B_{N,i}$ ) such that for all  $i = 1, \dots, N$ ,  $\nu(B_{N,i}) = \frac{1}{N}$ .*

Without ambiguity, we will forget the upper index  $(N)$  and only write the partition  $(B_1, \dots, B_N)$ .

*Proof.* Denote by  $\nu^{(1)}$  the first marginal of  $\nu$  that is absolutely continuous with respect to Lebesgue measure on  $\mathbb{R}$ . Let  $F_1$  be its continuous probability distribution function. Then for all  $i = 1, \dots, N$ , defining  $B_i := (F_1^{-1}(\left(\frac{i-1}{N}, \frac{i}{N}\right]) \times \mathbb{R}^{d-1}) \cap I$  gives the result.  $\square$

For every weighted graph  $\mathcal{G}$  with weights  $(g_{ij})$ , we associate a step-function  $W^{\mathcal{G}}$  constructed, upon this partition, as follows (see e.g. Lovász (2012); Borgs *et al.* (2011)):

$$W^{\mathcal{G}}(u, v) = \sum_{i=1}^N \sum_{j=1}^N g_{ij} \mathbf{1}_{B_i}(u) \mathbf{1}_{B_j}(v), \quad (u, v) \in I^2. \quad (3.3.7)$$

**Definition 3.3.6.** We denote by  $\mathcal{G}_N^{(1)}$  the directed weighted graph with vertices  $\{1, \dots, N\}$  such that every edge  $j \rightarrow i$  is present, and with weight  $\kappa_i^{(N)} W_N(x_i, x_j)$ .

Here  $\mathcal{G}_N^{(1)}$  represents the average version of the graph  $\mathcal{G}^{(N)}$  (where the interaction  $\xi_{ij} \sim \mathcal{B}(W_N(x_i, x_j))$  has been replaced by  $\mathbb{E}(\xi_{ij})$ ), renormalized by the dilution coefficient  $\kappa_i^{(N)}$ . A key argument of Theorems 3.3.10 and 3.3.12 will be to show that  $\mathcal{G}^{(N)}$  and  $\mathcal{G}_N^{(1)}$  are close as  $N \rightarrow \infty$  through concentration arguments that require the following uniformity assumptions on  $W_N$ .

**Hypothesis 3.3.7.** We suppose that there exist  $\kappa_N \geq 1$  and  $w_N \in ]0, 1]$  such that:

$$\max_{i \in \llbracket 1, N \rrbracket} \left( \kappa_i^{(N)} \right) \leq \kappa_N, \quad (3.3.8)$$

$$\max_{i, j \in \llbracket 1, N \rrbracket} (W_N(x_i, x_j)) \leq w_N, \quad (3.3.9)$$

$$\frac{1}{\kappa_N} \leq w_N \leq 1, \quad \text{and asymptotically:} \quad (3.3.10)$$

$$\kappa_N^2 w_N \underset{N \rightarrow \infty}{=} o\left(\frac{N}{\log(N)}\right) \quad \text{and} \quad \frac{\kappa_N}{N} \underset{N \rightarrow \infty}{\rightarrow} 0. \quad (3.3.11)$$

We also suppose that there exists  $C_W > 0$  independent of  $N$  such that

$$\sup_{i \in \llbracket 1, N \rrbracket} \frac{1}{N} \sum_{j=1}^N \kappa_i^{(N)} W_N(x_i, x_j) \leq C_W. \quad (3.3.12)$$

To illustrate the above conditions, think of the case where  $W_N = \rho_N$  is a constant with  $\rho_N \xrightarrow[N \rightarrow \infty]{} 0$ . This corresponds to a diluted Erdős-Rényi random graph. In this case, we can take  $w_N = \rho_N$  and  $\kappa_i^{(N)} = \kappa_N = 1/\rho_N$ . Then (3.3.11) boils down to  $\rho_N \gg \log(N)/N$ . Inequality (3.3.12) is the microscopic counterpart of (3.2.6): we require that the weighted indegrees of vertices in  $\mathcal{G}_N^{(1)}$  are uniformly bounded.

### 3.3.3 Convergence

We study the proximity between the particle systems (3.2.2) and its macroscopic limit (3.3.1). We show two theorems that require different sets of hypotheses on the parameter functions, under two main scenarios.

**Definition 3.3.8.** We consider two different frameworks for the choices of the positions:

1. **Random spatial distribution:** For  $(\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_N, \dots)$  a random sequence of i.i.d. variables distributed according to  $\nu$  on  $I$ , we set for all  $N \geq 1$   $\underline{x}^{(N)} = (x_1^{(N)}, \dots, x_N^{(N)})$  as the lexicographic ordering of the  $N$  first positions  $(\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_N)$ . We assume that  $\|W\|_{L^\chi(I^2, \nu \otimes \nu)} < \infty$  for some  $\chi > 5$ .
2. **Deterministic regular distribution of the positions:** For every  $N \geq 1$  and  $1 \leq i \leq N$ , we set  $x_i^{(N)} = i/N$  and  $I = [0, 1]$  endowed with  $\nu(dx) = dx$ . We assume that  $W$  is piecewise continuous on  $[0, 1]^2$ .

The assumption  $\chi > 5$  of Scenario (1) is required in Proposition 3.8.6, as a sufficient hypothesis for a Borel-Cantelli argument.

#### 3.3.3.1 First case: convergence in average

**Hypothesis 3.3.9.** We suppose that the annealed graph  $\mathcal{G}_N^{(1)}$  converges to  $W$  for the cut-distance:

$$d_{\square, \nu} \left( W^{\mathcal{G}_N^{(1)}}, W \right) \xrightarrow[N \rightarrow \infty]{} 0, \text{ as well as} \quad (3.3.13)$$

$$\sup_{j \in [1, N]} \frac{1}{N} \sum_{i=1}^N \kappa_i^{(N)} W_N(x_i, x_j) \leq C_W. \quad (3.3.14)$$

Note that (3.3.13) implies that  $\lim_{N \rightarrow \infty} \|W^{\mathcal{G}_N^{(1)}} - W\|_{\infty \rightarrow 1, \nu} = 0$  (see Remark 3.3.3). The hypothesis (3.3.14) differs from (3.3.12) in the sense that (3.3.14) asks for a uniform bound on the outdegree (that is, the number of tail ends adjacent to a vertex) whereas (3.3.12) relates to a uniform bound on the indegree.

**Theorem 3.3.10.** *Let  $T > 0$ . Suppose that the sequence of positions  $(\underline{x}^{(N)})_N$  satisfies one of the scenarios of Definition 3.3.8. Then, under the set of Hypotheses 3.2.4, 3.2.6, 3.3.4, 3.3.7, and 3.3.9, we have*

$$\frac{1}{N} \sum_{i=1}^N \mathbf{E} \left[ \sup_{t \in [0, T]} |Z_{N,i}(t) - \bar{Z}_i(t)| \right] \xrightarrow[N \rightarrow \infty]{} 0 \quad (3.3.15)$$

for  $\mathbb{P}$ -almost realizations of the connectivity sequence and positions  $(\xi^{(N)}, \underline{x}^{(N)})_{N \geq 1}$ .

The proof of Theorem 3.3.10 will be given in Section 3.8.2.

### 3.3.3.2 Second case: convergence of the supremum

Some graphs do not satisfy (3.3.14), see the examples of Section 3.5.1.2. We propose here another result of convergence that does not require the control (3.3.14), but ask in return for a stronger convergence of the graphons.

**Hypothesis 3.3.11.** We suppose that

$$\|W^{\mathcal{G}_N^{(1)}} - W\|_{\infty \rightarrow \infty, \nu} \xrightarrow{N \rightarrow \infty} 0. \quad (3.3.16)$$

**Theorem 3.3.12.** *Let  $T > 0$ . Suppose that the sequence of positions  $(\underline{x}^{(N)})_N$  satisfies one of the scenarios of Definition 3.3.8. Consider the coupling introduced in Definition 3.3.1. Then, under the set of Hypotheses 3.2.4, 3.2.6, 3.3.4, 3.3.7 and 3.3.11, we have*

$$\max_{1 \leq i \leq N} \mathbf{E} \left[ \sup_{t \in [0, T]} |Z_{N,i}(t) - \bar{Z}_i(t)| \right] \xrightarrow{N \rightarrow \infty} 0. \quad (3.3.17)$$

$\mathbb{P}$ -almost surely.

The proof of Theorem 3.3.12 will be given in Section 3.8.3.

**Remark 3.3.13.** Theorems 3.3.10 and 3.3.12 are quenched results, and do not provide any speed of convergence. In this case, the speed of convergence is unknown. Nevertheless, if we integrate also with respect to the graph (annealed case), one can obtain explicit speed of convergence as follows:

$$\begin{aligned} \max_{1 \leq i \leq N} \mathbf{E} \mathbf{E} \left[ \sup_{t \in [0, T]} |Z_{N,i}(t) - \bar{Z}_i(t)| \right] &\leq C_T \frac{\kappa_N \sqrt{w_N}}{\sqrt{N}}, \\ \frac{1}{N} \sum_{i=1}^N \mathbf{E} \mathbf{E} \left[ \sup_{t \in [0, T]} |Z_{N,i}(t) - \bar{Z}_i(t)| \right] &\leq C_T \frac{\kappa_N \sqrt{w_N}}{\sqrt{N}}. \end{aligned}$$

Working in the annealed case simplifies considerably the proof (left to the reader), the previous estimates can be easily derived from the calculation done in the proofs of the previous theorems: we no longer have to deal with concentration estimates (see the term  $A_{i,t,3}^{(N)}$  below in (3.8.12) which becomes a simple variance term).

### 3.3.4 Consequence on the empirical measure

A direct consequence of Theorems 3.3.10 and 3.3.12 concerns the behavior as  $N \rightarrow \infty$  of the empirical distribution on the space  $S := \mathbb{D}([0, T], \mathbb{N}) \times I$  of trajectories and positions.



**Definition 3.3.14.** We define the following probability measures on  $S$ :

$$\mu_N(d\eta, dx) := \frac{1}{N} \sum_{i=1}^N \delta_{(Z_{N,i}([0,T]), x_i^{(N)})}(d\eta, dx), \text{ and} \quad (3.3.18)$$

$$\mu_\infty(d\eta, dx) := P_{[0,T],\infty}(d\eta|x) \nu(dx), \quad (3.3.19)$$

where  $P_{[0,T],\infty}(\cdot|x)$  is the law of an inhomogeneous Poisson point process with intensity  $(\lambda_t(x))_{0 \leq t \leq T}$  (solution of (3.1.2)). Note that  $\mu_N$  is random.

**Theorem 3.3.15.** *Under the assumptions of Theorem 3.3.10 or Theorem 3.3.12, for  $\mathbb{P}$ -almost realizations of the connectivity sequence and positions  $(\xi^{(N)}, \underline{x}^{(N)})_{N \geq 1}$  we have*

$$\mathbf{E} [d_{BL}(\mu_N, \mu_\infty)] \xrightarrow[N \rightarrow \infty]{} 0, \quad (3.3.20)$$

where  $d_{BL}$  is the bounded Lipschitz distance introduced in (3.2.1).

The proof of Theorem 3.3.15 will be given in Section 3.9.1. We can see this result as an extension of Theorems 1 and 2 of Chevallier *et al.* (2019), where the memory function is an exponential kernel and the interaction comes from a fixed interaction kernel that depends on the positions.

### 3.3.5 Spatial profile

Here we are under the conditions of Scenario (2) of Definition 3.3.8, where  $\lambda$  solves (3.1.2).

**Definition 3.3.16.** Define the random profile

$$U_N(t, x) := \sum_{i=1}^N U_{N,i}(t) \mathbf{1}_{x \in (\frac{i-1}{N}, \frac{i}{N}]}, \text{ where} \quad (3.3.21)$$

$$U_{N,i}(t) := v_t(x_i) + \frac{\kappa_i^{(N)}}{N} \sum_{j=1}^N \xi_{ij}^{(N)} \int_0^t h(t-s) dZ_{N,j}(s), \quad (3.3.22)$$

and the deterministic profile

$$u_t(x) := v_t(x) + \int_I W(x, y) \int_0^t h(t-s) \lambda_s(y) ds dy. \quad (3.3.23)$$

We see from Theorem 3.2.7 that  $u$  is continuous and bounded on  $[0, T] \times IS$ .

Note that  $\lambda_{N,i}(t) = f(U_{N,i}(t-))$  and that  $U_{N,i}(t)$  describes the accumulated activity of neuron  $i$  up to time  $t$ . A similar quantity has already been considered in Chevallier *et al.* (2019) for  $h(t) = e^{-\alpha t}$  with a deterministic graph of interaction. In this case, with  $v_t(x) = e^{-\alpha t} \rho(x)$ , (3.3.23) is the solution of the scalar neural field equation

$$\partial_t u_t(x) = -\alpha u_t(x) + \int_I W(x, y) f(u_t(y)) dy.$$

It has been extensively studied in the literature as an important example of macroscopic structured model with non local interaction (see Amari (1977); Wilson & Cowan (1972); Bressloff (2012)).

**Proposition 3.3.17.** *Under the Hypotheses of Theorem 3.3.10,*

$$\mathbf{E} \left[ \int_0^T \int_0^1 |U_N(t, x) - u_t(x)| dx dt \right] \xrightarrow{N \rightarrow \infty} 0, \quad (3.3.24)$$

for  $\mathbb{P}$ -almost realizations of the connectivity sequence and positions  $(\xi^{(N)}, \underline{x}^{(N)})_{N \geq 1}$ .

The proof of Proposition 3.3.17 will be given in Section 3.9.2.

## 3.4 Large time behavior of the limit process in the linear case

We want to see how the limiting intensity (3.1.2) behaves as  $t \rightarrow \infty$ . We restrict here to the following *linear case*, that is, when  $f(x) = x$ :

$$\lambda_t(x) = v_t(x) + \int_I W(x, y) \int_0^t h(t-s) \lambda_s(y) ds \nu(dy), \quad (3.4.1)$$

for  $(t, x) \in \mathbb{R}_+ \times I$ . The case without spatial interaction, that is,  $\lambda_t = \mu + \int_0^t h(t-s) \lambda_s ds$  with  $\mu \geq 0$  is standard and has been thoroughly studied in (Delattre *et al.*, 2016a, Th. 10 and 11). Depending on the value of  $\|h\|_1$ , there is a phase transition in the behavior of such  $\lambda$  when  $t \rightarrow \infty$ : in the subcritical case ( $\|h\|_1 < 1$ ),  $\lambda_t \xrightarrow{t \rightarrow \infty} \frac{\mu}{1 - \|h\|_1}$  and in the supercritical case ( $\|h\|_1 > 1$ ),  $\lambda_t \xrightarrow{t \rightarrow \infty} \infty$ . The point of the present paragraph is to extend this result to the spatial case. We require the following assumptions:

**Hypothesis 3.4.1.** Suppose that we are in the linear case of Hypothesis 3.2.4. In addition to Hypotheses 3.2.4 and 3.2.6, we suppose that  $h$  is in  $L^1(\mathbb{R}^+)$  and piecewise continuously differentiable. We also suppose that  $t \mapsto v_t$  is continuously differentiable and that there exists  $C_v > 0$  such that

$$\sup_{x \in I} \left\| t \mapsto \frac{\partial v_t(x)}{\partial t} \right\|_1 = \sup_{x \in I} \int_{\mathbb{R}^+} \left| \frac{\partial v_s(x)}{\partial s} \right| ds \leq C_v < \infty. \quad (3.4.2)$$

We also suppose that there exists  $v_\infty$  Lipschitz continuous on  $I$  such that  $\limsup_{N \rightarrow \infty} \sup_{x \in I} |v_t(x) - v_\infty(x)| = 0$ . Hence, when it does not depend on time, we simply suppose  $v_t = v = v_\infty$ .

To describe the phase transition, we introduce the following linear operator

$$\begin{aligned} T_W : L^\infty(I) &\longrightarrow L^\infty(I) \\ g &\longmapsto (T_W g : x \longmapsto \int_I W(x, y) g(y) \nu(dy)). \end{aligned} \quad (3.4.3)$$

The continuity of  $T_W$  follows directly from (3.2.6), and we have  $\|T_W\| \leq C_W^{(1)}$ . We denote by  $r_\infty(T_W)$  the spectral radius of  $T_W$ :

$$r_\infty := r_\infty(T_W) = \sup_{\sigma \in Sp(T_W)} |\sigma| = \lim_{n \rightarrow \infty} \|T_W^n\|^{\frac{1}{n}}. \quad (3.4.4)$$

The phase transition is given in terms of  $\|h\|_1 r_\infty < 1$  (subcritical) and  $\|h\|_1 r_\infty > 1$  (supercritical). The two cases are described separately below, after dealing with the usual exponential case.

### 3.4.1 The exponential case

Previous works (Chevallier *et al.*, 2019) have considered  $h(t) = e^{-\alpha t}$  with  $\alpha > 0$  (hence  $\|h\|_1 = 1/\alpha$ ). The term  $\alpha$  is then called the *leakage rate*. Note that in this case, the dynamics becomes Markovian (Dion *et al.*, 2021). At the large population limit, the spatial profile seen in Section 3.3.5 is in this case linked to the scalar neural field equation (Chevallier *et al.*, 2019). In the exponential case, with the introduction of the operator  $T_W$  we can give an explicit solution of (3.4.1).

**Proposition 3.4.2.** *In the exponential case  $h(t) = e^{-\alpha t}$  with  $v_t = v$ , the solution of (3.4.1) is explicitly given by*

$$\lambda_t(x) = e^{-\alpha t} e^{tT_W} v(x) + \alpha \int_0^t e^{-\alpha(t-s)} e^{(t-s)T_W} v(x) ds, \quad (3.4.5)$$

where  $(e^{tT_W})_{t \geq 0}$  is the semigroup of the bounded operator  $T_W$  defined as

$$e^{tT_W} u := \sum_{k=0}^{\infty} \frac{t^k}{k!} T_W^k u, \quad u \in L^\infty(I). \quad (3.4.6)$$

*Proof.* Define for  $t \geq 0$   $A(t) := x \mapsto e^{\alpha t} \lambda_t(x)$ . Multiplying (3.4.1) by  $e^{\alpha t}$ , we obtain that  $A(t)$  solves in  $L^\infty(I)$  the differential equation  $\frac{d}{dt} A(t) = \alpha e^{\alpha t} v + T_W A(t)$  with  $A(0) = \lambda_0(\cdot) = v$ . As  $t \rightarrow e^{tT_W} u$  is the unique solution of  $X'(t) = T_W X(t)$  with initial condition  $X(0) = u$  for  $u \in L^\infty(I)$ , a variation of constants formula gives  $A(t) = e^{tT_W} v + \alpha \int_0^t e^{(t-s)T_W} e^{\alpha s} v ds$ , and (3.4.5) follows by definition of  $A$ .  $\square$

**Example 3.4.3.** Consider the particular case of Expected Degree Distribution (EED) (see Chung & Lu (2002); Ouadah *et al.* (2019)), where  $W(x, y) = f(x)g(y)$  with  $f$  and  $g$  two positive functions on  $I$  such that  $f, g \in L^2(I, \nu)$ . Without any loss of generality, we assume  $\int_I g d\nu = 1$  and then  $D(x) = f(x)$  and  $r_\infty = \langle f, g \rangle$ . When  $\alpha \neq \langle f, g \rangle$ , the solution of (3.4.5) is given by

$$\lambda_t(x) = v(x) + \frac{\langle g, v \rangle}{\alpha - \langle f, g \rangle} (1 - e^{t(\langle f, g \rangle - \alpha)}) f(x).$$

The large time behavior depends then explicitly on the sign of  $\langle f, g \rangle - \alpha$ :

$$\begin{aligned} \langle f, g \rangle > \alpha &\Rightarrow \forall x \in I, \quad \lambda_t(x) \xrightarrow[t \rightarrow \infty]{} +\infty \text{ and} \\ \langle f, g \rangle < \alpha &\Rightarrow \forall x \in I, \quad \lambda_t(x) \xrightarrow[t \rightarrow \infty]{} v(x) + \frac{\langle g, v \rangle}{\alpha - \langle f, g \rangle} f(x). \end{aligned}$$

*Proof.* Recall that we have here  $v_t(x) = v(x)$ . By induction, we have explicitly that for  $k \geq 1$ ,  $T_W^k v = f \langle g, v \rangle \langle f, g \rangle^{k-1}$ . Since  $e^{tT_W} u = u + \sum_{k=1}^{\infty} \frac{t^k}{k!} T_W^k u$  when  $u \in L^\infty(I)$ ,

together in (3.4.5), we obtain

$$\begin{aligned} \lambda_t(x) &= e^{-\alpha t} \left( v(x) + \sum_{k=1}^{\infty} \frac{t^k}{k!} f(x) \langle g, v \rangle \langle f, g \rangle^{k-1} \right) \\ &\quad + \alpha \int_0^t e^{-\alpha(t-s)} \left( v(x) + \sum_{k=1}^{\infty} \frac{(t-s)^k}{k!} f(x) \langle g, v \rangle \langle f, g \rangle^{k-1} \right) ds \\ &= v(x) + f(x) \frac{\langle g, v \rangle}{\langle f, g \rangle} \left( e^{-t(\alpha - \langle f, g \rangle)} - e^{-\alpha t} + \frac{\alpha (1 - e^{-t(\alpha - \langle f, g \rangle)})}{\alpha - \langle f, g \rangle} - 1 + e^{-\alpha t} \right) \end{aligned}$$

which gives then the result.  $\square$

We now consider the general case.

### 3.4.2 Subcritical case

We assume that:

$$\|h\|_1 r_\infty < 1. \quad (3.4.7)$$

The main result is the following

**Theorem 3.4.4.** *Assume (3.4.7). Under Hypotheses 3.2.6 and 3.4.1*

- *there exists a unique function  $\ell : I \mapsto \mathbb{R}^+$  solution of*

$$\ell(x) = v_\infty(x) + \|h\|_1 \int_I W(x, y) \ell(y) \nu(dy), \quad (3.4.8)$$

*continuous and bounded on  $I$ . Moreover, there exists  $C_\ell > 0$  such that for all  $(x, y) \in I^2$ ,*

$$|\ell(x) - \ell(y)| \leq C_\ell \phi(\|x - y\|), \quad (3.4.9)$$

*where  $\phi$  is given in (3.2.7).*

- *for any  $x \in I$ , we have the convergence*

$$\lambda_t(x) \xrightarrow[t \rightarrow \infty]{} \ell(x). \quad (3.4.10)$$

The proof of Theorem 3.4.4 will be given in Section 3.10.1. We are now in position to address the question that motivates our paper: to what extent does the inhomogeneity of the underlying graph influence the macroscopic dynamics?

**Proposition 3.4.5.** *In the subcritical case (3.4.7),  $\ell$  solution of (3.4.8) is explicitly defined by*

$$\ell = \sum_{k=0}^{\infty} \|h\|_1^k T_W^k v_\infty. \quad (3.4.11)$$

*In particular, if  $(t, x) \mapsto v_t(x)$  is constant (i.e. for all  $(t, x)$ ,  $v_t(x) = v_\infty(x) = v$ ),  $\ell$  is uniform (i.e.  $\ell(x) = \ell$  for every  $x \in I$ ) if and only if the indegree is uniform (i.e.  $D(x) = \int_I W(x, y) \nu(dy) = D$  for every  $x \in I$ ). In such case,  $r_\infty = D$ .*

Note that (3.4.11) informs us about the influence of the macroscopic graph  $W$  on the dynamics: when  $v_\infty$  is constant, we have

$$\ell(x) = v_\infty \left( \sum_{k=0}^{\infty} \|h\|_1^k D^{(k)}(x) \right), \quad (3.4.12)$$

where  $D^{(0)} = 1$ ,  $D^{(1)} = D(x)$  and  $D^{(k+1)} = T_W D^{(k)}$ . We see from (3.4.12) that in order to understand  $\ell(x)$ , one needs to explore the structure of the macroscopic graph around  $x$ .

*Proof.* Equation (3.4.8) can be written  $\ell = v_\infty + \|h\|_1 T_W \ell$  which leads to

$$\|h\|_1 \left( \frac{Id}{\|h\|_1} - T_W \right) \ell = v_\infty.$$

As  $r_\infty < \frac{1}{\|h\|_1}$  in the subcritical case,  $\left( \frac{Id}{\|h\|_1} - T_W \right)$  is invertible (recall that  $r_\infty = \sup_{\sigma \in Sp(T_W)} |\sigma|$ ) and then  $\ell = (Id - \|h\|_1 T_W)^{-1} v_\infty = \sum_{k=0}^{\infty} \|h\|_1^k T_W^k v_\infty$ .

We take now  $v_\infty$  constant. Theorem 3.4.4 gives the existence of a unique  $\ell$  satisfying (3.4.8). Assume that this solution is a constant function  $\ell_0$ , then for all  $x \in I$  we have from (3.4.8)  $\ell_0 = v_\infty + \|h\|_1 \ell_0 \int_I W(x, y) \nu(dy)$  thus  $\int_I W(x, y) \nu(dy)$  is constant and is equal to  $\frac{\ell_0 - v_\infty}{\ell_0 \|h\|_1}$ . Conversely, assume  $\int_I W(x, y) \nu(dy)$  constant and equal to  $D$ . Then, a direct computation gives  $\|T_W f\|_\infty \leq \|f\|_\infty D$  hence (as  $r_\infty = \lim_{n \rightarrow \infty} \|T_W^n\|^{1/n}$ )  $r_\infty \leq D$ . As  $T_W \mathbf{1} = D \mathbf{1}$  (where  $\mathbf{1}(x) \equiv 1$ ), we have  $D \leq r_\infty$  thus  $D = r_\infty$ . The subcritical case can then be written as  $\|h\|_1 D < 1$  and we can define  $\ell_0 := \frac{v_\infty}{1 - \|h\|_1 D} > 0$ . The constant function  $\ell_0$  is continuous, bounded and solution of (3.4.8) which is unique: thus the solution of (3.4.8) is indeed constant.  $\square$

### 3.4.3 Supercritical case

We assume that:

$$\|h\|_1 r_\infty > 1. \quad (3.4.13)$$

Note again that, without space interaction (i.e.  $W = 1$ ), (3.4.13) reduces to  $\|h\|_1 > 1$  and it can be shown (see Delattre *et al.* (2016a), Theorem 11) that  $\lambda_t \sim \alpha e^{\beta t} \rightarrow \infty$  for some  $\alpha, \beta > 0$ . In our context with nontrivial  $W$ , one does not expect to have  $\lambda_t(x) \xrightarrow[t \rightarrow \infty]{} \infty$  uniformly on  $x$  as one can see from the obvious following example: take  $W(x, y) = \alpha \mathbf{1}_{[0, \frac{1}{2})^2}(x, y) + \beta \mathbf{1}_{[\frac{1}{2}, 1]^2}(x, y)$  for  $\alpha > \beta$ , then  $r_\infty = \frac{\alpha}{2}$ . This corresponds to two disconnected mean-field components  $A$  (for neurons with positions in  $I_A = [0, \frac{1}{2})$ ) and  $B$  (for neurons with positions in  $I_B = [\frac{1}{2}, 1]$ ). The critical parameter for population  $A$  (resp.  $B$ ) is hence  $\alpha_c = \frac{2}{\|h\|_1}$  (resp.  $\beta_c = \frac{2}{\|h\|_1}$ ). Taking now  $\alpha > \alpha_c$  and  $\beta < \beta_c$ , (3.4.13) is satisfied but one does not have  $\lambda_t(x) \xrightarrow[t \rightarrow \infty]{} \infty$  uniformly on  $x$  as the population  $B$  is subcritical, we only have  $\lambda_t(x) \xrightarrow[t \rightarrow \infty]{} \infty$  for  $x \in I_A$ .

In order to avoid such trivial examples, we assume that the graphon  $W$  is sufficiently connected in the following way. Defining for  $k \geq 1$ :

$$W^{(k)}(x, y) := \int_{I \times \dots \times I} W(x, x_1) \cdots W(x_{k-1}, y) dx_1 \cdots dx_{k-1},$$

we assume primitivity of  $W$  i.e. that there exists  $k$  such that

$$W^{(k)} > 0. \quad (3.4.14)$$

Note that  $W^{(k)}$  is the kernel of the operator  $T_W^k$ . To understand (3.4.14), think of the finite dimensional case with  $N$  particles interacting through a connectivity matrix  $A$ . In this context,  $A$  being primitive means the existence of some  $k \geq 1$  such that  $A^k(i, j) > 0$  for all  $i, j$ . Hypothesis (3.4.14) is the exact counterpart in infinite dimension. We also assume the more technical assumptions:

**Hypothesis 3.4.6.**

$$\sup_x \int_I W(x, y)^2 \nu(dy) =: C_W^{(2)} < \infty, \quad (3.4.15)$$

and

$$W(x, y) = W(y, x), \quad \text{for all } (x, y) \in I^2. \quad (3.4.16)$$

We also assume that we can define the Laplace transform of  $h$  for any  $z \geq 0$  :  $\mathcal{L}(h)(z) := \int_0^\infty e^{-tz} h(t) dt < \infty$ . Having  $h$  of polynomial growth works for instance.

**Proposition 3.4.7.** *Under Hypothesis 3.4.6, for all  $p \geq 1$ , the linear operator  $T_W^p$  is continuous from  $L^2(I)$  to  $L^2(I)$ , is compact, self-adjoint, its spectrum is the union of  $\{0\}$  and a discrete sequence of eigenvalues  $(\mu_n^{(p)})_{n \geq 1}$  such that  $\mu_n^{(p)} \rightarrow 0$  as  $n \rightarrow \infty$ . Moreover, the spectral radius  $r_2(T_W^p)$  verifies*

$$r_2(T_W^p) = r_\infty^p \quad (3.4.17)$$

where  $r_\infty$  defined in (3.4.4).

Secondly, if one assumes further hypothesis (3.4.14) for  $p = k$ ,  $\mu_0^{(k)} := r_2(T_W^k) > 0$  is an eigenvalue of  $T_W^k$  with a unique normalized eigenfunction  $h_0^{(k)}$  that is bounded, continuous and strictly positive on  $I$ . Moreover, every other eigenvalue  $\mu_n^{(k)}$  of  $T_W^k$  has modulus  $|\mu_n^{(k)}| < r_2(T_W^k)$ .

**Proposition 3.4.8.** *Suppose that we are in the supercritical case (3.4.13). Under Hypotheses 3.4.1 and 3.4.6,  $\int_I \lambda_t(x)^2 \nu(dx) \xrightarrow[t \rightarrow \infty]{} \infty$ .*

The proofs of Propositions 3.4.7 and 3.4.8 will be given in Section 3.10.2.

**Remark 3.4.9.** Proposition 3.4.8 provides a divergence result in  $L^2$  norm, that is not uniform in  $x$ . But under more restrictive hypotheses on the connectivity of  $W$  (without supposing  $W$  symmetric), one can easily derive uniform divergence result. Assume  $0 < \inf_{x \in I} u(x) =: \underline{u} < \infty$  and  $\|h\|_1 \inf_{x \in I} \|W(x, \cdot)\|_{1, \nu} > 1$ , we have then that  $\inf_{x \in I} \lim_{t \rightarrow \infty} \lambda(t, x) = +\infty$ . Note that by Fatou Lemma,  $\inf_{x \in I} \|W(x, \cdot)\|_{1, \nu} \leq r_\infty$ , hence it also implies the result of Proposition 3.4.8.

*Proof of Remark 3.4.9.* Let  $u(t, x) := \inf_{s \geq t} \lambda_s(x)$ . For all  $x \in I$ , set

$$\underline{\ell}(x) = \liminf_{t \rightarrow \infty} \lambda_t(x).$$

We have for all  $t > 0$ , using the positivity of  $W, h$  and  $\lambda$ , and the fact that  $\lambda_s(y) \geq u(\frac{t}{2}, y)$  for all  $s \in [\frac{t}{2}, t]$ ,

$$\begin{aligned} \lambda_t(x) &= v_t(x) + \int_0^{\frac{t}{2}} \int_I W(x, y) h(t-s) \lambda_s(y) \nu(dy) ds \\ &\quad + \int_{\frac{t}{2}}^t \int_I W(x, y) h(t-s) \lambda_s(y) \nu(dy) ds \\ &\geq v_t(x) + \int_0^{\frac{t}{2}} h(s) ds \int_I W(x, y) u\left(\frac{t}{2}, y\right) \nu(dy), \end{aligned}$$

then taking  $\liminf_{t \rightarrow \infty}$ , we obtain as  $u(\cdot, y)$  is non decreasing by monotone convergence

$$\inf_{x \in I} \underline{\ell}(x) \geq \underline{v}_\infty + \inf_{y \in I} \underline{\ell}(y) \|h\|_1 \inf_{x \in I} \int_I W(x, y) \nu(dy).$$

As  $\underline{v}_\infty$  is positive and  $\|h\|_1 \inf_{x \in I} \|W(x, \cdot)\|_{1, \nu} > 1$  (in the subcritical case), it implies that  $\inf_{x \in I} \underline{\ell}(x) = \inf_{x \in I} \lim_{t \rightarrow \infty} \inf_{s \geq t} \lambda_s(x) = +\infty$  hence the result.  $\square$

## 3.5 Applications

We give here examples of graphs  $(\mathcal{G}^{(N)})$  and corresponding graphons that satisfy the hypothesis of the paper. The main class of examples we have in mind fall into the framework of  $W$ -random graphs, see [Lovász & Szegedy \(2006\)](#); [Luçon \(2020\)](#).

### 3.5.1 A general class of examples

Given a positive measurable kernel  $(x, y) \mapsto \mathcal{P}(x, y)$  on  $I^2$ , for any  $N \geq 1$  we consider the interaction kernel

$$W_N(x, y) := \rho_N \min\left(\frac{1}{\rho_N}, \mathcal{P}(x, y)\right) \quad (3.5.1)$$

with  $\rho_N > 0$ . If  $\mathcal{P}$  is bounded, by modifying  $\rho_N$ , we can suppose with no loss of generality  $\|\mathcal{P}\|_\infty = 1$  and  $W_N(x, y) = \rho_N \mathcal{P}(x, y)$  whenever  $\rho_N \leq \frac{1}{\|\mathcal{P}\|_\infty}$ . Then, one distinguishes the dense case when  $\lim_{N \rightarrow \infty} \rho_N = \rho > 0$  and the diluted case when  $\rho_N \rightarrow 0$ .

#### 3.5.1.1 Uniformly bounded degrees

Suppose  $\sup_x \int_I \mathcal{P}(x, y) \nu(dy) < \infty$ . Recall that the prefactor  $\kappa_i^{(N)}$  in (3.2.3) was here to ensure that the interaction remains of order 1 as  $N \rightarrow \infty$ . In the dense case renormalization is not necessary, one can take  $\kappa_i^{(N)} = 1$ ; and in the diluted case we can take  $\kappa_i^{(N)} = \rho_N^{-1}$ . In either case, we take  $w_N = \rho_N$ . To satisfy

Hypothesis 3.3.7, we require  $\frac{N\rho_N}{\log(N)} \xrightarrow{N \rightarrow \infty} +\infty$ . Hypothesis 3.3.9 or Hypothesis 3.3.11 with  $W = \mathcal{P}$  are satisfied under regularity assumption on  $\mathcal{P}$ , see (Luçon, 2020, Propositions 3.2, 3.4, 3.6 and 3.9). Note that if  $\rho_N = 1$ , it is a direct consequence of Proposition 3.8.4 (in this case  $W^{\mathcal{G}_N^{(1)}} = W^{\mathcal{G}_N^{(2)}}$ , see Definition 3.8.3 for the graph  $\mathcal{G}_N^{(2)}$ ). Typical examples include the classic Erdős-Rényi graph with  $\mathcal{P} = 1$  (hence  $W_N = \rho_N$  is uniform), interaction with the P-nearest neighbors (see Omelchenko *et al.* (2012)), or the EDD model previously defined in Example 3.4.3. These examples are thoroughly detailed in the next part.

### 3.5.1.2 Unbounded degrees

Suppose that  $\int_I \mathcal{P}(x, y)^2 \nu(dy) < \infty$  for all  $x \in I$  and  $\mathcal{P}_* := \inf_{z \in I} \int \mathcal{P}(z, y) \nu(dy) > 0$ , but  $\sup_{x \in I} \int_I \mathcal{P}(x, y) \nu(dy) = \infty$ . Then we take the dilution parameter  $\kappa_i^{(N)} = N \left( \rho_N \sum_{j=1}^N \min \left( \frac{1}{\rho_N}, \mathcal{P}(x_i, x_j) \right) \right)^{-1}$ , and the macroscopic interaction kernel is  $W(x, y) = \frac{\mathcal{P}(x, y)}{\int_I \mathcal{P}(x, z) \nu(dz)}$ . For such examples, see (Luçon, 2020, Section 3.4). For instance, consider  $\mathcal{P}(x, y) = x^{-\alpha} g(y)$  with  $g$  a probability measure on  $[0, 1]$  and  $\alpha < \frac{1}{2}$ .

We present in the following different concrete examples of application of our results. We focus on the framework  $I = [0, 1]$  with the regular distribution of the positions  $x_i^{(N)} = \frac{i}{N}$ ,  $1 \leq i \leq N$  and  $\nu$  the Lebesgue measure. We take  $f(x) = x$  to apply the results of Section 3.4.

### 3.5.2 Example: Erdős-Rényi graph

Taking  $\mathcal{P} \equiv 1$  with  $\rho_N \in [0, 1]$ , (3.5.1) becomes  $W_N \equiv \rho_N$ . This corresponds to the case where  $\mathcal{G}^{(N)}$  is a (possibly diluted) Erdős-Rényi random graph: the dense case corresponds to  $\rho_N \rightarrow \rho \in (0, 1]$  (and one takes  $\kappa_i \equiv 1$  for all  $i$ ) whereas the diluted case corresponds to  $\rho_N \rightarrow 0$  (and one chooses  $\kappa_i \equiv \frac{1}{\rho_N}$ ). The dilution condition (3.10) reduces to

$$\frac{N\rho_N}{\log N} \xrightarrow{N \rightarrow \infty} +\infty. \quad (3.5.2)$$

Note that the condition (3.5.2) is the very same condition already met in the similar context of diffusions interacting on Erdős-Rényi random graphs (see Delattre *et al.* (2016b); Bet *et al.* (2020)), in the quenched case (i.e. where the randomness of the graph is frozen). In the (technically simpler) annealed case (where one integrates also with respect to the randomness of the graph), it is possible to get rid of this supplementary  $\log N$  term (that is required, in the present quenched setting, for our Borel-Cantelli arguments to work) and assume only  $N\rho_N \rightarrow \infty$  (as this has been done for diffusions in an annealed setting e.g. in Bhamidi *et al.* (2019); Coppini *et al.* (2019); Bayraktar *et al.* (2023)). Here, the limiting graphon is given by



$W \equiv \rho$  (with  $\rho = 1$  in the diluted case). Condition (3.2.5) is then trivially satisfied and one can apply Theorems 3.3.10 and 3.3.12 (the convergence of graphs is seen in Proposition 3.8.4). As the degree is constant, Proposition 3.4.5 gives  $r_\infty = \rho$ . As Hypothesis 3.4.6 is satisfied, there is a transition phase around  $\rho_c = \frac{1}{\|h\|_1}$ .

In the subcritical case  $\|h\|_1 \rho < 1$ , Theorem 3.4.4 gives that for any  $x \in I$ ,  $\lambda_t(x) \xrightarrow[t \rightarrow \infty]{} \ell(x) = \frac{v_\infty(x)(1 - \|h\|_1 \rho) + \|v_\infty\|_{I, \nu, 1} \|h\|_1 \rho}{1 - \|h\|_1 \rho}$ . Note that if  $v_\infty$  is constant,  $\ell = \frac{v_\infty}{1 - \|h\|_1 \rho}$ . Corresponding simulations are given in Figures 3.1 and 3.2.

In the supercritical case  $\|h\|_1 \rho > 1$ , as  $W$  is constant, we can directly apply Remark 3.4.9 and obtain  $\inf_{x \in I} \lim_{t \rightarrow \infty} \lambda_t(x) = +\infty$ .

### 3.5.3 Example: P-nearest neighbor model (Omelchenko *et al.*, 2012)

Consider the kernel  $W(x, y) = \mathbf{1}_{d_{\mathcal{S}_1}(x, y) < r}$  for any  $(x, y) \in I^2$  and for some fixed  $r \in (0, \frac{1}{2})$ , with

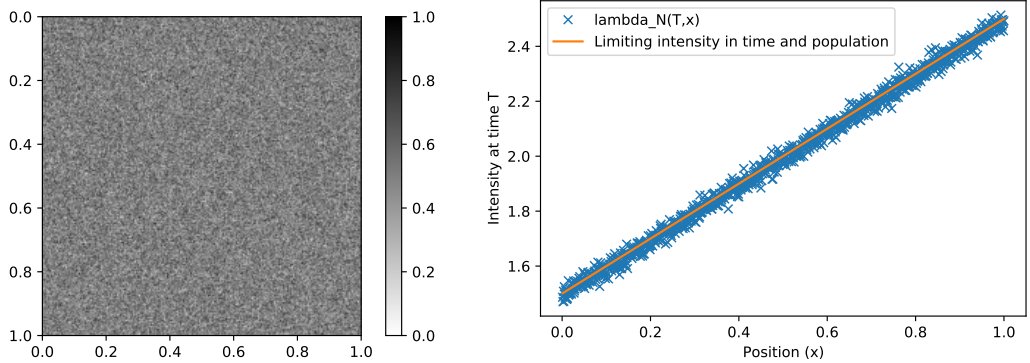
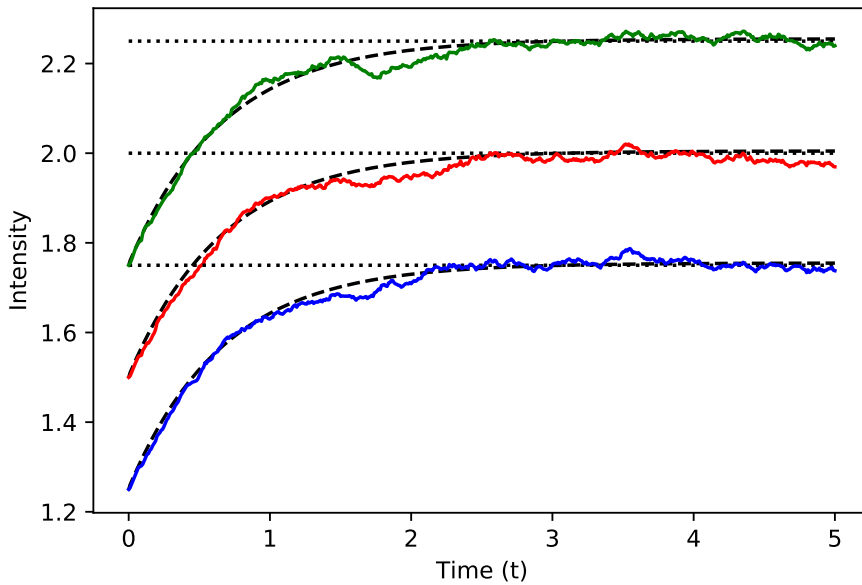
$$d_{\mathcal{S}_1}(x, y) = \min(|x - y|, 1 - |x - y|). \quad (3.5.3)$$

It means that the particles at positions  $x$  and  $y$  interact if and only if they are at distance less than  $r$  on the circle  $\mathcal{S}_1 := \mathbb{R}/_{[0,1]}$ . This corresponds to a deterministic graph. As (3.2.5) is satisfied - for any  $(x, x') \in I^2$ ,  $\int_I |W(x, y) - W(x', y)| \nu(dy) = \int_0^1 |\mathbf{1}_{|x-y|<r} - \mathbf{1}_{|x'-y|<r}| dy \leq 4|x - x'|$ , we can apply Theorems 3.3.10 and 3.3.12. As for any  $x \in I$ ,  $\int_I W(x, y) dy = 2r$ , Proposition 3.4.5 gives that  $r_\infty = 2r$ . The assumptions (3.4.15) and (3.4.16) are trivially verified, and as  $W^{(k)}$  is positive for  $k := \inf \{n \geq 0, nr \geq \frac{1}{2}\}$ , Hypothesis 3.4.6 is satisfied and there is a transition phase around  $r_c = \frac{1}{2\|h\|_1}$ . In the subcritical case ( $r < r_c$ ), Proposition 3.4.5 gives that when  $v_t(x) = v$  for all  $(t, x) \in \mathbb{R}_+ \times I$ , the limiting intensity is explicit and  $\ell = \frac{v}{1 - 2r\|h\|_1}$ . We give an example of simulation in this case in Figure 3.3. In the supercritical case ( $r > r_c$ ), as the degree is constant, we can directly apply Remark 3.4.9 and obtain  $\inf_{x \in I} \lim_{t \rightarrow \infty} \lambda_t(x) = +\infty$ .

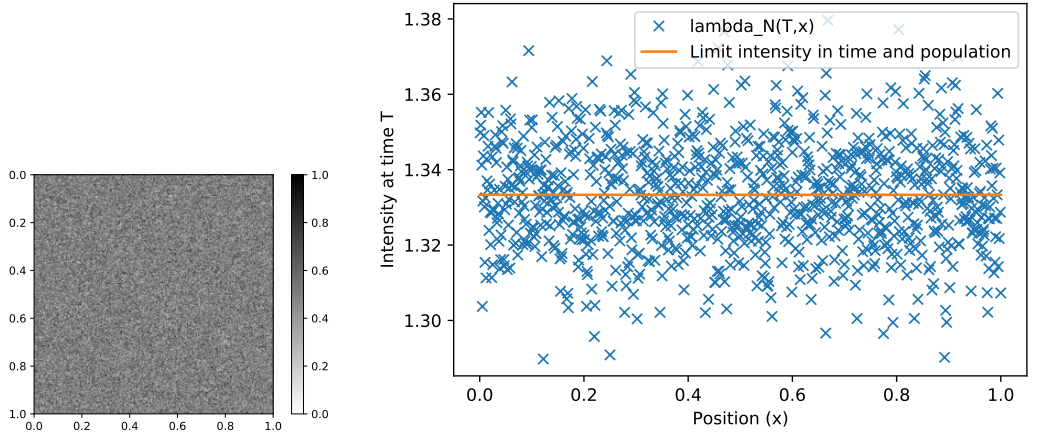
### 3.5.4 Example: Inhomogeneous graph with EDD (Chung & Lu, 2002)

Recall Example 3.4.3:  $W(x, y) = f(x)g(y)$  with  $f$  and  $g$  two positive bounded functions on  $I$  such that  $f, g \in L^2(I, \nu)$  and  $\int_I g d\nu = 1$ . We also suppose that  $f$  satisfies a Hölder condition for  $\vartheta \in (0, 1]$  and is bounded. Note that the indegree is  $D(x) = f(x)$ . Hypothesis 3.2.6 is satisfied and we can apply Theorems 3.3.10 and 3.3.12. The operator  $T_W$  is then defined as  $T_W k(x) = f(x)\langle g, k \rangle$  for  $k \in L^\infty$ . An iteration gives  $T_W^n = \langle f, g \rangle^{n-1} T_W$  for all  $n \geq 1$ , and then  $r_\infty = \langle f, g \rangle$ , so that the phase transition is given in term of  $\langle f, g \rangle \|h\|_1 < 1$  or  $\langle f, g \rangle \|h\|_1 > 1$  (and we retrieve Example 3.4.3 in the exponential case).

In the subcritical case  $\|h\|_1 \langle f, g \rangle < 1$ , Theorem 3.4.4 gives that for any  $x \in I$ ,  $\lambda_t(x) \xrightarrow[t \rightarrow \infty]{} \ell(x)$  where  $\ell$  is the solution of (3.4.8), that is  $\ell(x) = v_\infty(x) +$

(a) Matrix of  $\mathcal{G}^{(N)}$ (b) Each dot represents  $\lambda_{N,i}(T)$  at position  $x_i^{(N)}$  for  $1 \leq i \leq N$ , and the plain line corresponds to the macroscopic limit  $\ell(x)$ .(c) Evolution of microscopic and macroscopic intensities of three particles at positions  $x=0.25$  (blue - the lowest),  $0.5$  (red) and  $0.75$  (green - the highest). In each case, the colored line represents  $\lambda_{N,i}(t)$ , the dashed line represents  $\lambda_t(x)$  and the dotted line represents the limit  $\ell(x)$ .Figure 3.1: Simulation of Example 3.5.2 with inhomogeneous  $v_t$ 

We chose  $h(t) = e^{-\alpha t}$  with  $\alpha = 2$ ,  $\rho = 0.5$  for the Erdős-Rényi graph and  $v_t(x) = x + 1$ . We are in the subcritical case  $\|h\|_1 \rho < 1$  and the limiting intensity is given by  $\ell(x) = x + \frac{1}{2}$ . We run a simulation for  $N = 1000$  and  $T = 5$ : in 3.1a, we show the matrix of the Erdős-Rényi graph  $\mathcal{G}^{(N)}$ . In 3.1b, we represent the spatial distribution of intensities at fixed time  $T$ . In 3.1c, we show the time evolution of the intensities for different positions. Note here that the inhomogeneity of  $\ell(x)$  is due to the inhomogeneity of the  $v_t$ , not of the graph.


 (a) Matrix of  $\mathcal{G}^{(N)}$ 

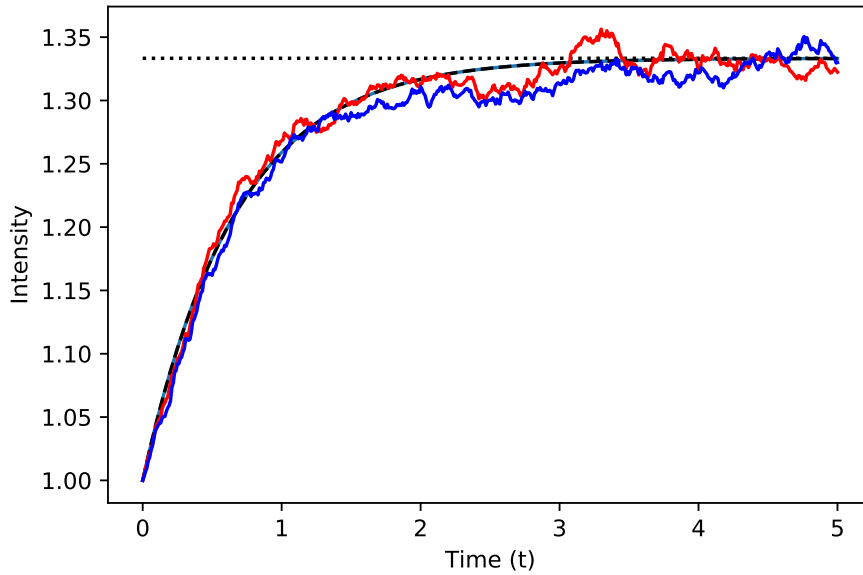
 (b) Each dot represents  $\lambda_{N,i}(T)$  at position  $x_i^{(N)}$  for  $1 \leq i \leq N$ , and the plain line corresponds to the macroscopic limit  $\ell(x)$ .

 (c) Evolution of microscopic and macroscopic intensities of two particles at positions  $x = 0.5$  (red) and  $0.75$  (blue). The colored lines represent  $\lambda_{N,i}(t)$ , the dashed line represents  $\lambda_t$  and the dotted line represents the limit  $\ell$ .

 Figure 3.2: Simulation of Example 3.5.2 with homogeneous  $v_t$ 

We chose  $h(t) = e^{-\alpha t}$  with  $\alpha = 2$ ,  $\rho = 0.5$  for the Erdős-Rényi graph and  $v_t(x) = 1$ , we are in the subcritical case ( $\|h\|_1 \rho = \frac{1}{4} < 1$ ). As the graph is homogeneous in space and the self-activity is constant, the limit solution of (3.1.2) does not depend of the position:  $\lambda_t = \frac{4}{3} - \frac{1}{3}e^{-\frac{2}{3}t}$ . The limiting intensity is constant  $\ell = \frac{4}{3}$ . We run a simulation for  $N = 1000$  and  $T = 5$ . In 3.2a, we show the matrix of the Erdős-Rényi graph  $\mathcal{G}^{(N)}$ . In 3.2c, we show the time evolution of the intensities for different positions. In 3.2b, we represent the spatial distribution of intensities at fixed time  $T$ .

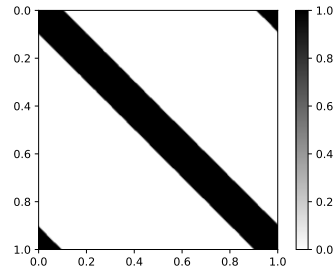
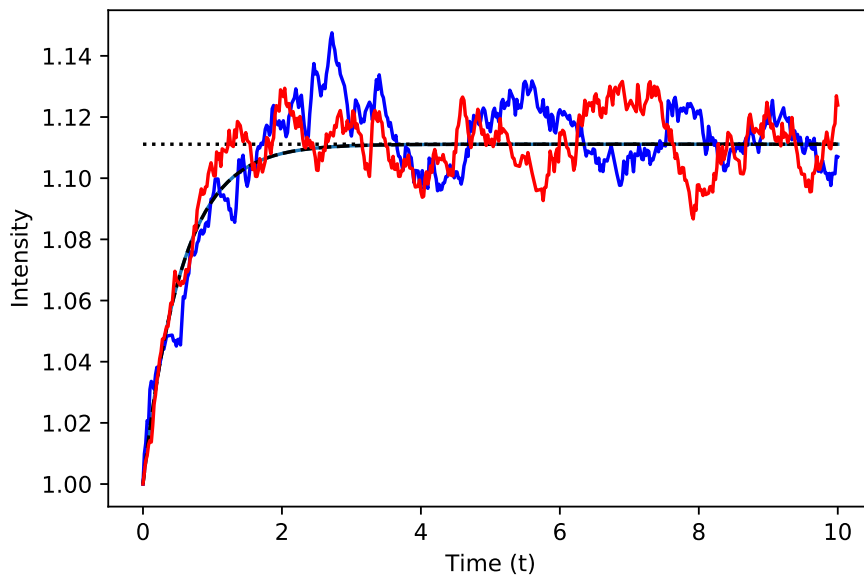
(a) Matrix of  $\mathcal{G}^{(N)}$ (b) Evolution of microscopic and macroscopic intensities of two particles at positions  $x=0.5$  (red) and  $0.1$  (blue). The colored lines represent  $\lambda_{N,i}(t)$ , the dashed line represents  $\lambda_t$  and the dotted line represents the limit  $\ell$ .

Figure 3.3: Simulation of Example 3.5.3 in the subcritical case

We chose  $h(t) = e^{-2t}$ ,  $r = 0.1$  and  $v_t(x) = 1$ , hence we are in the subcritical case as  $2r\|h\|_1 < 1$ . The graph is not homogeneous in space but has a symmetry and the self-activity  $v$  is constant, hence the solution of (3.1.2) does not depend of the position:  $\lambda_t = \frac{10}{9} - \frac{1}{9}e^{-\frac{2}{5}t}$ . The limiting intensity is constant  $\ell = \frac{10}{9}$ . We run a simulation for  $N = 500$  particles and a final time  $T = 10$ : in 3.3a, we show the matrix of the graph  $\mathcal{G}^{(N)}$  obtained. In 3.3b, we show the time evolution of the intensities for different positions. We see that the simulated intensities follow indeed the behavior expected, as they are close to  $\lambda_t(x)$  and converge toward a constant limit  $\ell$ .

$\|h\|_1 \frac{f(x)\langle v_\infty, g \rangle}{1 - \|h\|_1 \langle f, g \rangle}$ . We give an example of simulation in the case  $f = g$  in Figure 3.4.

### 3.5.5 Example: Multi-class interaction populations

Another interesting case concerns deterministic and inhomogeneous graphs modeling the macroscopic organization of neurons into vertical columns. A generic construction is the following: divide  $I = (0, 1]$  into  $P$  consecutive subintervals  $I_j$  with respective length  $\alpha_j > 0$ , that is,  $I_j = (\alpha_1 + \dots + \alpha_{j-1}, \alpha_1 + \dots + \alpha_j]$  and  $\alpha_1 + \dots + \alpha_P = 1$ . Take any connectivity matrix  $M$  between the  $P$  populations,  $M = (m_{ij})_{1 \leq i, j \leq P}$  with  $m_{ij} \in \{0, 1\}$  modeling the deterministic connection between subpopulations  $i$  and  $j$ . Take the self-activity fixed for each population, described by  $u_0(t) = (u_{0,i}(t))_{1 \leq i \leq P}$  and converging towards  $u = (u_i)_{1 \leq i \leq P}$  as  $t \rightarrow \infty$ . Define finally  $W(x, y) = \sum_{i, j=1}^P m_{ij} \mathbf{1}_{x \in I_i} \mathbf{1}_{y \in I_j}$ , as well as  $\widetilde{M} = (\alpha_j m_{ij})_{1 \leq i, j \leq P}$ . Then  $r_\infty = \rho(\widetilde{M})$  where  $\rho(\widetilde{M})$  is the spectral radius of  $\widetilde{M}$  so that the phase transition described above is given here in terms of  $\rho(\widetilde{M})\|h\|_1 < 1$  or  $\rho(\widetilde{M})\|h\|_1 > 1$ .

The limiting intensity  $\lambda_t(x)$  is constant on each population, described by the vector  $\widetilde{\lambda}(t) = (\lambda_i(t))_{1 \leq i \leq P}$  which solves  $\widetilde{\lambda}(t) = u_0(t) + \int_0^t h(t-s)\widetilde{M}\widetilde{\lambda}(s)ds$ . In the subcritical case, the limit  $\ell = (\ell_i)_{1 \leq i \leq P}$  is piecewise constant (on each population) and solves  $\ell = u + \|h\|_1 \widetilde{M}\ell$ . In the supercritical case,  $\sum_{i=1}^P \alpha_i \lambda_i(t)^2 \xrightarrow{t \rightarrow \infty} \infty$  when  $\widetilde{M}$  is symmetric and primitive.

*Remark.* A closer look at the proof of Theorem 2.3 and (2.5) of [Athreya & Murthy \(1976\)](#) shows that  $\lambda_i(t) \rightarrow \infty$  for all  $i \in \llbracket 1, p \rrbracket$  in the simpler case when  $M$  is only irreducible but not necessarily symmetric nor primitive (e.g. the case considered in [Ditlevsen & Löcherbach \(2017\)](#)).

## 3.6 Possible extensions

Inhibition is an important factor in neuronal dynamics. In the present model, we restricted ourselves for simplicity to a nonnegative interaction kernel  $W$ . Nevertheless, we can easily introduce a dependence in inhibition by considering signed spatial interaction  $W$ : take  $w_{ij}^{(N)}$  in (3.1.1) of the form  $w_{ij} = \kappa_i^{(N)} \xi_{ij}^{(N)} s(x_i, x_j)$ , where  $s(x, y) \in \{\pm 1\}$  expresses the nature of interaction between neurons located in  $x$  and  $y$ :  $s(x, y) = 1$  if the interaction is excitatory or  $-1$  if it is inhibitory. For instance, if the nature of the interaction only depends on the neuron sending the information, take  $s(x, y) = s(y)$ . The resulting macroscopic limit is now expressed in terms of a signed interaction kernel  $W$ . The results presented in this paper remain the same with appropriate regularity assumptions on  $s$ , up to notational changes in the norms where  $W$  is replaced by  $|W|$ .

Another possible extension concerns the memory kernel  $h$ . In our paper, this kernel is identical on the population. One could think that neurons can present an inhomogeneity in the way they remember the past information, that is considering a memory kernel depending also on the positions of the neurons  $h(t, x, y)$ . With enough regularity on such  $h$ , the same results hold up to notational changes.

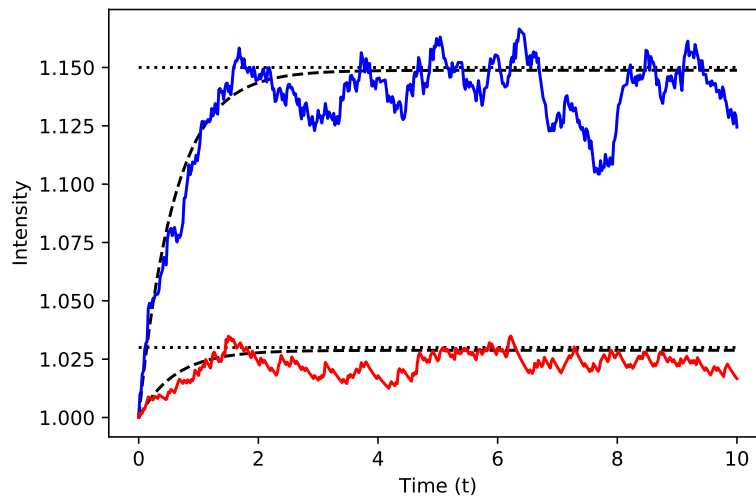
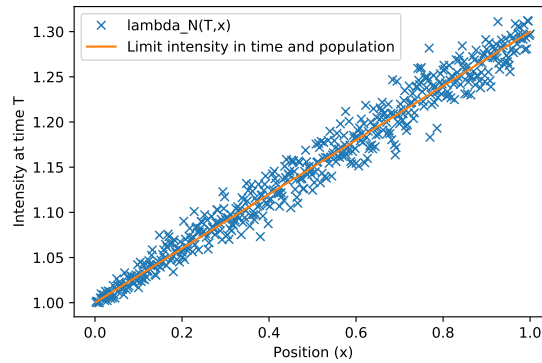
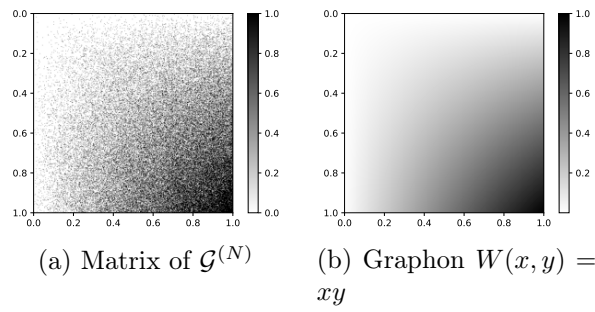


Figure 3.4: Simulation of Example 3.5.4

We chose  $h(t) = e^{-\alpha t}$  with  $\alpha = 2$ ,  $v_t(x) = 1$  and  $f(x) = g(x) = x$ , that is  $W(x, y) = xy$ : we are in the subcritical case ( $\|h\|_1 \langle f, f \rangle < 1$ ) and the limiting intensity is  $\ell(x) = 1 + \frac{3}{10}x$ . We run a simulation for  $N = 500$  and  $T = 10$ : in 3.4b, we represent the graphon  $W$ , and in 3.4a we show the matrix random graph  $\mathcal{G}^{(N)}$  obtained. In 3.4c, we represent the spatial distribution of intensities at fixed time  $T$ . In 3.4d, we show the time evolution of the intensities for different positions. Note here that the inhomogeneity of  $\ell(x)$  is only due to the inhomogeneity of the kernel  $W$ .

## 3.7 Existence and uniqueness of the model and its limit (proofs)

### 3.7.1 Proof of Proposition 3.2.5

We study the pathwise uniqueness by considering the total variation distance between two such processes. We show the existence by constructing a Cauchy sequence adapted and using a Picard iteration argument. We follow the structure of the proof proposed in (Delattre *et al.*, 2016a, Theorem 6). Let  $N \geq 1$ , we consider a family of independent Poisson measures  $(\pi_i(ds, dz))_{1 \leq i \leq N}$  with intensity  $dsdz$ . We start by showing uniqueness and we omit the notations  $^{(N)}$  and  $_N$  for simplicity. We set  $(Z_i(t))_{i \in \llbracket 1, N \rrbracket, t \geq 0}$  and  $(\underline{Z}_i(t))_{i \in \llbracket 1, N \rrbracket, t \geq 0}$  two solutions of the system (3.2.2) such that  $\mathbf{E}[Z_i(t)] < +\infty$  and  $\mathbf{E}[\underline{Z}_i(t)] < +\infty$  for any  $i \in \llbracket 1, N \rrbracket$  and  $t \geq 0$ . For any  $i \in \llbracket 1, N \rrbracket$ , we consider the total variation distance between  $Z_i$  and  $\underline{Z}_i$  on  $[0, t]$ :

$$\Delta_i(t) := \int_0^t |d(Z_i(s) - \underline{Z}_i(s))|.$$

$\Delta_i(t)$  counts the number of unshared jumps between  $Z_i$  and  $\underline{Z}_i$  on  $[0, t]$ . We denote respectively by  $\lambda_i$  and  $\underline{\lambda}_i$  the stochastic intensities of  $Z_i$  and  $\underline{Z}_i$ . As they are constructed on the same Poisson measure  $\pi_i$ , the unshared jumps are the points of  $\pi_i$  located between the two intensities, thus we have

$$\Delta_i(t) = \int_0^t \int_0^{+\infty} |\mathbf{1}_{\{z \leq \lambda_i(s)\}} - \mathbf{1}_{\{z \leq \underline{\lambda}_i(s)\}}| \pi_i(ds, dz).$$

Setting  $\delta_i(t) := \mathbf{E}[\Delta_i(t)]$ , we obtain with Fubini's Theorem

$$\delta_i(t) = \mathbf{E} \left[ \int_0^t \int_0^{+\infty} |\mathbf{1}_{\{z \leq \lambda_i(s)\}} - \mathbf{1}_{\{z \leq \underline{\lambda}_i(s)\}}| dz ds \right] = \int_0^t \mathbf{E}[|\lambda_i(s) - \underline{\lambda}_i(s)|] ds.$$

Using (3.2.3) and as  $f$  is Lipschitz continuous (Hypothesis 3.2.4), we have

$$\begin{aligned} \delta_i(t) &= \int_0^t \mathbf{E} \left[ \left| f \left( u_0(s, x_i) + \frac{1}{N} \sum_{j=1}^N w_{ij} \int_{]0, s[} h(s-u) dZ_j(u) \right) \right. \right. \\ &\quad \left. \left. - f \left( u_0(s, x_i) + \frac{1}{N} \sum_{j=1}^N w_{ij} \int_{]0, s[} h(s-u) d\underline{Z}_j(u) \right) \right| \right] ds \\ &\leq \int_0^t \mathbf{E} \left[ L_f \left| \frac{1}{N} \sum_{j=1}^N w_{ij} \int_{]0, s[} h(s-u) (dZ_j(u) - d\underline{Z}_j(u)) \right| \right] ds \\ &\leq L_f \frac{1}{N} \sum_{j=1}^N w_{ij} \mathbf{E} \left[ \int_0^t \int_{]0, s[} |h(s-u)| d\Delta_j(u) ds \right]. \end{aligned}$$

We apply Lemma 3.A.1 ( $\Delta_i$  is with finite variations,  $\Delta_i(0) = 0$  and  $h$  is locally integrable) and obtain

$$\delta_i(t) \leq L_f \frac{1}{N} \sum_{j=1}^N w_{ij} \int_0^t |h(t-s)| \delta_j(s) ds.$$

We set  $\delta(t) = \sum_{i=1}^N \delta_i(t)$  and  $W_N = \max_{(i,j) \in \llbracket 1, N \rrbracket^2} w_{ij}$ . Then, summing on  $i$ , we have

$$\delta(t) \leq L_f \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N W_N \int_0^t |h(t-s)| \delta_j(s) ds \leq L_f W_N \int_0^t |h(t-s)| \delta(s) ds.$$

Since  $h$  is locally integrable,  $\delta$  is nonnegative and locally bounded, we can apply Lemma 3.A.3 (i) and obtain that  $\delta(t) = 0$  for all  $t \geq 0$ . As each  $\Delta_i$  is nonnegative, we obtain that for all  $i \in \llbracket 1, N \rrbracket$  and  $t \geq 0$ ,  $\Delta_i(t) = 0$  almost surely. Hence  $Z_i(t) = \underline{Z}_i(t)$  almost surely for all  $i \in \llbracket 1, N \rrbracket$  and  $t \geq 0$ , which gives the uniqueness.

We show now the existence of a process satisfying (3.2.2). To do it, we proceed by iteration: for all  $i \in \llbracket 1, N \rrbracket$  and  $t \geq 0$ , let  $Z_{i,0}(t) = 0$ . Then, for all  $n \geq 0$  we set:

$$Z_{i,n+1}(t) = \int_0^t \int_0^{+\infty} \mathbf{1}_{\{z \leq f(v_t(x_i) + \frac{1}{N} \sum_{j=1}^N w_{ij} \int_0^{s-} h(s-u) dZ_{j,n}(u))\}} \pi_i(ds, dz).$$

With  $i$  and  $n$  fixed, such a process  $(Z_{i,n+1})$  exists: it is a counting process with stochastic intensity  $\lambda_{i,n+1}(t) = f\left(v_t(x_i) + \frac{1}{N} \sum_{j=1}^N w_{ij} \int_0^{t-} h(t-u) dZ_{j,n}(u)\right)$ . As for the uniqueness, we set for all  $i \in \llbracket 1, N \rrbracket$ ,  $n \geq 0$  and  $t \geq 0$ ,  $\delta_n(t) = \sum_{i=1}^N \delta_{i,n}(t)$

with  $\delta_{i,n}(t) = \mathbf{E} \left[ \int_0^t |dZ_{i,n+1}(s) - dZ_{i,n}(s)| \right]$ . As done previously, we find:

$$\begin{aligned} \delta_{i,n+1}(t) &= \mathbf{E} \left[ \int_0^t |dZ_{i,n+2}(s) - dZ_{i,n+1}(s)| \right] \\ &= \mathbf{E} \left[ \int_0^t \int_0^{+\infty} |\mathbf{1}_{\{z \leq \lambda_{i,n+2}(s)\}} - \mathbf{1}_{\{z \leq \lambda_{i,n+1}(s)\}}| dz ds \right] \\ &\leq \int_0^t \mathbf{E} \left[ L_f \left| \frac{1}{N} \sum_{j=1}^N w_{ij} \int_{]0,s[} h(s-u) (dZ_{j,n+1}(u) - dZ_{j,n}(u)) \right| \right] ds. \end{aligned}$$

Summing on  $i$  and using Lemma 3.A.1 we obtain

$$\delta_{n+1}(t) \leq L_f W_N \int_0^t |h(t-s)| \delta_n(s) ds. \quad (3.7.1)$$

We want to apply Lemma 3.A.3(ii), but for this we have to show that  $\delta_n$  is locally bounded. We note  $m_{i,n}(t) = \mathbf{E}[Z_{i,n}(t)]$  and  $\bar{m}_n(t) = \sum_{i=1}^N m_{i,n}(t)$ . By construction,

$$m_{i,n+1}(t) = \mathbf{E} \left[ \int_0^t \int_0^{+\infty} \mathbf{1}_{\{z \leq f(v_s(x_i) + \frac{1}{N} \sum_{j=1}^N w_{ij} \int_0^{s-} h(s-u) dZ_{j,n}(u))\}} \pi_i(ds, dz) \right].$$

As  $\pi_i$  is a random Poisson measure with intensity  $ds dz$ , we have

$$m_{i,n+1}(t) = \mathbf{E} \left[ \int_0^t f \left( v_s(x_i) + \frac{1}{N} \sum_{j=1}^N w_{ij} \int_0^{s-} h(s-u) dZ_{j,n}(u) \right) ds \right].$$



By Hypothesis 3.2.4, we have that  $f(y) \leq f(0) + L_f|y|$  for all  $y$  so that:

$$\begin{aligned} m_{i,n+1}(t) &\leq \mathbf{E} \left[ \int_0^t \left( f(0) + L_f \left| v_s(x_i) + \frac{1}{N} \sum_{j=1}^N w_{ij} \int_0^{s-} h(s-u) dZ_{j,n}(u) \right| \right) ds \right] \\ &\leq f(0)t + L_f \|v\|_\infty t + \frac{1}{N} \sum_{j=1}^N w_{ij} \int_0^t \int_0^{s-} |h(s-u)| dm_{j,n}(u) ds. \end{aligned}$$

Applying Lemma 3.A.1 and summing on  $i$  we obtain

$$\bar{m}_{n+1}(t) \leq Nt(f(0) + L_f \|v\|_\infty) + W_N \int_0^t |h(t-s)| \bar{m}_n(s) ds. \quad (3.7.2)$$

As  $\bar{m}_0 = 0$  and  $h$  is locally integrable, by induction  $\bar{m}_n$  is locally bounded for all  $n \geq 0$ . Yet  $\delta_n(t) = \sum_{i=1}^N \mathbf{E} \left[ \int_0^t |dZ_{i,n+1}(s) - dZ_{i,n}(s)| \right] \leq \bar{m}_{n+1}(t) + \bar{m}_n(t)$  hence  $\delta_n$  is indeed locally bounded for all  $n$ . Lemma 3.A.3(ii) and (3.7.1) give then that for all  $T > 0$ , there exists  $C_T$  such that  $\sup_{t \in [0, T]} \sum_{n \geq 0} \delta_n(t) \leq C_T < \infty$ . Thus we have

$$\sup_{t \in [0, T]} \sum_{n \geq 0} \sum_{i=1}^N \mathbf{E} \left[ \int_0^t |dZ_{i,n+1}(s) - dZ_{i,n}(s)| \right] \leq C_T < \infty.$$

Thus for  $i$  fixed, the sequence of random variables  $(Z_{i,n})_n$  is Cauchy in  $L^1$  on the space  $D([0, t], \mathbb{R})$  with the expectation of the total variation distance. Hence there exists a process  $Z_i$  such that  $\mathbf{E} \left[ \int_0^T |dZ_{i,n}(s) - dZ_i(s)| \right] \xrightarrow{n \rightarrow \infty} 0$ . From this convergence and a diagonal argument, there exists an extraction  $\varphi$  such that for all  $i$ ,

$$\int_0^T |dZ_{i,\varphi(n)}(s) - dZ_i(s)| \xrightarrow{n \rightarrow \infty} 0.$$

Since  $\int_0^T |dZ_{i,\varphi(n)}(s) - dZ_i(s)|$  is an integer,  $Z_{i,\varphi(n)}$  is a.s. stationary and one obtains from this that the right hand side of

$$Z_{i,\varphi(n)+1}(t) = \int_0^t \int_0^{+\infty} \mathbf{1}_{\{z \leq f(v_t(x_i) + \frac{1}{N} \sum_{j=1}^N w_{ij} \int_0^{s-} h(s-u) dZ_{j,\varphi(n)}(u))\}} \pi_i(ds, dz) \quad (3.7.3)$$

is equal to  $\int_0^t \int_0^{+\infty} \mathbf{1}_{\{z \leq f(v_t(x_i) + \frac{1}{N} \sum_{j=1}^N w_{ij} \int_0^{s-} h(s-u) dZ_j(u))\}} \pi_i(ds, dz)$ . Hence the left hand side of (3.7.3) converges too, towards some  $\tilde{Z}_i(t)$ . It remains to show that  $\tilde{Z} = Z$ . Fatou's Lemma gives

$$\mathbf{E} \left[ \int_0^T |dZ_i(s) - d\tilde{Z}_i(s)| \right] \leq \liminf_{n \rightarrow \infty} \mathbf{E} \left[ \int_0^T |dZ_{i,\varphi(n)}(s) - dZ_{i,\varphi(n)+1}(s)| \right] = 0$$

as  $(Z_{i,n})_n$  is a Cauchy sequence. We have then that the limit process verifies a.s.

$$Z_i(t) = \int_0^t \int_0^{+\infty} \mathbf{1}_{\{z \leq f(v_s(x_i) + \frac{1}{N} \sum_{j=1}^N w_{ij} \int_0^{s-} h(s-u) dZ_j(u))\}} \pi_i(ds, dz).$$

This gives the existence of  $(Z_1(t), \dots, Z_N(t))_{t \geq 0}$  satisfying (3.2.2). Now let us verify that  $t \mapsto \sup_{1 \leq i \leq N} \mathbf{E}[Z_{N,i}(t)]$  is locally bounded. Recall (3.7.2): as  $\bar{m}_n$  is locally bounded, by Lemma 3.A.3(iii) for all  $T > 0$ , there exists  $C_T$  such that  $\sup_{t \in [0, T]} \sup_{n \geq 0} \bar{m}_n(t) \leq C_T < +\infty$  hence

$$\sup_{t \in [0, T]} \sup_{n \geq 0} \sum_{i=1}^N \mathbf{E}[Z_{i,n}(t)] \leq C_T < +\infty$$

and by dominated convergence, for all  $T > 0$ ,  $\sup_{t \in [0, T]} \sum_{i=1}^N \mathbf{E}[Z_i(t)] < +\infty$  and the proof is concluded with the multivariate Hawkes process  $(Z_1(t), \dots, Z_N(t))_{t \geq 0} = (Z_{N,1}(t), \dots, Z_{N,N}(t))_{t \geq 0}$ .  $\square$

### 3.7.2 Proof of Theorem 3.2.7

We show existence and uniqueness of a continuous and bounded solution  $(t, x) \mapsto \lambda_t(x)$  to equation (3.1.2) on  $[0, T] \times I$ . For the purpose of the proof, we write  $\lambda_t(x) = \lambda(t, x)$ . We follow the proof proposed in (Chevallier *et al.*, 2019, Proposition 5), with major changes to accommodate our hypotheses. We apply Banach fixed-point Theorem. We consider the map  $F$  defined on  $\mathcal{C}_b([0, T] \times I, \mathbb{R})$ , that is the set of bounded continuous functions defined on  $[0, T] \times I$ , by for any  $g \in \mathcal{C}_b([0, T] \times I, \mathbb{R})$ :

$$F(g)(t, x) = f \left( v_t(x) + \int_I W(x, y) \int_0^t h(t-s)g(s, y) ds \nu(dy) \right), \quad (t, x) \in [0, T] \times I.$$

We check first that  $F$  takes values in  $\mathcal{C}_b([0, T] \times I, \mathbb{R})$ : consider  $g \in \mathcal{C}_b([0, T] \times I, \mathbb{R})$ . Let us show that  $F(g)$  is **bounded**. Fix  $(t, x) \in [0, T] \times I$ . As  $f$  is Lipschitz continuous and with Hypothesis 3.2.4, we have:

$$F(g)(t, x) \leq f(0) + L_f \|v\|_\infty + L_f \int_{\mathbb{R}^d} |W(x, y)| \int_0^t |h(t-s)|g(s, y) ds \nu(dy).$$

As  $g$  is bounded and  $h$  is locally integrable by Hypothesis 3.2.4, we have

$$\sup_{\substack{t \in [0, T] \\ x \in I}} F(g)(t, x) \leq f(0) + L_f \left( \|v\|_\infty + \|h\|_{[0, T], 1} \|g\|_\infty \sup_{x \in I} \int_{\mathbb{R}^d} W(x, y) \nu(dy) \right) < \infty,$$

where we used Hypothesis (3.2.6).

We check now that  $F(g)$  is **continuous**. We show the sequential continuity: we fix  $(t, x) \in [0, T] \times I$  and a sequence  $(t_n, x_n)$  converging to  $(t, x)$ . As  $f$  is Lipschitz continuous, we have:

$$\begin{aligned} |F(g)(t_n, x_n) - F(g)(t, x)| &\leq L_f |v_{t_n}(x_n) - v_t(x)| \\ &+ L_f \left| \int_I \left( W(x_n, y) \int_0^{t_n} h(t_n - s)g(s, y) ds - W(x, y) \int_0^t h(t - s)g(s, y) ds \right) \nu(dy) \right|. \end{aligned} \quad (3.7.4)$$

The first term  $L_f |v_{t_n}(x_n) - v_t(x)|$  tends to 0 when  $n$  tends to infinity as  $(t, x) \mapsto v_t(x)$  is continuous in time and space by Hypothesis 3.2.4. To show the convergence of the second term, we use the following bound:

$$\begin{aligned} & \left| \int_I W(x_n, y) \int_0^{t_n} h(t_n - s)g(s, y)ds\nu(dy) - \int_I W(x, y) \int_0^t h(t - s)g(s, y)ds\nu(dy) \right| \\ & \leq \left| \int_I (W(x_n, y) - W(x, y)) \int_0^{t_n} h(t_n - s)g(s, y)ds\nu(dy) \right| \\ & + \left| \int_I W(x, y) \left( \int_0^{t_n} h(t_n - s)g(s, y)ds - \int_0^t h(t - s)g(s, y)ds \right) \nu(dy) \right| =: A_n + B_n. \end{aligned}$$

As  $h$  is locally integrable,  $g$  bounded, we can upper bound  $A_n$  immediately:

$$\begin{aligned} A_n & \leq \|h\|_{[0, T], 1} \|g\|_\infty \int_I |W(x_n, y) - W(x, y)| \nu(dy) \\ & \leq \|h\|_{[0, T], 1} \|g\|_\infty C_w \|x - x_n\|^\vartheta \xrightarrow{n \rightarrow \infty} 0, \end{aligned}$$

using (3.2.5). To study the convergence of  $B_n$ , we do a substitution and split the integral in two:

$$\begin{aligned} B_n & \leq \int |W(x, y)| \left| \int_0^{t_n} h(u)g(t_n - u, y)du - \int_0^t h(u)g(t - u, y)ds \right| \nu(dy) \\ & \leq \int |W(x, y)| \left( \int_0^t |h(u)| |g(t_n - u, y) - g(t - u, y)| du \right) \nu(dy) \\ & + \int |W(x, y)| \left( \int_{\min(t, t_n)}^{\max(t, t_n)} |h(u)| |g(t_n - u, y)du \right) \nu(dy) =: a_n + b_n. \end{aligned}$$

Since  $g$  is continuous, we have  $\int_0^t |h(u)| |g(t_n - u, y) - g(t - u, y)| du \xrightarrow{n \rightarrow \infty} 0$  for all  $y \in I$ , and since  $\int_0^t |h(u)| |g(t_n - u, y) - g(t - u, y)| du \leq 2\|h\|_{[0, T], 1} \|g\|_\infty$ , we see from dominated convergence theorem that  $a_n \xrightarrow{n \rightarrow \infty} 0$ . We focus on the term  $b_n$ , we have

$$\begin{aligned} \int_{\min(t, t_n)}^{\max(t, t_n)} |h(u)| |g(t_n - u, y)du & = \int_0^T |h(u)| |g(t_n - u, y)| \mathbf{1}_{[\min(t, t_n), \max(t, t_n)]}(u) du \\ & \leq \|g\|_\infty \int_0^T |h(u)| \mathbf{1}_{[\min(t, t_n), \max(t, t_n)]}(u) du. \end{aligned}$$

Yet  $|h(u)| \mathbf{1}_{[\min(t, t_n), \max(t, t_n)]}(u) \xrightarrow{n \rightarrow \infty} 0$ , and we obtain  $b_n \xrightarrow{n \rightarrow \infty} 0$  by dominated convergence as  $h$  is locally integrable. We have shown that for all  $(t, x) \in [0, T] \times I$ ,  $\lim_{n \rightarrow \infty} |F(g)(t_n, x_n) - F(g)(t, x)| = 0$  for any sequence  $(t_n, x_n)$  tending to  $(t, x)$ :  $F(g)$  is continuous.

We show now that there exists a constant  $C > 0$  such that for all  $(t, x, z) \in [0, T] \times I^2$ :

$$|F(g)(t, x) - F(g)(t, z)| \leq C (\|x - z\| + \|x - z\|^\vartheta). \quad (3.7.5)$$

Let  $(t, x, z) \in [0, T] \times I \times I$ . As done previously ( $f$  and  $v_t$  are Lipschitz continuous with Hypothesis 3.2.4), we have:

$$|F(g)(t, x) - F(g)(t, z)| \leq L_f L_v \|x - z\| + L_f \int_I \int_0^t |h(t-s)| g(s, y) ds |W(x, y) - W(z, y)| \nu(dy).$$

Since  $g$  is bounded,  $h$  is locally integrable, using (3.2.5) we have

$$|F(g)(t, x) - F(g)(t, z)| \leq L_f L_v \|x - z\| + L_f \|g\|_\infty \|h\|_{[0, T], 1} C_w \|x - z\|^\vartheta,$$

which gives (3.7.5).

Hence,  $\mathcal{C}_b([0, T] \times I, \mathbb{R})$  is stable by  $F$ . We prove that  $F$  admits a unique fixed point, which is  $\lambda_t(x) = \lambda(t, x)$  satisfying (3.1.2). To do it, we show that some iteration of  $F$  is contractive, and the Banach fixed-point Theorem gives the result. Let  $t \in [0, T]$ ,  $g$  and  $\tilde{g}$  be two functions in  $\mathcal{C}_b([0, t] \times I, \mathbb{R})$ . We use the distance  $D_t(g, \tilde{g}) := \sup_{s \in [0, t]} \sup_{x \in I} |g(s, x) - \tilde{g}(s, x)|$  which makes the space  $\mathcal{C}_b([0, t] \times I, \mathbb{R})$  complete. Obviously, for all  $s \leq t$ ,  $D_s(g, \tilde{g}) \leq D_t(g, \tilde{g})$ . Let  $x \in \mathbb{R}^d$ . As previously,

$$\begin{aligned} |F(g)(t, x) - F(\tilde{g})(t, x)| &\leq L_f \left| \int_I W(x, y) \int_0^t h(t-s) (g(s, y) - \tilde{g}(s, y)) ds \nu(dy) \right| \\ &\leq L_f \left( \sup_{z \in I} \int_I |W(z, y)| \nu(dy) \right) \int_0^t |h(t-s)| D_s(g, \tilde{g}) ds. \end{aligned}$$

Using Cauchy-Schwarz inequality, as  $h$  is in  $L_{loc}^2$  under Hypothesis 3.2.4,

$$\int_0^t |h(t-s)| D_s(g, \tilde{g}) ds \leq \|h\|_{[0, T], 2} \left( \int_0^t (D_s(g, \tilde{g}))^2 ds \right)^{\frac{1}{2}}.$$

Using (3.2.6), we have then shown the existence of a constant  $C(f, W, \nu, h, T, p)$  such that for all mappings  $g$  and  $\tilde{g}$ , for all  $t \in [0, T]$ :

$$D_t(F(g), F(\tilde{g})) \leq C \left( \int_0^t (D_s(g, \tilde{g}))^2 ds \right)^{\frac{1}{2}}. \quad (3.7.6)$$

By induction on  $k \in \mathbb{N}$ , with (3.7.6), we show that for all  $t \in [0, T]$  and for any mappings  $g$  and  $\tilde{g}$ :  $D_t(F^k(g), F^k(\tilde{g})) \leq C^k \left( \frac{t^k}{k!} \right)^{\frac{1}{2}} D_t(g, \tilde{g})$ . The initialisation is immediate, and then for  $k \geq 0$ , using (3.7.6) and the induction hypothesis

$$\begin{aligned} D_t(F^{k+1}(g), F^{k+1}(\tilde{g})) &\leq C \left( \int_0^t (D_s(F^k(g), F^k(\tilde{g})))^2 ds \right)^{\frac{1}{2}} \\ &\leq C \left( \int_0^t C^{2k} \frac{s^k}{k!} D_s(g, \tilde{g})^2 ds \right)^{\frac{1}{2}} \leq C^{k+1} \left( \frac{t^{k+1}}{(k+1)!} \right)^{\frac{1}{2}} D_t(g, \tilde{g}), \end{aligned}$$

which concludes the induction. We have then that for all  $k$  and any functions  $g$  and  $\tilde{g}$  of  $\mathcal{C}_b([0, T] \times I, \mathbb{R}_+)$ , the  $k$ -th iteration of  $F$  verifies  $D_T(F^k(g), F^k(\tilde{g})) \leq$

$C^k \left( \frac{T^k}{k!} \right)^{\frac{1}{2}} D_T(g, \tilde{g})$ . Hence there exists a rank  $k$  such that  $F^k$  is contractive, thus has a unique fixed point which is also then the unique fixed point of  $F$  in  $\mathcal{C}_b([0, T] \times I, \mathbb{R})$  that we call  $\lambda$ . Furthermore, we have shown that any image by  $F$  verifies the property (3.7.5), so in particular  $\lambda$  verifies it too and (3.2.7) is then true (with  $C_\lambda$  the constant of equation (3.7.5) for  $g = \lambda$ ). Note that such a  $\lambda$  is necessarily nonnegative, as the iterative map  $F$  preserves positivity in both cases  $f \geq 0$  and  $f(x) = x$  with  $v_t, h \geq 0$ .

We focus now on the second part of Theorem 3.2.7: we consider  $(t, x) \mapsto v_t(x)$  continuously differentiable in time and  $(t, x) \mapsto \frac{\partial v_t}{\partial t}(x)$  bounded on  $[0, T] \times I$ ,  $h$  continuous and piecewise continuously differentiable, and  $f(x) = x$ . First, we ensure that (3.2.8) admits a unique continuous bounded solution. Then, by studying a sequence of functions that converges towards  $\lambda$ , we show that  $\lambda$  is differentiable in time and  $\frac{\partial \lambda}{\partial t}$  satisfies (3.2.8). Using the same method as above, we show that the map  $G$  defined on  $\mathcal{C}_b([0, T] \times I, \mathbb{R})$  by

$$\begin{aligned}
 G(g)(t, x) &= \frac{\partial v_t}{\partial t}(x) + h(t) \int_I W(x, y) \lambda(0, y) \nu(dy) \\
 &\quad + \int_I \int_0^t h(t-s) W(x, y) g(s, y) \nu(dy) ds
 \end{aligned}$$

admits a unique fixed point called  $\mu$ . Moreover, we can introduce a sequence of function  $(\mu_n)_n$  that converges uniformly towards  $\mu$  defined by iteration, with  $\mu_0 = 0$  and

$$\begin{aligned}
 \mu_{n+1}(t, x) &:= \frac{\partial v_t}{\partial t}(x) + h(t) \int_I W(x, y) \lambda(0, y) \nu(dy) \\
 &\quad + \int_I \int_0^t h(t-s) W(x, y) \mu_n(s, y) \nu(dy) ds.
 \end{aligned}$$

Similarly, we introduce a sequence of function  $(\lambda_n)_n$  that converges uniformly towards  $\lambda$  defined by iteration, with  $\lambda_0 = 0$  and  $\lambda_{n+1}(t, x) := v_t(x) + \int_I \int_0^t h(t-s) W(x, y) \lambda_n(s, y) \nu(dy) ds$ . By induction, for every  $n$ ,  $\lambda_n$  is differentiable in time and bounded and then, by integration by parts we obtain

$$\begin{aligned}
 \frac{\partial \lambda_{n+1}}{\partial t}(t, x) &= \frac{\partial v_t}{\partial t}(x) + h(t) \int_I W(x, y) \lambda_n(0, y) \nu(dy) \\
 &\quad + \int_0^t \int_I W(x, y) h(t-s) \frac{\partial \lambda_n}{\partial s}(s, y) \nu(dy) ds. \quad (3.7.7)
 \end{aligned}$$

Now, we can compare  $\mu_n$  and  $\frac{\partial \lambda_n}{\partial t}$ : setting  $\varpi_n(t, x) := \mu_n(t, x) - \frac{\partial \lambda_n}{\partial t}(t, x)$  for all

$n$ , for any  $(t, x) \in [0, T] \times I$ ,

$$|\varpi_{n+1}(t, x)| = \left| h(t) \int_I W(x, y) (\lambda(0, y) - \lambda_n(0, y)) \nu(dy) + \int_0^t \int_I W(x, y) h(t-s) \varpi_n(s, y) \nu(dy) ds \right|,$$

so that  $\|\varpi_{n+1}(t, \cdot)\|_\infty \leq h(t) C_W^{(1)} \|\lambda(0, \cdot) - \lambda_n(0, \cdot)\|_\infty + C_W^{(1)} \int_0^t h(t-s) \|\varpi_n(s, \cdot)\|_\infty ds$ . We obtain, as  $(\lambda_n)$  converges uniformly to  $\lambda$ , that  $\limsup_{n \rightarrow \infty} \|\varpi_{n+1}(t, \cdot)\|_\infty \leq C_W^{(1)} \int_0^t h(t-s) \limsup_{n \rightarrow \infty} \|\varpi_n(s, \cdot)\|_\infty ds$ . This gives, from Lemma 3.A.3 (i) and provided that one has verified that  $\limsup_{n \rightarrow \infty} \|\varpi_n(s, \cdot)\|_\infty < +\infty$  is finite,  $\sup_{t \in [0, T]} \limsup_{n \rightarrow \infty} \|\varpi_n(t, \cdot)\|_\infty = 0$ . It implies that, as  $(\mu_n)$  converges uniformly to  $\mu$ , so does  $\left(\frac{\partial \lambda_n}{\partial t}\right)_n$ , and then as  $\lambda$  is differentiable,  $\lambda_n \xrightarrow[n \rightarrow \infty]{\text{uniformly}} \lambda$  and  $\frac{\partial \lambda_n}{\partial t} \xrightarrow[n \rightarrow \infty]{\text{uniformly}} \mu$ , we obtain  $\frac{\partial \lambda}{\partial t} = \mu$ . It remains to check that  $\limsup_{n \rightarrow \infty} \|\varpi_n(s, \cdot)\|_\infty$  is indeed finite. As  $(\lambda_n)$  converges to  $\lambda$ , it is uniformly bounded and as  $(t, x) \mapsto \frac{\partial v_t}{\partial t}(x)$  is bounded, we can find  $g$  locally bounded such that, from (3.7.7),  $\|\partial_t \lambda_{n+1}(t, \cdot)\|_\infty \leq g(t) + C_W^{(1)} \int_0^t h(t-s) \|\partial_s \lambda_n(s, \cdot)\|_\infty ds$ . Lemma 3.A.3 (iii) gives then that  $\sup_{s \in [0, T]} \sup_{n \geq 0} \|\partial_s \lambda_n(s, \cdot)\|_\infty < \infty$ . We can do the same for  $(\mu_n)$  and obtain  $\sup_{s \in [0, T]} \sup_{n \geq 0} \|\mu_n(s, \cdot)\|_\infty < \infty$  which concludes to  $\limsup_{n \rightarrow \infty} \|\varpi_n(s, \cdot)\|_\infty < \infty$  for any  $s \in [0, T]$ .  $\square$

## 3.8 Convergence of the mean-field process (proofs)

### 3.8.1 Toolbox

We present useful results that come up in the main proofs.

**Proposition 3.8.1.** *Recall the definitions of  $\kappa_N$  and  $w_N$  in Hypothesis 3.3.7. Let  $(\alpha_{ij})$  and  $(\alpha_{ijk})$  such that for every  $(i, j, l) \in \llbracket 1, N \rrbracket^3$ ,  $|\alpha_{ij}| \leq 1$  and  $|\alpha_{ijl}| \leq 1$ . Define*

$$X_j := \frac{\kappa_N}{N} \sum_{l=1}^N \alpha_{lj} \bar{\xi}_{lj}, \quad \tilde{X}_i := \frac{\kappa_N}{N} \sum_{l=1}^N \alpha_{il} \bar{\xi}_{il}, \quad X_{ij} := \frac{\kappa_N}{N} \sum_{l=1}^N \alpha_{ijl} \bar{\xi}_{il},$$

with  $\bar{\xi}_{lk} := \xi_{lk}^{(N)} - W_N(x_l, x_k)$ . Then, under Hypothesis 3.3.7,  $\mathbb{P}$ -almost surely if  $N$  is large enough:

$$\sup_{1 \leq j \leq N} |X_j| \leq \varepsilon_N, \quad \sup_{1 \leq i \leq N} |\tilde{X}_i| \leq \varepsilon_N \quad \text{and} \quad \sup_{1 \leq i, j \leq N} |X_{ij}| \leq \varepsilon_N,$$

for  $\varepsilon_N := 32 \frac{\kappa_N^2 w_N}{N} \log(N)$ .

Note that under Hypothesis 3.3.7,  $\varepsilon_N \xrightarrow[N \rightarrow \infty]{} 0$ .

*Proof.* We rely on Lemma 3.A.7. We derive a uniform bound on  $(X_j)_{j \in [1, N]}$ : fixing  $j$ , we apply Lemma 3.A.7 for the choice  $U_l = \xi_{lj}^{(N)}$ ,  $p_l = W_N(x_l, x_j)$  (note that (3.3.9) yields that  $p_l \leq w_N$ ),  $d_l = \alpha_{lj}$  and the constant  $\kappa_N > 0$ . We obtain, taking the supremum on  $j$  and a union bound:

$$\mathbb{P} \left( \sup_{j \in [1, N]} |X_j| > \varepsilon_N \right) \leq 2N \exp \left( -16 \log(N) B \left( 4\sqrt{2} \left( \frac{\log(N)}{Nw_N} \right)^{\frac{1}{2}} \right) \right).$$

As  $B(u) = u^{-2} ((1+u) \log(1+u) - u) \rightarrow \frac{1}{2}$  when  $u \rightarrow 0$  and  $\frac{\log(N)}{Nw_N} \leq \frac{\log(N)}{N} \kappa_N^2 w_N \rightarrow 0$  when  $N \rightarrow \infty$  using (3.3.10) and (3.3.11), we can choose a

deterministic  $p$  such that for all  $N \geq p$ ,  $B \left( 4\sqrt{2} \left( \frac{\log(N)}{Nw_N} \right)^{\frac{1}{2}} \right) \geq \frac{3}{16}$ . We then

have if  $N \geq p$ :  $\mathbb{P}(\sup_{j \in [1, N]} |X_j| > \varepsilon_N) \leq 2N \exp(-3 \log(N)) = \frac{2}{N^2}$ . Hence,

by Borel-Cantelli Lemma, there exists  $\tilde{\mathcal{O}} \in \mathcal{F}$  such that  $\mathbb{P}(\tilde{\mathcal{O}}) = 1$  and on  $\tilde{\mathcal{O}}$ , there exists  $\tilde{N} < \infty$  such that if  $N \geq \tilde{N}$ ,  $\sup_{j \in [1, N]} |X_j| \leq \varepsilon_N$ . We can show

similarly that  $\sup_{1 \leq i \leq N} |\tilde{X}_i| \leq \varepsilon_N$ . To show the result on  $(X_{ij})$ , we use the same

Lemma 3.A.7 but we need to lower-bound  $B \left( 4\sqrt{2} \left( \frac{\log(N)}{Nw_N} \right)^{\frac{1}{2}} \right)$  differently: we

can choose a deterministic  $\tilde{p}$  for all  $N \geq \tilde{p}$ ,  $B \left( 4\sqrt{2} \left( \frac{\log(N)}{Nw_N} \right)^{\frac{1}{2}} \right) \geq \frac{1}{4}$  and

then the same argument as before works to obtain  $\mathbb{P}(\sup_{i, j \in [1, N]} |X_{ij}| > \varepsilon_N) \leq 2N^2 \exp \left( -16 \log(N) \frac{1}{4} \right) \leq \frac{2}{N^2}$ , and we conclude by Borel-Cantelli Lemma.  $\square$

**Corollary 3.8.2.** *Under Hypothesis 3.3.7, we have  $\mathbb{P}$ -almost surely if  $N$  is large enough:*

$$\sup_{1 \leq j \leq N} \left( \sum_{i=1}^N \frac{\kappa_i^{(N)}}{N} \xi_{ij}^{(N)} \right) \leq 1 + \sup_{1 \leq j \leq N} \left( \sum_{i=1}^N \frac{\kappa_i^{(N)}}{N} W_N(x_i, x_j) \right) \quad (3.8.1)$$

$$\sup_{1 \leq i \leq N} \left( \sum_{j=1}^N \frac{\kappa_i^{(N)}}{N} \xi_{ij}^{(N)} \right) \leq 1 + \sup_{1 \leq i \leq N} \left( \sum_{j=1}^N \frac{\kappa_i^{(N)}}{N} W_N(x_i, x_j) \right) \quad (3.8.2)$$

$$\frac{1}{N^3} \sum_{i, j=1}^N \left( \kappa_i^{(N)} \right)^2 \xi_{ij}^{(N)} \leq \frac{\kappa_N}{N} \left( 1 + \sum_{i, j=1}^N \frac{\kappa_i^{(N)}}{N^2} W_N(x_i, x_j) \right). \quad (3.8.3)$$

*Proof.* It is a direct application of Proposition 3.8.1 (as for any  $i$ ,  $\left| \frac{\kappa_i^{(N)}}{\kappa_N} \right| \leq 1$  with (3.3.8)), as  $\varepsilon_N$  defined in Proposition 3.8.1 tends to 0 as  $N \rightarrow \infty$  under

Hypothesis 3.3.7, hence for  $N$  large enough  $\varepsilon_N \leq 1$ . For instance for (3.8.3), with

$$X_j = \frac{\kappa_N}{N} \sum_{i=1}^N \left( \frac{\kappa_i^{(N)}}{\kappa_N} \right)^2 \left( \xi_{ij}^{(N)} - W_N(x_i, x_j) \right) \text{ we have}$$

$$\frac{1}{N^3} \sum_{i,j=1}^N \left( \kappa_i^{(N)} \right)^2 \xi_{ij}^{(N)} \leq \frac{\kappa_N}{N^2} \sum_{j=1}^N X_j + \frac{\kappa_N}{N^3} \sum_{i,j=1}^N \kappa_i^{(N)} W_N(x_i, x_j) \leq \frac{\kappa_N}{N} (\varepsilon_N + C_W),$$

where we used Proposition 3.8.1 and (3.3.12).  $\square$

We introduce the following auxiliary graph.

**Definition 3.8.3.** We denote by  $\mathcal{G}_N^{(2)}$  the directed weighted graph with vertices  $\{1, \dots, N\}$  such that every edge  $j \rightarrow i$  is present, and with weight  $W(x_i, x_j)$ .

The proof of the following technical Proposition is postponed in Section 3.8.4.

**Proposition 3.8.4.** *Under the Scenarios of Definition 3.3.8,*

$$d_{\square, \nu} \left( W^{\mathcal{G}_N^{(2)}}, W \right) \xrightarrow[N \rightarrow \infty]{} 0, \quad (3.8.4)$$

$$\|W^{\mathcal{G}_N^{(2)}} - W\|_{\infty \rightarrow \infty, \nu} \xrightarrow[N \rightarrow \infty]{} 0, \quad (3.8.5)$$

$$\max_{1 \leq i \leq N} \int_0^T \left| \int_I W(x_i, x) \gamma(s, x) (\nu^{(N)}(dx) - \nu(dx)) \right| ds \xrightarrow[N \rightarrow \infty]{} 0 \quad (3.8.6)$$

and

$$\frac{1}{N} \sum_{i=1}^N \int_0^T \left| \int_I W(x_i, x) \gamma(s, x) (\nu^{(N)}(dx) - \nu(dx)) \right| ds \xrightarrow[N \rightarrow \infty]{} 0, \quad (3.8.7)$$

where  $\gamma(s, x) := \int_0^s h(s-u) \lambda_u(x) du$ .

If (3.8.4) and (3.8.7) are satisfied in another configuration of positions than in the Scenarios of Definition 3.3.8, Theorem 3.3.10 still applies. Likewise, if (3.8.5) and (3.8.6) are satisfied, Theorem 3.3.12 still applies.

### 3.8.2 Proof of Theorem 3.3.10

Recall the definitions of  $Z_{N,i}$  and  $\bar{Z}_i$  in (3.2.2) and (3.3.1). We remind that we consider the sequences  $(\underline{x}^{(N)})_{N \geq 1}$  and  $(\xi_{ij}^{(N)})_{\substack{N \geq 1 \\ i,j \in \llbracket 1, N \rrbracket}}$  fixed (our result is quenched).

Let  $t \in [0, T]$ . For each  $i \in \llbracket 1, N \rrbracket$ , let  $\Delta_i^{(N)}(t)$  be the total variation distance between  $Z_{N,i}$  and  $\bar{Z}_i$  on  $[0, t]$ :

$$\Delta_i^{(N)}(t) = \int_0^t |d(Z_{N,i}(s) - \bar{Z}_i(s))|. \quad (3.8.8)$$



Remark that we always have  $\sup_{t \in [0, T]} |Z_{N,i}(t) - \bar{Z}_i(t)| \leq \Delta_i^{(N)}(T)$ . We have taking the expectation

$$\begin{aligned} \mathbf{E} \left[ \Delta_i^{(N)}(t) \right] &= \mathbf{E} \left[ \int_0^t \int_0^\infty |\mathbf{1}_{\{z \leq \lambda_{N,i}(s)\}} - \mathbf{1}_{\{z \leq \lambda_s(x_i)\}}| \pi_i(ds, dz) \right] \\ &= \int_0^t \mathbf{E} [|\lambda_{N,i}(s) - \lambda_s(x_i)|] ds. \end{aligned}$$

Using the Lipschitz continuity of  $f$  and recalling the definition of  $\lambda_{N,i}$  in (3.2.3) and of  $\lambda$  in (3.1.2), we obtain:

$$\mathbf{E} \left[ \Delta_i^{(N)}(t) \right] \leq L_f \left( \sum_{k=1}^5 A_{i,t,k}^{(N)} \right), \quad (3.8.9)$$

$$\text{where } A_{i,t,1}^{(N)} := \int_0^t \mathbf{E} \left[ \left| \frac{\kappa_i^{(N)}}{N} \sum_{j=1}^N \xi_{ij}^{(N)} \int_0^{s-} h(s-u) (dZ_{N,j}(u) - d\bar{Z}_j(u)) \right| \right] ds, \quad (3.8.10)$$

$$A_{i,t,2}^{(N)} := \int_0^t \mathbf{E} \left[ \left| \frac{\kappa_i^{(N)}}{N} \sum_{j=1}^N \xi_{ij}^{(N)} \int_0^{s-} h(s-u) (d\bar{Z}_j(u) - \lambda_u(x_j) du) \right| \right] ds, \quad (3.8.11)$$

$$A_{i,t,3}^{(N)} := \int_0^t \left| \frac{\kappa_i^{(N)}}{N} \sum_{j=1}^N \left( \xi_{ij}^{(N)} - W_N(x_i, x_j) \right) \int_0^s h(s-u) \lambda_u(x_j) du \right| ds, \quad (3.8.12)$$

$$A_{i,t,4}^{(N)} := \int_0^t \left| \frac{1}{N} \sum_{j=1}^N \left( \kappa_i^{(N)} W_N(x_i, x_j) - W(x_i, x_j) \right) \int_0^s h(s-u) \lambda_u(x_j) du \right| ds \text{ and} \quad (3.8.13)$$

$$\begin{aligned} A_{i,t,5}^{(N)} &:= \int_0^t \left| \frac{1}{N} \sum_{j=1}^N W(x_i, x_j) \int_0^s h(s-u) \lambda_u(x_j) du \right. \\ &\quad \left. - \int_I W(x_i, y) \int_0^s h(s-u) \lambda_u(y) du \nu(dy) \right| ds. \end{aligned} \quad (3.8.14)$$

We are going to control each term  $\frac{1}{N} \sum_{i=1}^N A_{i,t,k}^{(N)}$ . The term  $A_{i,t,1}^{(N)}$  captures the proximity between the particle system  $Z_{N,i}$  with its mean-field counterpart  $\bar{Z}_i$  at the same position. We have, as the graph  $(\xi^{(N)})$  is fixed,

$$\begin{aligned} A_{i,t,1}^{(N)} &= \int_0^t \mathbf{E} \left[ \left| \frac{\kappa_i^{(N)}}{N} \sum_{j=1}^N \xi_{ij}^{(N)} \int_0^{s-} h(s-u) (dZ_{N,j}(u) - d\bar{Z}_j(u)) \right| \right] ds \\ &\leq \frac{1}{N} \sum_{j=1}^N \kappa_i^{(N)} \xi_{ij}^{(N)} \mathbf{E} \left[ \int_0^t \int_0^{s-} |h(s-u)| \left| d \left( \Delta_j^{(N)}(u) \right) \right| ds \right]. \end{aligned}$$

We use Lemma 3.A.1 so that

$$A_{i,t,1}^{(N)} \leq \frac{1}{N} \sum_{j=1}^N \kappa_i^{(N)} \xi_{ij}^{(N)} \mathbf{E} \left[ \int_0^t |h(t-s)| \Delta_j^{(N)}(s) ds \right], \quad (3.8.15)$$

then we have, by summation:

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N A_{i,t,1}^{(N)} &\leq \frac{1}{N} \sum_{j=1}^N \left( \sum_{i=1}^N \frac{\kappa_i^{(N)}}{N} \xi_{ij}^{(N)} \right) \int_0^t |h(t-s)| \mathbf{E} \left[ \Delta_j^{(N)}(s) \right] ds \\ &\leq \sup_{1 \leq j \leq N} \left( \sum_{i=1}^N \frac{\kappa_i^{(N)}}{N} \xi_{ij}^{(N)} \right) \int_0^t |h(t-s)| \mathbf{E} \left[ \frac{1}{N} \sum_{j=1}^N \Delta_j^{(N)}(s) \right] ds. \end{aligned}$$

We use (3.8.1) and (3.3.14) to obtain  $\mathbb{P}$ -almost surely if  $N$  is large enough the bound

$$\frac{1}{N} \sum_{i=1}^N A_{i,t,1}^{(N)} \leq (1 + C_W) \int_0^t |h(t-s)| \mathbf{E} \left[ \frac{1}{N} \sum_{j=1}^N \Delta_j^{(N)}(s) \right] ds. \quad (3.8.16)$$

The second term  $A_{i,t,2}^{(N)}$  captures the proximity between the limit process and its expectation. We have that

$$A_{i,t,2}^{(N)} = \int_0^t \mathbf{E} \left[ \left| \frac{1}{N} \sum_{j=1}^N (V_j^i(s) - \mathbf{E} [V_j^i(s)]) \right| \right] ds,$$

where  $V_j^i(s) = \kappa_i^{(N)} \xi_{ij}^{(N)} \int_0^{s-} h(s-u) d\bar{Z}_j(u)$  is a family of independent random variables (by independence of the  $\pi_i$ ). Note that  $\mathbf{E} [V_j^i(s)] = \kappa_i^{(N)} \xi_{ij}^{(N)} \int_0^s h(s-u) \lambda_u(x_j) du$ . Define  $M_j^i(s) := V_j^i(s) - \mathbf{E} [V_j^i(s)]$ , which can also be written as

$$\begin{aligned} M_j^i(s) &= \int_0^s \int_0^\infty \mathbf{1}_{\{z \leq \lambda_u(x_j)\}} \kappa_i^{(N)} \xi_{ij}^{(N)} h(s-u) \pi_i(du, dz) \\ &\quad - \int_0^s \int_0^\infty \mathbf{1}_{\{z \leq \lambda_u(x_j)\}} \kappa_i^{(N)} \xi_{ij}^{(N)} h(s-u) dudz, \end{aligned}$$

so that

$$\begin{aligned} \text{Var} (V_j^i(s)) &= \mathbf{E} [M_j^i(s)^2] = \mathbf{E} \left[ \int_0^s \int_0^\infty \left( \mathbf{1}_{\{z \leq \lambda_u(x_j)\}} \kappa_i^{(N)} \xi_{ij}^{(N)} h(s-u) \right)^2 dudz \right] \\ &= \int_0^s \left( \kappa_i^{(N)} \right)^2 \xi_{ij}^{(N)} h(s-u)^2 \lambda_u(x_j) du. \end{aligned}$$

Thus summing on  $i$  and using Lemma 3.A.2, we have

$$\frac{1}{N} \sum_{i=1}^N A_{i,t,2}^{(N)} \leq \frac{1}{N} \sum_{i=1}^N \int_0^t \frac{1}{N} \sqrt{\sum_{j=1}^N \int_0^s \left( \kappa_i^{(N)} \right)^2 \xi_{ij}^{(N)} h(s-u)^2 \lambda_u(x_j) du} ds.$$

We apply Jensen's inequality to both uniform measures on  $\{1, \dots, N\}$  and  $[0, t]$  to obtain:

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N A_{i,t,2}^{(N)} &\leq \frac{t}{N} \int_0^t \sqrt{\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N \int_0^s \left( \kappa_i^{(N)} \right)^2 \xi_{ij}^{(N)} h(s-u)^2 \lambda_u(x_j) du} \frac{ds}{t} \\ &\leq \frac{t}{N} \sqrt{\frac{1}{Nt} \sum_{i,j=1}^N \left( \kappa_i^{(N)} \right)^2 \xi_{ij}^{(N)} \int_0^t \int_0^s h(s-u)^2 \lambda_u(x_j) duds} \end{aligned}$$

By Hypothesis 3.2.4 on  $h$ , we have

$$\frac{1}{N} \sum_{i=1}^N A_{i,t,2}^{(N)} \leq t \|h\|_{t,2} \sqrt{\|\lambda\|_{[0,t] \times I, \infty}} \sqrt{\frac{1}{N^3} \sum_{i,j=1}^N \left(\kappa_i^{(N)}\right)^2 \xi_{ij}^{(N)}}.$$

We use (3.8.3) and (3.3.12) to obtain  $\mathbb{P}$ -almost surely if  $N$  is large enough the bound:

$$\frac{1}{N} \sum_{i=1}^N A_{i,t,2}^{(N)} \leq t \|h\|_{t,2} \sqrt{\|\lambda\|_{[0,t] \times I, \infty}} \sqrt{\frac{\kappa_N}{N}} (1 + C_W). \quad (3.8.17)$$

The term  $A_{i,t,3}^{(N)}$  captures the proximity between the realization of the graph  $(\xi^{(N)})$  and its expectation. We define for  $(s, x, y) \in [0, T] \times I \times I$ :

$$\gamma(s, x) := \int_0^s h(s-u) \lambda_u(x) du, \quad (3.8.18)$$

$$\Gamma_T(x, y) := \int_0^T \gamma(s, x) \gamma(s, y) ds. \quad (3.8.19)$$

Note that we always have  $|\gamma(s, x)| \leq \|h\|_{s,1} \|\lambda\|_{[0,s] \times I, \infty} =: \gamma_{s,\infty}$  and  $0 \leq \Gamma_T(x, y) \leq T \gamma_{T,\infty}^2$ . Recall that  $\bar{\xi}_{ij} := \xi_{ij}^{(N)} - W_N(x_i, x_j)$ . Then

$$A_{i,t,3}^{(N)} = \int_0^t \left| \frac{\kappa_i^{(N)}}{N} \sum_{j=1}^N \bar{\xi}_{ij} \gamma(s, x_j) \right| ds \leq \int_0^T \left| \frac{\kappa_i^{(N)}}{N} \sum_{j=1}^N \bar{\xi}_{ij} \gamma(s, x_j) \right| ds.$$

Note that one cannot apply Proposition 3.8.1 directly in the integrand since we would not get an a.s. result. Therefore, we control its square, by Jensen's inequality:

$$A_{i,t,3}^{(N)2} \leq T \int_0^T \left( \frac{\kappa_i^{(N)}}{N} \sum_{j=1}^N \bar{\xi}_{ij} \gamma(s, x_j) \right)^2 ds = T^2 \gamma_{T,\infty}^2 \frac{\kappa_i^{(N)}}{N} \sum_{j=1}^N \bar{\xi}_{ij} X_{ij},$$

where we set  $X_{ij} := \frac{\kappa_i^{(N)}}{N} \sum_{l=1}^N \frac{\Gamma_T(x_j, x_l)}{\gamma_{T,\infty}^2}$ . Now, by Proposition 3.8.1,  $\mathbb{P}$ -almost surely for  $N$  large enough,  $\sup_{1 \leq i,j \leq N} |X_{ij}| \leq \varepsilon_N$ , thus

$$\begin{aligned} A_{i,t,3}^{(N)2} &\leq T^2 \gamma_{T,\infty}^2 \frac{\kappa_i^{(N)}}{N} \sum_{j=1}^N \left( \xi_{ij}^{(N)} + W_N(x_i, x_j) \right) \sup_{i,j} |X_{ij}| \\ &\leq T^2 \gamma_{T,\infty}^2 \varepsilon_N \frac{\kappa_i^{(N)}}{N} \sum_{j=1}^N \left( \xi_{ij}^{(N)} + W_N(x_i, x_j) \right). \end{aligned} \quad (3.8.20)$$

Taking the square root then summing on  $i$ , we use the discrete Jensen's inequality to obtain

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N A_{i,t,3}^{(N)} &\leq \sqrt{\varepsilon_N} T \gamma_{T,\infty} \frac{1}{N} \sum_{i=1}^N \sqrt{\frac{\kappa_i^{(N)}}{N} \sum_{j=1}^N \left( \xi_{ij}^{(N)} + W_N(x_i, x_j) \right)} \\ &\leq \sqrt{\varepsilon_N} T \gamma_{T,\infty} \sqrt{\sum_{i=1}^N \frac{\kappa_i^{(N)}}{N^2} \sum_{j=1}^N \left( \xi_{ij}^{(N)} + W_N(x_i, x_j) \right)}, \end{aligned}$$

if  $N$  is large enough  $\mathbb{P}$ -almost surely. Using (3.8.2) and (3.3.12), we have

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N A_{i,t,3}^{(N)} &\leq \sqrt{\varepsilon_N T} \gamma_{T,\infty} \sqrt{\frac{1}{N} \sum_{i=1}^N \frac{\kappa_i^{(N)}}{N} \sum_{j=1}^N \xi_{ij}^{(N)} + \frac{1}{N} \sum_{i=1}^N \frac{\kappa_i^{(N)}}{N} \sum_{j=1}^N W_N(x_i, x_j)} \\ &\leq \sqrt{\varepsilon_N T} \gamma_{T,\infty} \sqrt{1 + C_W + \sup_{i \in [1, N]} \left( \frac{\kappa_i^{(N)}}{N} \sum_{j=1}^N W_N(x_i, x_j) \right)} \end{aligned} \quad (3.8.21)$$

$$\leq \sqrt{\varepsilon_N T} \gamma_{T,\infty} \sqrt{1 + 2C_W}. \quad (3.8.22)$$

The term  $A_{i,t,4}^{(N)}$  captures the proximity between the law of the graph on  $N$  particles and the limit graphon  $W$ . Recall the definition of  $\gamma$  in (3.8.18) and the graphs introduced in Definitions 3.3.6 and 3.8.3. Denoting by  $c(s) = (c_j(s))_{1 \leq j \leq N} = \left( \frac{\gamma(s, x_j)}{\gamma_{t,\infty}} \right)_{1 \leq j \leq N} \in [-1, 1]^N$ , we obtain using (3.3.7) and introducing for any  $c = (c_1, \dots, c_N) \in [-1, 1]^N$  the step function  $g^c(v) = \sum_{l=1}^N c_l \mathbf{1}_{B_l}(v)$  for  $v \in I$ , after summation:

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N A_{i,t,4}^{(N)} &= \frac{\gamma_{t,\infty}}{N^2} \int_0^t \sum_{i=1}^N \left| \sum_{j=1}^N \left( \kappa_i^{(N)} W_N(x_i, x_j) - W(x_i, x_j) \right) c_j(s) \right| ds \\ &= \gamma_{t,\infty} \int_0^t \int \left| \int \left( W^{\mathcal{G}_N^{(1)}}(u, v) - W^{\mathcal{G}_N^{(2)}}(u, v) \right) g^{c(s)}(v) \nu(dv) \right| \nu(du) ds \\ &\leq T \gamma_{T,\infty} \|W^{\mathcal{G}_N^{(1)}} - W^{\mathcal{G}_N^{(2)}}\|_{\infty \rightarrow 1, \nu}, \end{aligned}$$

where  $\|\cdot\|_{\infty \rightarrow 1, \nu}$  is defined in (3.3.3). Hence, with Remark 3.3.3 we obtain:

$$\frac{1}{N} \sum_{i=1}^N A_{i,t,4}^{(N)} \leq 4T \gamma_{T,\infty} \left( d_{\square, \nu} \left( W^{\mathcal{G}_N^{(1)}}, W \right) + d_{\square, \nu} \left( W^{\mathcal{G}_N^{(2)}}, W \right) \right). \quad (3.8.23)$$

We use (3.8.4) to deal with  $d_{\square, \nu} \left( W^{\mathcal{G}_N^{(2)}}, W \right)$ .

The term  $\frac{1}{N} \sum_{i=1}^N A_{i,t,5}^{(N)}$  captures the proximity between the empirical measure of the positions of  $N$  particles  $\mu^{(N)}$  and its limit  $\nu$ . We control  $A_{t,5}^{(N)} := \frac{1}{N} \sum_{i=1}^N A_{i,t,5}^{(N)}$  with (3.8.7). Combining (3.8.16), (3.8.17), (3.8.21) and (3.8.23), we obtain if  $N$  is large enough  $\mathbb{P}$ -almost surely for every  $t \in [0, T]$ :

$$\begin{aligned} \mathbf{E} \left[ \frac{1}{N} \sum_{i=1}^N \Delta_i^{(N)}(t) \right] &\leq C_1 \int_0^t |h(t-s)| \mathbf{E} \left[ \frac{1}{N} \sum_{j=1}^N \Delta_j^{(N)}(s) \right] ds + C_2 \sqrt{\frac{\kappa_N}{N}} + C_3 \sqrt{\varepsilon_N} \\ &\quad + C_4 d_{\square, \nu} \left( W^{\mathcal{G}_N^{(1)}}, W \right) + C_4 \gamma_{T,\infty} d_{\square, \nu} \left( W^{\mathcal{G}_N^{(2)}}, W \right) + L_f A_{t,5}^{(N)}, \end{aligned} \quad (3.8.24)$$

with  $C_1, C_2, C_3, C_4$  constants depending on  $L_f, C_W, h$  and  $T$ . Recall that the total variation  $\sup_{t \in [0, T]} |Z_{N,i}(t) - \bar{Z}_i(t)| \leq \Delta_i^{(N)}(T)$ , we apply Lemma 3.A.5 with

$u(t) = \mathbf{E} \left[ \frac{1}{N} \sum_{i=1}^N \Delta_i^{(N)}(t) \right]$  on  $[0, T]$  to obtain  $\mathbb{P}$ -almost surely on the realization of  $(\xi^{(N)})$  if  $N$  is large enough:

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \mathbf{E} \left[ \sup_{t \in [0, T]} \left| Z_i^{(N)}(t) - \bar{Z}_i(t) \right| \right] &\leq C \left( \sqrt{\frac{\kappa_N}{N}} + \sqrt{\varepsilon_N} + d_{\square, \nu} \left( W^{\mathcal{G}_N^{(1)}}, W \right) \right. \\ &\quad \left. + d_{\square, \nu} \left( W^{\mathcal{G}_N^{(2)}}, W \right) + A_{T, 5}^{(N)} \right), \end{aligned} \quad (3.8.25)$$

with  $C = \sqrt{2} \max(C_2, C_3, C_4, L_f) \exp(C_1^2 \|h\|_{T, 2}^2 T)$ . By (3.3.11),  $\lim_{N \rightarrow \infty} \varepsilon_N = 0$ ,  $\lim_{N \rightarrow \infty} \frac{\kappa_N}{N} = 0$  and by (3.3.13)  $d_{\square, \nu} \left( W^{\mathcal{G}_N^{(1)}}, W \right) \xrightarrow{N \rightarrow \infty} 0$ . Combined with Proposition 3.8.4, we conclude the proof of (3.3.15).  $\square$

### 3.8.3 Proof of Theorem 3.3.12

It is almost the same as for Theorem 3.3.10 with changes due to the fact that we take now the maximum on  $i$ . Let us go back to the inequality (3.8.9). We are going to control each term  $\max_{1 \leq i \leq N} A_{i, t, k}^{(N)}$ . Concerning  $A_{i, t, 1}^{(N)}$ , the same estimate (3.8.15) as in the proof of Theorem 3.3.10 leads now to

$$\max_{1 \leq i \leq N} A_{i, t, 1}^{(N)} \leq \max_{1 \leq i \leq N} \left( \frac{1}{N} \sum_{j=1}^N \kappa_i^{(N)} \xi_{ij}^{(N)} \right) \int_0^t |h(t-s)| \max_{1 \leq i \leq N} \mathbf{E} \left[ \Delta_i^{(N)}(s) \right] ds.$$

We use (3.8.2) and (3.3.12) to obtain  $\mathbb{P}$ -almost surely if  $N$  is large enough:

$$\max_{1 \leq i \leq N} A_{i, t, 1}^{(N)} \leq (1 + C_W) \int_0^t |h(t-s)| \max_{1 \leq i \leq N} \mathbf{E} \left[ \Delta_i^{(N)}(s) \right] ds. \quad (3.8.26)$$

Note that here, we do not use the same control as is the proof of Theorem 3.3.10, we only need the uniformly bounded indegree. Concerning  $A_{i, t, 2}^{(N)}$ , we obtain as in the proof of Theorem 3.3.10

$$A_{i, t, 2}^{(N)} \leq \int_0^t \frac{1}{N} \sqrt{\sum_{j=1}^N \int_0^s \left( \kappa_i^{(N)} \right)^2 \xi_{ij}^{(N)} h(s-u)^2 \lambda_u(x_j) du} ds.$$

We use Jensen's inequality on the probability measure  $\frac{1}{t} dt$  on  $[0, t]$  and then the boundedness of  $h$  and  $\lambda$  to obtain

$$\begin{aligned} A_{i, t, 2}^{(N)} &\leq \frac{t}{N} \int_0^t \sqrt{\sum_{j=1}^N \int_0^s \left( \kappa_i^{(N)} \right)^2 \xi_{ij}^{(N)} h(s-u)^2 \lambda_u(x_j) du} \frac{ds}{t} \\ &\leq \frac{t}{N \sqrt{t}} \kappa_i^{(N)} \sqrt{\sum_{j=1}^N \xi_{ij}^{(N)} \int_0^t \int_0^s h(s-u)^2 \lambda_y(x_j) du ds} \\ &\leq \|h\|_{t, 2} \sqrt{\|\lambda\|_{[0, t] \times \mathbb{R}^d, \infty}} \frac{\sqrt{t}}{N} \kappa_i^{(N)} \sqrt{\sum_{j=1}^N \xi_{ij}^{(N)}}, \end{aligned}$$

and taking the maximum leads to

$$\max_{1 \leq i \leq N} A_{i,t,2}^{(N)} \leq \sqrt{\frac{\kappa_N}{N}} \sqrt{\max_{1 \leq i \leq N} \frac{\kappa_i^{(N)}}{N} \sum_{j=1}^N \xi_{ij}^{(N)} \|h\|_{t,2} \sqrt{t \|\lambda\|_{[0,t] \times \mathbb{R}^d, \infty}}}.$$

Using as before (3.8.2) and (3.3.12), we obtain  $\mathbb{P}$ -almost surely if  $N$  is large enough

$$\max_{1 \leq i \leq N} A_{i,t,2}^{(N)} \leq \sqrt{\frac{\kappa_N}{N}} \sqrt{1 + C_W} \|h\|_{t,2} \sqrt{t \|\lambda\|_{[0,t] \times \mathbb{R}^d, \infty}}. \quad (3.8.27)$$

Concerning  $A_{i,t,3}^{(N)}$ , we obtain as in the proof of Theorem 3.3.10 (see (3.8.20), with (3.8.2) and (3.3.12)) that  $\mathbb{P}$ -almost surely

$$A_{i,t,3}^{(N)} \leq T \gamma_{T,\infty} \sqrt{\varepsilon_N \frac{\kappa_i^{(N)}}{N} \sum_{j=1}^N \left( \xi_{ij}^{(N)} + W_N(x_i, x_j) \right)},$$

hence taking the maximum and using (3.8.2) and (3.3.12), we obtain  $\mathbb{P}$ -almost surely if  $N$  is large enough

$$\max_{1 \leq i \leq N} A_{i,t,3}^{(N)} \leq T \gamma_{T,\infty} \sqrt{\varepsilon_N} \sqrt{1 + 2C_W}. \quad (3.8.28)$$

Concerning  $A_{i,t,4}^{(N)}$ , we recognise

$$A_{i,t,4}^{(N)} = \gamma_{t,\infty} \int_0^t \left| \frac{1}{N} \sum_{j=1}^N \left( \kappa_i^{(N)} W_N(x_i, x_j) - W(x_i, x_j) \right) c_j(s) \right| ds.$$

We obtain, using Definitions 3.3.6 and 3.8.3 with Lemma 3.3.5 that as

$$\begin{aligned} & \sup_{1 \leq i \leq N} \left| \frac{1}{N} \sum_{j=1}^N \left( \kappa_i^{(N)} W_N(x_i, x_j) - W(x_i, x_j) \right) c_j(s) \right| \\ &= \sup_{u \in I} \left| \sum_{i=1}^N \mathbf{1}_{B_{N,i}}(u) \int_I \sum_{j=1}^N \left( \kappa_i^{(N)} W_N(x_i, x_j) - W(x_i, x_j) \right) \mathbf{1}_{B_{N,j}}(v) c_j(s) \nu(dv) \right| \\ &= \sup_{u \in I} \left| \int \left( W^{\mathcal{G}_N^{(1)}}(u, v) - W^{\mathcal{G}_N^{(2)}}(u, v) \right) g^{c(s)}(v) \nu(dv) \right|, \end{aligned}$$

we have

$$\begin{aligned} \sup_{1 \leq i \leq N} A_{i,t,4}^{(N)} &= \gamma_{t,\infty} \sup_{1 \leq i \leq N} \int_0^t \left| \frac{1}{N} \sum_{j=1}^N \left( \kappa_i^{(N)} W_N(x_i, x_j) - W(x_i, x_j) \right) c_j(s) \right| ds \\ &\leq \gamma_{t,\infty} \int_0^t \sup_{\|g\|_\infty \leq 1} \sup_{u \in I} \left| \int \left( W^{\mathcal{G}_N^{(1)}}(u, v) - W^{\mathcal{G}_N^{(2)}}(u, v) \right) g(v) \nu(dv) \right| ds \\ &\leq T \gamma_{T,\infty} \|W^{\mathcal{G}_N^{(1)}} - W^{\mathcal{G}_N^{(2)}}\|_{\infty \rightarrow \infty, \nu} \\ &\leq T \gamma_{T,\infty} \left( \|W^{\mathcal{G}_N^{(1)}} - W\|_{\infty \rightarrow \infty, \nu} + \|W - W^{\mathcal{G}_N^{(2)}}\|_{\infty \rightarrow \infty, \nu} \right). \quad (3.8.29) \end{aligned}$$

Concerning  $A_{i,t,5}^{(N)}$ , we denote by  $\tilde{A}_{t,5}^{(N)} = \max_{1 \leq i \leq N} A_{i,t,5}^{(N)}$ . It is controlled with (3.8.6). Combining (3.8.26), (3.8.27), (3.8.28) and (3.8.29), we obtain if  $N$  is large enough  $\mathbb{P}$ -almost surely for every  $t \in [0, T]$ :

$$\begin{aligned} \mathbf{E} \left[ \max_{1 \leq i \leq N} \Delta_i^{(N)}(t) \right] &\leq C_1 \int_0^t |h(t-s)| \mathbf{E} \left[ \max_{1 \leq i \leq N} \Delta_i^{(N)}(s) \right] ds + C_2 \sqrt{\frac{\kappa_N}{N}} \\ &+ C_3 \sqrt{\varepsilon_N} + C_4 \|W^{\mathcal{G}_N^{(1)}} - W\|_{\infty \rightarrow \infty, \nu} + C_4 \|W^{\mathcal{G}_N^{(2)}} - W\|_{\infty \rightarrow \infty, \nu} + L_f \tilde{A}_{t,5}^{(N)}, \end{aligned} \quad (3.8.30)$$

with  $C_1, C_2, C_3$  and  $C_4$  constants depending on  $h, f, C_W$  and  $T$ . Recall (3.8.8), as  $\sup_{t \in [0, T]} |Z_{N,i}(t) - \bar{Z}_i(t)| \leq \Delta_i^{(N)}(T)$ , Lemma 3.A.5 with  $u(t) = \mathbf{E} \left[ \max_{1 \leq i \leq N} \Delta_i^{(N)}(t) \right]$  on  $[0, T]$  gives that  $\mathbb{P}$ -almost surely on the realization of  $(\xi^{(N)})$  if  $N$  is large enough

$$\begin{aligned} \max_{1 \leq i \leq N} \mathbf{E} \left[ \sup_{t \in [0, T]} |Z_{N,i}(t) - \bar{Z}_i(t)| \right] &\leq C \left( \sqrt{\frac{\kappa_N}{N}} + \sqrt{\varepsilon_N} + \|W^{\mathcal{G}_N^{(1)}} - W\|_{\infty \rightarrow \infty, \nu} \right. \\ &\left. + \|W^{\mathcal{G}_N^{(2)}} - W\|_{\infty \rightarrow \infty, \nu} + \tilde{A}_{t,5}^{(N)} \right), \end{aligned} \quad (3.8.31)$$

with  $C = \sqrt{2} \max(C_2, C_3, C_4, L_f) \exp(C_1^2 \|h\|_{T,2}^2 T)$ . By (3.3.11),  $\lim_{N \rightarrow \infty} \varepsilon_N = 0$ ,  $\lim_{N \rightarrow \infty} \frac{\kappa_N}{N} = 0$  and by (3.3.16)  $\lim_{N \rightarrow \infty} \|W^{\mathcal{G}_N^{(1)}} - W\| = 0$ . Combining with Proposition 3.8.4, it concludes the proof of (3.3.17).  $\square$

### 3.8.4 Proofs: Application to the Scenarios of Definition

#### 3.3.8

In this section, we prove Proposition 3.8.4. We start with auxiliary results that come up in the main proof.

#### 3.8.4.1 Toolbox

**Lemma 3.8.5.** *Let  $(\tilde{x}_i)_{i \geq 1}$  be a sequence of i.i.d positions on  $[0, 1]$  with distribution  $\mathcal{U}[0, 1]$ . For all  $N \geq 1$  and for  $i = 1, \dots, N$ , define  $x_i = \tilde{x}_{(i)}$  as the order statistics of  $(\tilde{x}_1, \dots, \tilde{x}_N)$  (i.e.  $\{\tilde{x}_1, \dots, \tilde{x}_N\} = \{x_1, \dots, x_N\}$  and  $x_1 < \dots < x_N$ ). Then, for any borelian sets  $A$  and  $B$  of  $(0, 1]$ ,*

$$\frac{1}{N} \sum_{i=1}^N \mathbf{1}_{x_i \in A, \frac{i}{N} \in B} \xrightarrow[N \rightarrow \infty]{} \text{Leb}(A \cap B) \quad a.s. \quad (3.8.32)$$

where  $\text{Leb}$  denotes the Lebesgue measure on  $[0, 1]$ .

*Proof.* It is sufficient to show that for all  $(t, t') \in (0, 1]^2$ ,

$$\frac{1}{N} \sum_{i=1}^N \mathbf{1}_{x_i \leq t, \frac{i}{N} \leq t'} \xrightarrow[N \rightarrow \infty]{} \min(t, t') \quad a.s.$$

We introduce the uniform sample quantile function as in Csörgő (1983): define for any  $y \in [0, 1]$

$$U_N(y) = \begin{cases} 0 & \text{if } y = 0 \\ x_k & \text{if } \frac{k-1}{N} < y \leq \frac{k}{N}, \quad k \in \llbracket 1, N \rrbracket. \end{cases} \quad (3.8.33)$$

First, we show that  $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{x_i \leq t, \frac{i}{N} \leq t'} = \lim_{N \rightarrow \infty} \int_0^{t'} \mathbf{1}_{U_N(y) \leq t} dy$ . We note  $k$  the integer such that  $x_k \leq t < x_{k+1}$  (and  $k = 0$  if  $x_1 > t$ ). We have two cases. If  $t' \geq \frac{k}{N}$ , then  $\frac{1}{N} \sum_{i=1}^N \mathbf{1}_{x_i \leq t, \frac{i}{N} \leq t'} = \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{x_i \leq t, i \leq Nt'} = \frac{1}{N} \sum_{i=1}^{Nt'} \mathbf{1}_{x_i \leq t} = \frac{k}{N}$ , and  $\int_0^{t'} \mathbf{1}_{U_N(y) \leq t} dy = \int_0^{t'} \mathbf{1}_{y \leq \frac{k}{N}} dy = \frac{k}{N}$ . If  $t' < \frac{k}{N}$ ,  $\int_0^{t'} \mathbf{1}_{U_N(y) \leq t} dy = \int_0^{t'} \mathbf{1}_{y \leq \frac{k}{N}} dy = t'$  and  $\frac{1}{N} \sum_{i=1}^N \mathbf{1}_{x_i \leq t, \frac{i}{N} \leq t'} = \frac{1}{N} \sum_{i=1}^k \mathbf{1}_{i \leq Nt'} = \frac{\lfloor Nt' \rfloor}{N} \xrightarrow{N \rightarrow \infty} t'$ . Then, we know from Csörgő (1983) that  $\sup_{0 \leq y \leq 1} |U_N(y) - y| \xrightarrow[N \rightarrow \infty]{a.s.} 0$ , and hence almost surely, for any fixed  $y \in [0, 1]$ ,  $U_N(y) \xrightarrow[N \rightarrow \infty]{p.s.} y$  and by dominated convergence  $\int_0^{t'} \mathbf{1}_{U_N(y) \leq t} dy \xrightarrow[N \rightarrow \infty]{} \int_0^{t'} \mathbf{1}_{y \leq t} dy = \min(t, t')$ , which concludes the proof.  $\square$

**Proposition 3.8.6.** *Under the Scenario (1) of Definition 3.3.8, for any function  $g$  such that  $\|g\|_{L^X(I \times I), \nu \times \nu} < \infty$  with  $\chi > 5$ ,*

$$\sup_{1 \leq i \leq N} \int_I g(x_i, y) (\nu^{(N)}(dy) - \nu(dy)) \xrightarrow{N \rightarrow \infty} 0 \quad (3.8.34)$$

$\mathbb{P}$ -almost surely on the realization of the sequence  $(\underline{x}^{(N)})_N$ .

*Proof.* Fix  $M > 0$ , and define the function  $p_M(u) = u \mathbf{1}_{|u| \leq M} + M \mathbf{1}_{u > M} - M \mathbf{1}_{u < -M}$  on  $\mathbb{R}$ . Set  $g_M = p_M \circ g$ . The following arguments come from Luçon (2020) in the proof of Proposition 3.4. We have

$$\begin{aligned} \sup_{1 \leq i \leq N} \int_I g(x_i, y) (\nu^{(N)}(dy) - \nu(dy)) &\leq \sup_{1 \leq i \leq N} \frac{1}{N} \sum_{j=1}^N |g(x_i, x_j) - g_M(x_i, x_j)| \\ &+ \sup_{1 \leq i \leq N} \int_I |g(x_i, y) - g_M(x_i, y)| \nu(dy) + \sup_{1 \leq i \leq N} \int_I g_M(x_i, y) (\nu^{(N)}(dy) - \nu(dy)) \\ &=: (I) + (II) + (III). \end{aligned}$$

To study (I), note that  $|g(x, y) - g_M(x, y)| = |g(x, y) - g_M(x, y)| \mathbf{1}_{|g(x, y)| > M} \leq$



$2|g(x, y)|\mathbf{1}_{|g(x, y)| > M}$ , and that for any independent  $X, Y$  with distribution  $\nu$  we have

$$\begin{aligned} & \mathbb{E} [|g(X, Y)| \mathbf{1}_{|g(X, Y)| > M}] \\ &= \sum_{l=0}^{+\infty} \mathbb{E} [|g(X, Y)| \mathbf{1}_{2^l M < |g(X, Y)| \leq 2^{l+1} M}] \\ &\leq \sum_{l=0}^{+\infty} 2^{l+1} M (\mathbb{P}(|g(X, Y)| > 2^l M) - \mathbb{P}(|g(X, Y)| > 2^{l+1} M)) \\ &= 2M\mathbb{P}(|g(X, Y)| > M) + \sum_{l=1}^{+\infty} 2^l M \mathbb{P}(|g(X, Y)| > 2^l M) \\ &\leq \mathbb{E} [|g(X, Y)|^\chi] \left( \frac{2}{M^{\chi-1}} + \sum_{l=1}^{+\infty} \frac{2^l M}{(2^l M)^\chi} \right) \leq \frac{3\mathbb{E} [|g(X, Y)|^\chi]}{M^{\chi-1}}, \end{aligned}$$

using Markov inequality. Since

$$\mathbb{E} \left[ \frac{1}{N} \sum_{l=1}^N |g(x_i, x_l) - g_M(x_i, x_l)| \right] \leq \frac{2}{N} \sum_{l=1}^N \mathbb{E} [|g(x_i, x_l)| \mathbf{1}_{|g(x_i, x_l)| > M}],$$

it implies for the choice of  $M = N^{\delta_1}$  with  $\delta_1 > 0$  to be defined later, using Markov inequality and a union bound that  $\mathbb{P} \left( (I) > \frac{1}{N^{\delta_2}} \right) \leq \frac{6\mathbb{E} [|g(X, Y)|^\chi]}{N^{\delta_1(\chi-1) - \delta_2 - 1}}$ . Similarly, we can show that  $\mathbb{P} \left( (II) > \frac{1}{N^{\delta_2}} \right) \leq \frac{6\mathbb{E} [|g(X, Y)|^\chi]}{N^{\delta_1(\chi-1) - \delta_2 - 1}}$ . We will use the two previous bounds with Borel-Cantelli Lemma to deduce that  $\mathbb{P}$ -almost surely,  $(I) + (II) \xrightarrow[N \rightarrow \infty]{} 0$  by asking  $\delta_1(\chi - 1) - \delta_2 - 1 > 1$ . To deal with  $(III)$  we use the boundedness of  $g_M$ . Note that  $(III)$  can be re-written  $\sup_{1 \leq i \leq N} \frac{1}{N} \sum_{l=1}^N Y_l^{(i), M}$  with

$$Y_l^{(i), M} := g_M(x_i, x_l) - \int_I g_M(x_i, y) \nu(dy) = g_M(x_i, x_l) - \mathbb{E} [g_M(x_i, Y) | x_i].$$

We set  $\mathcal{F}_l^{(i)} = \sigma(x_i, x_1, \dots, x_l)$ . We have for  $l \neq i$

$$\mathbb{E} [Y_l^{(i), M} | \mathcal{F}_{l-1}^{(i)}] = \mathbb{E} [U_M(x_i, x_l) - \mathbb{E}_Y [U_M(x_i, Y) | x_i] | \mathcal{F}_{l-1}^{(i)}] = 0.$$

As  $|Y_l^{(i), M}| \leq 2M$ , we can then apply Lemma 3.A.6: for any  $x > 0$ ,

$$\mathbb{P} \left( \frac{1}{N-1} \sum_{\substack{l=1 \\ l \neq i}}^N \frac{Y_l^{(i), M}}{2M} \geq x \right) \leq \exp \left( -(N-1) \frac{x^2}{2} B(x) \right),$$

with the function  $B$  defined in (3.A.1). We consider a sequence  $\varepsilon_N$  such that  $\varepsilon_N \xrightarrow[N \rightarrow \infty]{} 0$  (we precise later on which one), and we apply the previous result with

$x = \frac{\varepsilon_N N}{2M(N-1)}$ . As  $B(u) = u^{-2}((1+u)\log(1+u) - u) \rightarrow \frac{1}{2}$  when  $u \rightarrow 0$ , we can choose a deterministic  $p$  such that for all  $N \geq p$ ,  $B\left(\frac{\varepsilon_N N}{2M(N-1)}\right) \geq \frac{1}{4}$ . We then have if  $N \geq p$ :  $\mathbb{P}\left(\frac{1}{N} \sum_{\substack{l=1 \\ l \neq i}}^N Y_l^{(i),M} \geq \varepsilon_N\right) \leq \exp\left(-\frac{1}{32M^2} \frac{\varepsilon_N^2 N^2}{N-1}\right)$ , doing the same for  $-Y_l^{(i)}$  and with a union bound we obtain

$$\mathbb{P}\left(\sup_{1 \leq i \leq N} \left| \frac{1}{N} \sum_{\substack{l=1 \\ l \neq i}}^N Y_l^{(i),M} \right| \geq \varepsilon_N\right) \leq 2N \exp\left(-\frac{1}{32M^2} \frac{\varepsilon_N^2 N^2}{N-1}\right).$$

It is sufficient to find  $\varepsilon_N$  such that  $\varepsilon_N \xrightarrow{N \rightarrow \infty} 0$  and  $\sum_N 2N \exp\left(-\frac{1}{32M^2} \frac{\varepsilon_N^2 N^2}{N-1}\right) < \infty$  to conclude by Borel-Cantelli's Lemma,  $\mathbb{P}$ -almost surely if  $N$  is large enough  $\sup_{1 \leq i \leq N} \left| \frac{1}{N} \sum_{\substack{l=1 \\ l \neq i}}^N Y_l^{(i),M} \right| \leq \varepsilon_N$ . We set then  $\varepsilon_N^2 := 32M^2(N-1)N^\gamma$ , and require  $-2 < \gamma < -1 - 2\delta_1$ . As  $Y_i^{(i),M}$  is bounded (by  $2M$ ), adding the term  $\frac{1}{N} Y_i^{(i),M}$  does not change the convergence if  $\delta_1 < 1$  which was already asked for the conditions on  $\varepsilon_N$  (recall  $M = N^{\delta_1}$ ). We are left with finding parameters  $(\delta_1, \delta_2, \gamma)$  such that  $\delta_1 > 0$ ,  $\delta_2 > 0$ ,  $\delta_1(\chi - 1) - \delta_2 - 1 > 1$ ,  $-2 < \gamma < -1 - 2\delta_1$  (to ensure that the probabilities obtained with (I), (II) and (III) are summable and the sufficient conditions on  $\varepsilon_N$ ). As  $\chi > 5$ , any choice such that  $\delta_1 \in (0, \frac{1}{2})$  and  $\delta_2 \in (0, 1)$  works (as  $\delta_1(\chi - 1) - 1 < 1$ ) with  $\gamma \in (-2, -1 - 2\delta_1)$ , and we obtain (3.8.34)  $\mathbb{P}$ -almost surely.  $\square$

**Corollary 3.8.7.** *Under Scenario (1) of Definition 3.3.8, we define*

$$\epsilon_{i,1} := \int_{I \times I} W(x_i, y) W(x_i, z) \Gamma(y, z) (\nu^{(N)}(dy) \nu^{(N)}(dz) - \nu(dy) \nu^{(N)}(dz)) \quad (3.8.35)$$

$$\epsilon_{i,2} := \int_{I \times I} W(x_i, y) W(x_i, z) \Gamma(y, z) (\nu(dy) \nu^{(N)}(dz) - \nu(dy) \nu(dz)), \quad (3.8.36)$$

where  $\Gamma$  is defined in (3.8.19). Then under Hypothesis 3.3.7,  $\mathbb{P}$ -almost surely,

$$\sup_{1 \leq i \leq N} \epsilon_{i,1} \xrightarrow{N \rightarrow \infty} 0 \quad \text{and} \quad \sup_{1 \leq i \leq N} \epsilon_{i,2} \xrightarrow{N \rightarrow \infty} 0.$$

*Proof.* Define  $\phi(x, z) := W(x, z) \int_I W(x, y) \Gamma(y, z) \nu(dy)$ , then (3.8.36) becomes  $\epsilon_{i,2} = \int_I \phi(x_i, z) (\nu^{(N)}(dz) - \nu(dz))$ . As  $\Gamma$  is bounded,  $|\phi(x, z)| \leq |W(x, z)| \|\Gamma\|_\infty C_W^{(1)}$  and since  $W \in L^X(I^2, \nu \times \nu)$ ,  $\|\phi\|_{L^X(I \times I), \nu \times \nu} < \infty$  thus  $\sup_{1 \leq i \leq N} \epsilon_{i,2} \xrightarrow{N \rightarrow \infty} 0$  is an immediate application of Proposition 3.8.6. Similarly, define  $g_N(x, y) := W(x, y) \int_I W(x, z) \Gamma(y, z) \nu^{(N)}(dz)$  and  $g(x, y) := W(x, y) \int_I W(x, z) \Gamma(y, z) \nu(dz)$ , (3.8.35) becomes

$$\epsilon_{i,1} = \int_I (g_N(x_i, y) - g(x_i, y)) (\nu^{(N)}(dy) - \nu(dy)) + \int_I g(x_i, y) (\nu^{(N)}(dy) - \nu(dy)).$$

As done with (3.8.36), we have immediately that

$$\sup_{1 \leq i \leq N} \int_I g(x_i, y) (\nu^{(N)}(dy) - \nu(dy)) \xrightarrow{N \rightarrow \infty} 0.$$

The other term can be written  $\epsilon_{i,3} := \int_I W(x_i, y) \alpha_N(x_i, y) (\nu^{(N)}(dy) - \nu(dy))$  where  $\alpha_N(x_i, y) := \int_I W(x_i, z) \Gamma(y, z) (\nu^{(N)}(dz) - \nu(dz))$ . As  $\Gamma$  is bounded, Proposition 3.8.6 (and its proof) gives then that  $\alpha_N(x_i, y) \xrightarrow{N \rightarrow \infty} 0$  uniformly in  $i$  and  $y$ . Another application of Proposition 3.8.6 gives that  $\sup_{1 \leq i \leq N} \epsilon_{i,3} \xrightarrow{N \rightarrow \infty} 0$ , which concludes the proof.  $\square$

### 3.8.4.2 Proof of Proposition 3.8.4 for Scenario (1)

We treat the estimates (3.8.4), (3.8.5), (3.8.6) and (3.8.7) separately.

**Proof of (3.8.4)** We remind that we want to prove  $d_{\square, \nu} \left( W^{\mathcal{G}_N^{(2)}}, W \right) \xrightarrow{N \rightarrow \infty} 0$ , when the positions are i.i.d. according to  $\nu$  on  $I$ . Recall the definition of  $(x_1, \dots, x_N)$  as the lexicographic reordering of the i.i.d. sample  $(\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_N)$ . The proof is organised as follow: we start by looking at the case  $d = 1$ ,  $I = [0, 1]$  and  $\nu$  is the Lebesgue measure on  $I$ , and then extend to the general case

*Step 1 - Approximation of  $W$  in norm  $L^1$ .* We first prove that for  $\varepsilon > 0$ , there exists  $m \geq 1$  sufficiently large such that  $\|W - W_{\mathcal{P}_m}\|_{L^1(I^2)} \leq \varepsilon$ . We fix  $\varepsilon > 0$ . As  $W \in L^1(I^2, \nu)$ , there exists  $\widetilde{W}$  continuous such that  $\|W - \widetilde{W}\|_{1, \nu} \leq \frac{\varepsilon}{3}$ . As  $\widetilde{W}$  is also uniformly continuous, there exists  $\eta > 0$  such that if  $\|u - u'\| + \|v - v'\| \leq \eta$ ,  $|\widetilde{W}(u, v) - \widetilde{W}(u', v')| \leq \frac{\varepsilon}{3}$ . We fix  $m$  large enough such that  $\frac{1}{m} \leq \eta$ , and denote by  $\mathcal{P}_m = \sqcup_{i=1}^m J_i$  the partition with  $J_i = \left(\frac{i-1}{m}, \frac{i}{m}\right]$ . It verifies then, for each  $i \in (1, \dots, m)$   $\text{Diam}(J_i) \leq \eta$ . We define the step function (which average the values of  $W$  over cells obtained with the partition)

$$W_{\mathcal{P}_m}(u, v) := m^2 \sum_{i,j=1}^m \int_{J_i \times J_j} W(x, y) \nu(dx) \nu(dy) \mathbf{1}_{J_i}(u) \mathbf{1}_{J_j}(v). \quad (3.8.37)$$

We note  $\mathcal{G}_N^{(3)}$  the directed weighted graph with vertices  $\{1, \dots, N\}$  such that every edge  $j \rightarrow i$  is present, with weight  $W_{\mathcal{P}_m}(x_i, x_j)$ . We use it to upper-bound the cut-distance between  $W$  and  $W^{\mathcal{G}_N^{(2)}}$ :

$$\begin{aligned} d_{\square, \nu} \left( W^{\mathcal{G}_N^{(2)}}, W \right) &\leq \|W^{\mathcal{G}_N^{(2)}} - W\|_{1, \nu} \\ &\leq \|W^{\mathcal{G}_N^{(2)}} - W^{\mathcal{G}_N^{(3)}}\|_{1, \nu} + \|W^{\mathcal{G}_N^{(3)}} - W_{\mathcal{P}_m}\|_{1, \nu} + \|W_{\mathcal{P}_m} - W\|_{1, \nu}. \end{aligned} \quad (3.8.38)$$

We are going to control each term of the right hand side of (3.8.38) in the following steps.

*Step 2 - Control of  $\|W_{\mathcal{P}_m} - W\|_{1,\nu}$ .* We have

$$\|W_{\mathcal{P}_m} - W\|_{1,\nu} \leq \|W_{\mathcal{P}_m} - \widetilde{W}_{\mathcal{P}_m}\|_{1,\nu} + \|\widetilde{W}_{\mathcal{P}_m} - \widetilde{W}\|_{1,\nu} + \|\widetilde{W} - W\|_{1,\nu}.$$

As  $\|W - \widetilde{W}\|_{1,\nu} \leq \frac{\varepsilon}{3}$ , and as for any partition  $\mathcal{P}$ ,  $\|W_{\mathcal{P}}\|_{1,\nu} \leq \|W\|_{1,\nu}$ , we have  $\|W_{\mathcal{P}_m} - \widetilde{W}_{\mathcal{P}_m}\|_{1,\nu} \leq \frac{\varepsilon}{3}$  and

$$\begin{aligned} \|\widetilde{W}_{\mathcal{P}_m} - \widetilde{W}\|_{1,\nu} &= \sum_{i,j=1}^m \int_{J_i} \int_{J_j} \left| \widetilde{W}(u,v) - m^2 \int_{J_i} \int_{J_j} \widetilde{W}(x,y) \nu(dx) \nu(dy) \right| \nu(du) \nu(dv) \\ &\leq \sum_{i,j=1}^m \int_{J_i} \int_{J_j} m^2 \int_{J_i} \int_{J_j} \left| \widetilde{W}(u,v) - \widetilde{W}(x,y) \right| \nu(dx) \nu(dy) \nu(du) \nu(dv) \\ &\leq \sum_{i,j=1}^m \int_{J_i} \int_{J_j} m^2 \int_{J_i} \int_{J_j} \frac{\varepsilon}{3} \nu(dx) \nu(dy) \nu(du) \nu(dv) \leq \frac{\varepsilon}{3}, \end{aligned}$$

hence  $\|W_{\mathcal{P}_m} - W\|_{1,\nu} \leq \varepsilon$  (recall here that  $\nu$  is the Lebesgue measure on  $[0, 1]$ ).

*Step 3 - Control of  $\|W^{\mathcal{G}_N^{(2)}} - W^{\mathcal{G}_N^{(3)}}\|_{1,\nu}$ .* For all  $N \geq 1$ , we recall from Lemma 3.3.5  $(B_1^{(N)}, \dots, B_N^{(N)})$  the partition of  $I$  with  $B_i = \left(\frac{i-1}{N}, \frac{i}{N}\right]$  (we omit by simplicity the upper index  $(N)$ ). Using the notation introduced in (3.3.7), we have

$$\begin{aligned} \|W^{\mathcal{G}_N^{(2)}} - W^{\mathcal{G}_N^{(3)}}\|_{1,\nu} &= \sum_{i,j=1}^N \int_{B_i} \int_{B_j} |W(x_i, x_j) - W_{\mathcal{P}_m}(x_i, x_j)| \nu(du) \nu(dv) \\ &= \frac{1}{N^2} \sum_{i,j=1}^N |W(x_i, x_j) - W_{\mathcal{P}_m}(x_i, x_j)| =: \frac{1}{N^2} \sum_{i,j=1}^N F(x_i, x_j). \end{aligned} \tag{3.8.39}$$

We use the following proposition to show that it converges almost surely to  $\|W_{\mathcal{P}_m} - W\|_{1,\nu}$ .

**Proposition 3.8.8 (Hoeffding (1961)).** *Let  $X_1, X_2, \dots$  be a sequence of i.i.d. random variables with distribution  $\nu$ , and  $f$  a real-valued measurable function. Then if  $\mathbb{E}[|f(X_1, X_2)|] < +\infty$ ,*

$$\frac{1}{N(N-1)} \sum_{\substack{i,j=1 \\ i \neq j}}^N f(X_i, X_j) \xrightarrow[N \rightarrow \infty]{a.s.} \mathbb{E}[f(X_1, X_2)] = \int \int f(x, y) \nu(dx) \nu(dy). \tag{3.8.40}$$

We have indeed

$$\sum_{i,j=1}^N F(x_i, x_j) = \frac{1}{N^2} \sum_{i=1}^N F(x_i, x_i) + \frac{N(N-1)}{N^2} \frac{1}{N(N-1)} \sum_{\substack{i,j=1 \\ i \neq j}}^N F(x_i, x_j),$$

where the second term converges as  $N \rightarrow \infty$  to  $\iint F(x, y)\nu(dx)\nu(dy)$  a.s. and

$$\frac{1}{N^2} \sum_{i=1}^N F(x_i, x_i) \leq \frac{1}{N} \left( \frac{1}{N} \sum_{i=1}^N |W(x_i, x_i)| + \frac{1}{N} \sum_{i=1}^N |W_{\mathcal{P}_m}(x_i, x_i)| \right) \xrightarrow{N \rightarrow \infty} 0$$

as the sums are controlled by Hypothesis 3.2.6.

*Step 4 - Control of  $\|W^{\mathcal{G}_N^{(3)}} - W_{\mathcal{P}_m}\|_{1,\nu}$ .* We have

$$\|W^{\mathcal{G}_N^{(3)}} - W_{\mathcal{P}_m}\|_{1,\nu} = \sum_{i,j=1}^N \int_{B_i} \int_{B_j} |W_{\mathcal{P}_m}(x_i, x_j) - W_{\mathcal{P}_m}(x, y)| \nu(dx)\nu(dy). \quad (3.8.41)$$

Recalling (3.8.37) and setting  $\alpha_{kl} = m^2 \int_{J_k \times J_l} W(u, v)\nu(du)\nu(dv)$  we have

$$\begin{aligned} & \|W^{\mathcal{G}_N^{(3)}} - W_{\mathcal{P}_m}\|_{1,\nu} \\ &= \sum_{i,j=1}^N \int_{B_i} \int_{B_j} \left| \sum_{k,l} \alpha_{kl} \mathbf{1}_{J_k \times J_l}(x_i, x_j) - \sum_{k',l'} \alpha_{k'l'} \mathbf{1}_{J_{k'} \times J_{l'}}(x, y) \right| \nu(dx)\nu(dy) \\ &= \sum_{i,j=1}^N \sum_{k,l} \sum_{k',l'} |\alpha_{kl} - \alpha_{k'l'}| \mathbf{1}_{J_k \times J_l}(x_i, x_j) \int_{B_i} \int_{B_j} \mathbf{1}_{J_{k'} \times J_{l'}}(x, y) \nu(dx)\nu(dy) \\ &= \sum_{k,l} \sum_{k',l'} |\alpha_{kl} - \alpha_{k'l'}| \sum_{i,j=1}^N \mathbf{1}_{J_k \times J_l}(x_i, x_j) \nu(J_{k'} \cap B_i) \nu(J_{l'} \cap B_j). \end{aligned}$$

We consider  $N$  large enough ( $N > m$ ) such that every box  $B_i = (\frac{i-1}{N}, \frac{i}{N}]$  (of size  $\frac{1}{N}$ ) is inside a larger box  $J_{k'} = (\frac{k'-1}{m}, \frac{k'}{m}]$  (of size  $\frac{1}{m}$ ) (there might be some  $B_i$  that are on two different parts of the partition  $\mathcal{P}_m$ , but we can neglect this contribution - at most of order  $\frac{m}{N} \xrightarrow{N \rightarrow \infty} 0$ ). Then  $\nu(J_{k'} \cap B_i) = \mathbf{1}_{\{B_i \subset J_{k'}\}} \nu(B_i) = \frac{1}{N} \mathbf{1}_{\{\frac{i}{N} \in J_{k'}\}}$ , and

$$\begin{aligned} & \|W^{\mathcal{G}_N^{(3)}} - W_{\mathcal{P}_m}\|_{1,\nu} \\ & \leq \sum_{k,l=1}^m \sum_{k',l'=1}^m |\alpha_{kl} - \alpha_{k'l'}| \sum_{i,j=1}^N \frac{1}{N^2} \mathbf{1}_{\{(x_i, x_j) \in J_k \times J_l, B_i \subset J_{k'}, B_j \subset J_{l'}\}} \\ & \leq \sum_{k,l=1}^m \sum_{k',l'=1}^m |\alpha_{kl} - \alpha_{k'l'}| \sum_{i,j=1}^N \frac{1}{N^2} \mathbf{1}_{\{(x_i, x_j) \in J_k \times J_l, \frac{i}{N} \in J_{k'}, \frac{j}{N} \in J_{l'}\}} \\ & \leq \sum_{k,l=1}^m \sum_{k',l'=1}^m |\alpha_{kl} - \alpha_{k'l'}| \left( \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{\{x_i \in J_k, \frac{i}{N} \in J_{k'}\}} \right) \left( \frac{1}{N} \sum_{j=1}^N \mathbf{1}_{\{x_j \in J_l, \frac{j}{N} \in J_{l'}\}} \right). \end{aligned}$$

Then, from Lemma 3.8.5,  $\frac{1}{N} \sum_{i=1}^N \mathbf{1}_{\{x_i \in J_k, \frac{i}{N} \in J_{k'}\}} \xrightarrow{N \rightarrow \infty} \lambda(J_k \cap J_{k'}) = \frac{1}{m} \mathbf{1}_{k=k'}$ , hence we obtain that

$$\limsup_{N \rightarrow \infty} \|W^{\mathcal{G}_N^{(3)}} - W_{\mathcal{P}_m}\|_{1,\nu} \leq \frac{1}{m^2} \sum_{k,l=1}^m \sum_{k',l'=1}^m |\alpha_{k,l} - \alpha_{k',l'}| \mathbf{1}_{k=k'} \mathbf{1}_{l=l'}.$$

The claim is that the above bound is uniformly 0 for all  $m$ : the sum reduces to  $k = k'$  and  $l = l'$  hence the prefactor  $|\alpha_{k,l} - \alpha_{k',l'}|$  gives that this last contribution is 0, thus almost-surely (on the realization of the sequence of positions) we have  $\|W^{\mathcal{G}_N^{(3)}} - W_{\mathcal{P}_m}\|_{1,\nu} \xrightarrow{N \rightarrow \infty} 0$ .

*Conclusion when the positions are uniformly drawn* - From (3.8.38) and Steps 3 and 4, we obtain that  $\limsup_{N \rightarrow \infty} d_{\square} \left( W^{\mathcal{G}_N^{(2)}}, W \right) \leq 2 \|W_{\mathcal{P}_m} - W\|_{1,\nu}$ . Choosing now  $m$  as in Step 2 gives that  $\limsup_{N \rightarrow \infty} d_{\square} \left( W^{\mathcal{G}_N^{(2)}}, W \right) \leq \varepsilon$  for all  $\varepsilon > 0$ , which concludes the proof for the case  $x_i \sim \mathcal{U}(0, 1)$ .

*Generalisation: from  $[0, 1]$  to  $[0, 1]^d$*  - Consider the case  $x_i = \left( u_i^{(1)}, \dots, u_i^{(d)} \right)$  where  $\left( u_i^{(j)} \right)_{1 \leq j \leq d}$  are drawn uniformly on  $(0, 1]$  (but not necessarily independent),

and the partition  $I = (0, 1]^d = \bigsqcup_{i=1}^N B_i = \bigsqcup_{i=1}^N \left( \left( \frac{i-1}{N}, \frac{i}{N} \right] \times (0, 1]^{d-1} \right)$ . Proposition 3.8.8 still apply, and the treatment of the terms  $\|W^{\mathcal{G}_N^{(2)}} - W^{\mathcal{G}_N^{(3)}}\|_{1,\nu}$  and  $\|W_{\mathcal{P}_m} - W\|_{1,\nu}$  in (3.8.38) remains the same. For the term  $\|W^{\mathcal{G}_N^{(3)}} - W_{\mathcal{P}_m}\|_{1,\nu}$ , it suffices to note that the chosen partition  $\bigsqcup B_i$  only affects the first coordinates to conclude by the same arguments.

*General case* - Consider  $\nu$  absolutely continuous with respect to Lebesgue measure, and  $I \subset \mathbb{R}^d$ . From Sklar's theorem (Nelsen, 1999, Theorem 2.3.3), we have:

$$f_{\nu}(x^{(1)}, \dots, x^{(d)}) = c(F_1(x^{(1)}), \dots, F_d(x^{(d)})) f_1(x^{(1)}) \dots f_d(x^{(d)}),$$

where  $c$  is the copula density function of  $\nu$ ,  $f_i$  the  $i$ -th marginal probability density functions,  $F_i$  the  $i$ -th marginal cumulative distribution functions and  $f_{\nu}$  the density of  $\nu$  with respect to Lebesgue measure:  $\nu(dx) = f_{\nu}(x^{(1)}, \dots, x^{(d)}) dx^{(1)} \dots dx^{(d)}$ . It implies, by the change of variables  $u = (F_1(x^{(1)}), \dots, F_d(x^{(d)}))$  that  $c(u) du = f_{\nu}(x) dx$ . Define also  $u_i = \left( F_1(x_i^{(1)}), \dots, F_d(x_i^{(d)}) \right)$  and

$$W_F(u, v) := W \left( (F_1^{-1}(u^{(1)}), \dots, F_d^{-1}(u^{(d)})), (F_1^{-1}(v^{(1)}), \dots, F_d^{-1}(v^{(d)})) \right),$$

the previous change of variable gives then that

$$\begin{aligned} \|W^{\mathcal{G}_N^{(2)}} - W\|_{1,\nu} &= \sum_{i,j=1}^N \int_{B_i} \int_{B_j} |W(x_i, x_j) - W(x, y)| \nu(dx) \nu(dy) \\ &= \sum_{i,j=1}^N \int_{\left( \frac{i-1}{N}, \frac{i}{N} \right] \times (0, 1]^{d-1}} \int_{\left( \frac{j-1}{N}, \frac{j}{N} \right] \times (0, 1]^{d-1}} |W_F(u_i, u_j) - W_F(u, v)| c(u) c(v) du dv, \end{aligned}$$

with

$$B_i := \left( F_1^{-1} \left( \frac{i-1}{N}, \frac{i}{N} \right] \times F_2^{-1}((0, 1]) \times \dots \times F_d^{-1}((0, 1]) \right)$$

(note that this partition corresponds to the one introduced in Lemma 3.3.5). The previous case gives immediately the result.

**Proof of (3.8.5)** We remind that we want to prove  $\|W^{\mathcal{G}_N^{(2)}} - W\|_{\infty \rightarrow \infty, \nu} \xrightarrow{N \rightarrow \infty} 0$ . As in the proof of (3.8.4), we start with the case  $I = [0, 1]$ ,  $\tilde{x}_i \sim \mathcal{U}(0, 1)$  i.i.d. (then  $\nu$  is the Lebesgue measure). What changes is that we no longer integrate with respect to the first variable, but we take the supremum. The approximation in  $L^1(I^2)$  is not adapted anymore, thus we approximate  $W$  differently. Recall that

$$\|W^{\mathcal{G}_N^{(2)}} - W\|_{\infty \rightarrow \infty, \nu} = \sup_{\|g\|_{\infty} \leq 1} \sup_{u \in I} \left| \int_I \left( W^{\mathcal{G}_N^{(2)}}(u, v) - W(u, v) \right) g(v) \nu(dv) \right|.$$

*Step 1 - A first bound.* Fixing  $g$  such that  $\|g\|_{\infty} \leq 1$  and  $u \in I$ , for any  $N$  there exists a unique  $i$  such that  $u \in B_{N,i} = \left( \frac{i-1}{N}, \frac{i}{N} \right]$ . Then

$$\begin{aligned} \int_I \left( W^{\mathcal{G}_N^{(2)}}(u, v) - W(u, v) \right) g(v) \nu(dv) &= \sum_{j=1}^N \int_{B_j} \left( W(x_i, x_j) - W(u, v) \right) g(v) \nu(dv) \\ &\leq \left| \sum_{j=1}^N \int_{B_j} \left( W(x_i, x_j) - W(x_i, v) \right) g(v) \nu(dv) \right| + \left| \int_I \left( W(x_i, v) - W(u, v) \right) g(v) \nu(dv) \right| \\ &=: A(g, u) + B(g, u). \end{aligned}$$

*Step 2 - Upper-bound of  $A(g, u)$  by approximated functions independent of  $g$ .* As  $\|g\|_{\infty} \leq 1$ , we have  $A(g, u) \leq \sum_{j=1}^N \int_{B_j} |W(x_i, x_j) - W(x_i, v)| dv$ . Note that it does not depend anymore on  $g$  and it depends on  $u$  only by the index  $i$ . To control this term, we first approximate  $W$  by a stepfunction in  $L^1(I)$ ,  $\widetilde{W}_{\mathcal{P}_m}$ . Introduce  $(\varphi_{\eta})_{\eta > 0}$  as  $\varphi_{\eta}(x) = \eta^{-1} \phi(\frac{x}{\eta})$  where  $\phi$  is a nonnegative continuous function of  $I$  with  $\int \phi = 1$ . Define for all  $x \in I$   $\widetilde{W}_{\eta}(x, \cdot) := W(x, \cdot) * \varphi_{\eta}$ . Note that  $y \mapsto \widetilde{W}_{\eta}(x, y) \in \mathbb{R}$  is a continuous function for all  $x \in [0, 1]$ . As for any  $(x, x') \in I^2$ ,  $\|W(x, \cdot) - W(x', \cdot)\|_1 \leq C_w \|x - x'\|^{\vartheta}$  using (3.2.5),  $x \mapsto W(x, \cdot)$  is continuous from  $[0, 1]$  to  $L^1(I)$ , so that the set of functions  $F := \{W(x, \cdot), x \in [0, 1]\}$  is compact. Hence, for  $\varepsilon > 0$ , we can find  $p \geq 1$  and  $p$  positions  $y_1, \dots, y_p$  such that  $F \subset \cup_{k=1}^p B_{L^1}(W(y_k, \cdot), \varepsilon)$ . Then, there exists  $\eta > 0$  such that for all  $k \leq p$ ,  $\|\widetilde{W}_{\eta}(y_k, \cdot) - W(y_k, \cdot)\|_{L^1} \leq \varepsilon$ . From now, we may omit the notation  $\eta$  for  $\widetilde{W}$ . Let  $m \geq 1$  and  $\mathcal{P}_m = \sqcup_{i=1}^m J_i$  for  $J_r = \left( \frac{r-1}{m}, \frac{r}{m} \right]$  the regular partition of  $I$  of order  $m$ . For any kernel  $H$  on  $I^2$ , define

$$H_{\mathcal{P}_m}(x, v) := m \sum_{r=1}^m \left( \int_{J_r} H(x, y) dy \right) \mathbf{1}_{J_r}(v). \quad (3.8.42)$$

The function  $H \mapsto H_{\mathcal{P}_m}$  is continuous:  $\|H_{\mathcal{P}_m}\|_{L^1(I^2)} \leq \|H\|_{L^1(I^2)}$ . Note that this definition is different from the one used in the proof of (3.8.4) where we integrated on both variables. By continuity of  $y \mapsto \widetilde{W}(y_k, y)$  for all  $k = 1 \dots p$ , there exists  $m \geq 1$  such that  $\sup_{1 \leq l \leq p} \|\widetilde{W}(y_l, \cdot) - \widetilde{W}_{\mathcal{P}_m}(y_l, \cdot)\|_{\infty} \leq \varepsilon$ , and thus

$\sup_{1 \leq l \leq p} \int \left| \widetilde{W}(y_l, y) - \widetilde{W}_{\mathcal{P}_m}(y_l, y) \right| dy \leq \varepsilon$ . Then, for any  $x \in I$ ,

$$\begin{aligned} \|W(x, \cdot) - \widetilde{W}_{\mathcal{P}_m}(x, \cdot)\|_{I,1} &\leq \underbrace{\|W(x, \cdot) - W(y_l, \cdot)\|_{I,1}}_{\leq \varepsilon \text{ by the cover of } F} + \underbrace{\|W(y_l, \cdot) - \widetilde{W}(y_l, \cdot)\|_{I,1}}_{\leq \varepsilon \text{ by the choice of } \eta} \\ &\quad + \underbrace{\|\widetilde{W}(y_l, \cdot) - \widetilde{W}_{\mathcal{P}_m}(y_l, \cdot)\|_{I,1}}_{\leq \varepsilon \text{ by the choice of } m} + \|\widetilde{W}_{\mathcal{P}_m}(y_l, \cdot) - \widetilde{W}_{\mathcal{P}_m}(x, \cdot)\|_{I,1} \\ &\leq 3\varepsilon + \|\widetilde{W}(y_l, \cdot) - \widetilde{W}(x, \cdot)\|_{I,1} \leq 4\varepsilon, \end{aligned}$$

where we used the fact that for any partition  $\mathcal{P}$ ,  $\|\widetilde{W}_{\mathcal{P}}\|_{I,1} \leq \|\widetilde{W}\|_{I,1}$  and because

$$\begin{aligned} \left\| \widetilde{W}(y_l, \cdot) - \widetilde{W}(x, \cdot) \right\|_{I,1} &= \|(W(y_l, \cdot) - W(x, \cdot)) * \varphi_\eta\|_{I,1} \\ &\leq \|W(y_l, \cdot) - W(x, \cdot)\|_{I,1} \|\varphi_\eta\|_{I,1} \end{aligned}$$

by Young's inequality. By compactness of  $F$  and since  $B_{L^1}(W(y_l, \cdot))_{l=1, \dots, p}$  is an  $\varepsilon$ -covering of  $F$ , this last term is smaller than  $\varepsilon$ . Using this approximation, we can now upper bound  $A(g, u)$  independently of the choice of  $g$  and relying on the choice of  $u$  only by the index  $i$  such that  $u \in B_{N,i}$ : we have

$$\begin{aligned} A(g, u) &\leq \frac{1}{N} \sum_{j=1}^N \left| W(x_i, x_j) - \widetilde{W}_{\mathcal{P}_m}(x_i, x_j) \right| + \sum_{j=1}^N \int_{B_j} \left| \widetilde{W}_{\mathcal{P}_m}(x_i, x_j) - \widetilde{W}_{\mathcal{P}_m}(x_i, v) \right| dv \\ &\quad + \sum_{j=1}^N \int_{B_j} \left| \widetilde{W}_{\mathcal{P}_m}(x_i, v) - W(x_i, v) \right| dv =: A_1^{(i)} + A_2^{(i)} + A_3^{(i)}. \end{aligned}$$

*Step 3 - Uniform control of the  $A_k^{(i)}$ .* As  $A_3^{(i)} = \|\widetilde{W}_{\mathcal{P}_m}(x_i, \cdot) - W(x_i, \cdot)\|_{I,1}$ , we control it by the work done previously independently of the index  $i$  (see Step 2):  $\sup_i A_3^{(i)} \xrightarrow{N \rightarrow \infty} 0$ . Set  $g(x, y) := W(x, y) - \widetilde{W}_{\mathcal{P}_m}(x, y)$ , and as  $W \in L^X(I^2)$ , so does  $g$ . We can then apply Proposition 3.8.6 and we obtain

$$\sup_{1 \leq i \leq N} \left| A_3^{(i)} - A_1^{(i)} \right| = \sup_{1 \leq i \leq N} \int_I g(x_i, y) (\nu^{(N)}(dy) - \nu(dy)) \xrightarrow{N \rightarrow \infty} 0.$$

We focus now on  $A_2^{(i)}$  and show that  $\sup_x \sum_{j=1}^N \int_{B_j} \left| \widetilde{W}_{\mathcal{P}_m}(x, x_j) - \widetilde{W}_{\mathcal{P}_m}(x, v) \right| dv$  tends to 0: denoting by  $\alpha_k(x) = m \int_{J_k} \widetilde{W}(x, y) dy$ , we have

$$\begin{aligned} &\sum_{j=1}^N \int_{B_j} \left| \widetilde{W}_{\mathcal{P}_m}(x, x_j) - \widetilde{W}_{\mathcal{P}_m}(x, v) \right| dv \\ &= \sum_{j=1}^N \int_{B_j} \left| \sum_{k=1}^m \alpha_k(x) \mathbf{1}_{J_k}(x_j) - \sum_{k'=1}^m \alpha_{k'}(x) \mathbf{1}_{J_{k'}}(v) \right| dv \\ &\leq \sum_{k, k'=1}^m |\alpha_k(x) - \alpha_{k'}(x)| \sum_{j=1}^N \mathbf{1}_{J_k}(x_j) |J_{k'} \cap B_j|. \end{aligned}$$



Similarly to what has been done in Step 4 for the proof of (3.8.4), we consider  $N$  large enough ( $N > m$ ) such that every box  $B_i = \left(\frac{i-1}{N}, \frac{i}{N}\right]$  is inside a larger box  $J_{k'} = \left(\frac{k'-1}{m}, \frac{k'}{m}\right]$ , then  $\nu(J_{k'} \cap B_j) = \mathbf{1}_{\{B_j \subset J_{k'}\}} \nu(B_j) = \frac{1}{N} \mathbf{1}_{\{\frac{j}{N} \in J_{k'}\}}$  and

$$\sum_{j=1}^N \int_{B_j} \left| \widetilde{W}_{\mathcal{P}_m}(x, x_j) - \widetilde{W}_{\mathcal{P}_m}(x, v) \right| dv \leq \sum_{\substack{k, k'=1 \\ k \neq k'}}^m |\alpha_k(x) - \alpha_{k'}(x)| \sum_{j=1}^N \frac{1}{N} \mathbf{1}_{\{x_j \in J_k, \frac{j}{N} \in J_{k'}\}}.$$

As  $\alpha_k(x) \leq m \int_I \widetilde{W}(x, y) dy \leq m C_W^{(1)}$  which is independent of  $x$  and  $k$ ,

$$\sum_{j=1}^N \int_{B_j} \left| \widetilde{W}_{\mathcal{P}_m}(x, x_j) - \widetilde{W}_{\mathcal{P}_m}(x, v) \right| dv \leq 2m C_W^{(1)} \sum_{\substack{k, k'=1 \\ k \neq k'}}^m \sum_{j=1}^N \frac{1}{N} \mathbf{1}_{\{x_j \in J_k, \frac{j}{N} \in J_{k'}\}}.$$

From Lemma 3.8.5,  $\frac{1}{N} \sum_{j=1}^N \mathbf{1}_{\{x_j \in J_k, \frac{j}{N} \in J_{k'}\}} \xrightarrow[N \rightarrow \infty]{a.s.} \lambda(J_k \cap J_{k'}) = \frac{1}{m} \mathbf{1}_{k=k'}$ , thus a.s. (on

the realization of the sequence of positions),  $\sum_{j=1}^N \int_{B_j} \left| \widetilde{W}_{\mathcal{P}_m}(x, x_j) - \widetilde{W}_{\mathcal{P}_m}(x, v) \right| dv$  tends to 0 independently on the choice of  $x$ . We have shown that  $\mathbb{P}$ -a.s.,

$$\sup_{g, \|g\|_\infty \leq 1} \sup_{u \in I} A(g, u) \xrightarrow[N \rightarrow \infty]{} 0.$$

*Step 4 - Control of  $B(g, u)$  and conclusion.* Using (3.2.5) from Hypothesis 3.2.6, we have

$$B(g, u) \leq \int_I |W(x_i, v) - W(u, v)| \|g\|_\infty \nu(dv) \leq C_w \|x_i - u\|^\vartheta.$$

Let us show that  $\sup_{x \in I} \sum_{i=1}^N \mathbf{1}_{B_i}(x) \|x_i - x\|^\vartheta \xrightarrow[N \rightarrow \infty]{} 0$ . Recall (3.8.33): we have

$\sum_{i=1}^N \mathbf{1}_{B_i}(x) \|x_i - x\|^\vartheta = \|U_N(x) - x\|^\vartheta$  by definition of  $U_N$ , the uniform sample quantile function. As we know from Csörgő (1983) that  $\sup_{0 \leq y \leq 1} |U_N(y) - y| \xrightarrow[N \rightarrow \infty]{a.s.} 0$ , almost surely  $\sup_{g, \|g\|_\infty \leq 1} \sup_{u \in I} B(g, u) \xrightarrow[N \rightarrow \infty]{} 0$ . It concludes the proof for (3.8.5).

**Proof of (3.8.7) and (3.8.6)** The term of interest is  $A_{i,T,5}^{(N)}$ , defined in (3.8.14), we have by Jensen's inequality

$$\begin{aligned}
 A_{i,T,5}^{(N)2} &\leq \left( \int_0^T \left| \int_I F(x_i, y, s) (\nu^{(N)}(dy) - \nu(dy)) \right| ds \right)^2 \\
 &\leq T \int_0^T \left( \frac{1}{N} \sum_{j=1}^N \left( F(x_i, x_j, s) - \int_I F(x_i, y, s) \nu(dy) \right) \right)^2 ds \\
 &\leq \frac{T}{N^2} \sum_{j,l=1}^N \int_0^T \left( F(x_i, x_j, s) F(x_i, x_l, s) + \left( \int_I F(x_i, y, s) \nu(dy) \right)^2 \right. \\
 &\quad \left. - 2F(x_i, x_j, s) \int_I F(x_i, y, s) \nu(dy) \right) ds.
 \end{aligned}$$

As  $F(x_i, y, s) F(x_i, z, s) = W(x_i, y)W(x_i, z)\gamma(s, y)\gamma(s, z)$  for any  $y$  and  $z$ , denoting by  $\Gamma(y, z) := \int_0^T \gamma(s, y)\gamma(s, z)ds$  we obtain

$$\begin{aligned}
 A_{i,T,5}^{(N)2} &\leq \frac{T}{N^2} \sum_{j,l=1}^N W(x_i, x_j)W(x_i, x_l)\Gamma(x_j, x_l) \\
 &\quad - \frac{2T}{N} \sum_{j=1}^N \int_I W(x_i, x_j)W(x_i, y)\Gamma(x_j, y)\nu(dy) \\
 &\quad + T \int_{I^2} W(x_i, y)W(x_i, z)\Gamma(y, z)\nu(dy)\nu(dz), \quad (3.8.43)
 \end{aligned}$$

which is equal to

$$T \int_{I \times I} W(x_i, y)W(x_i, z)\Gamma(y, z) (\nu^{(N)}(dy)\nu^{(N)}(dz) - 2\nu^{(N)}(dy)\nu(dz) + \nu(dy)\nu(dz))$$

and  $T(\epsilon_{i,1} + \epsilon_{i,2})$ , recognising  $\epsilon_{i,1}$  and  $\epsilon_{i,2}$  from Corollary 3.8.7. Taking the square root and then summing on  $i$  or taking the supremum, (3.8.7) and (3.8.6) follow.

**Remark 3.8.9.** If we ask for more regularity of  $W$ , we can have a more direct proof of (3.8.7). Assume that there exist  $L_W > 0$  and  $M_W > 0$  such that

$$\sup_{x \in I} \sup_{y \neq y'} \frac{|W(x, y) - W(x, y')|}{\|y - y'\|} \leq L_W \text{ and } \sup_{x, y \in I} |W(x, y)| \leq M_W.$$

Hypothesis 3.2.6 is trivially satisfied with  $\vartheta = 1$  and  $C_w = L_W$ , which implies that  $\lambda$  is uniformly Lipschitz continuous in the second variable (in (3.2.7),  $\phi(x) = 2\|x\|$ ). We show first that  $F$  defined above in (3.8.39) is also uniformly Lipschitz continuous in the second variable: for any  $(x, y, y', s) \in I^3 \times [0, T]$ ,

$$\begin{aligned}
 F(x, y, s) - F(x, y', s) &= (W(x, y) - W(x, y'))\gamma(s, y) \\
 &\quad + W(x, y') \int_0^s h(s-u) (\lambda_u(y) - \lambda_u(y')) du,
 \end{aligned}$$

then

$$\begin{aligned}
 & |F(x, y, s) - F(x, y', s)| \\
 & \leq |W(x, y) - W(x, y')| |\gamma(s, y)| + |W(x, y')| \int_0^s |h(s-u)| |\lambda_u(y) - \lambda_u(y')| du \\
 & \leq \|h\|_{T,1} \|y - y'\| (\|\lambda\|_{[0,T] \times I, \infty} L_W + 2M_W C_\lambda) =: L_F \|y - y'\|,
 \end{aligned}$$

with  $L_F > 0$  independent of the choice of  $s$  and  $x$ . Recall (3.8.14), as  $F$  is uniformly Lipschitz continuous in the second variable with constant  $L_F$ , we have

$$\begin{aligned}
 \frac{1}{N} \sum_{i=1}^N A_{i,T,5}^{(N)} & \leq L_F \sup_{g \in BL} \int_0^T \left| \int_I g(y) (\nu^{(N)}(dy) - \nu(dy)) \right| ds \\
 & \leq T L_F d_{BL}(\nu^{(N)}, \nu) \xrightarrow{N \rightarrow \infty} 0
 \end{aligned}$$

by Varadarajan Theorem (see (Dudley, 2002, Theorem 11.4.1) and Varadarajan (1958)).

### 3.8.4.3 Proof of Proposition 3.8.4 for Scenario (2)

Recall that  $I = [0, 1]$ ,  $x_i^{(N)} = \frac{i}{N}$ , and  $\nu(dx) = dx$ . We focus on the case  $W$  continuous. When  $W$  is piecewise continuous, the same results follow as we can work on each rectangle where  $W$  can be extended to a continuous function, and these rectangles are in finite number.

**Proof of (3.8.4)** Using Remark 3.3.3 and (3.3.3), we have

$$\begin{aligned}
 d_{\square}(W^{\mathcal{G}_N^{(2)}}, W) & \leq \|W^{\mathcal{G}_N^{(2)}} - W\|_{\infty \rightarrow 1} = \sup_{\|g\|_{\infty} \leq 1} \int \left| \int (W^{\mathcal{G}_N^{(2)}} - W)(x, y) g(y) dy \right| dx \\
 & \leq \int \int \left| (W^{\mathcal{G}_N^{(2)}} - W)(x, y) \right| dx dy = \|W^{\mathcal{G}_N^{(2)}} - W\|_{L^1, [0,1]^2} \\
 & = \sum_{i,j=1}^N \int_{\frac{i-1}{N}}^{\frac{i}{N}} \int_{\frac{j-1}{N}}^{\frac{j}{N}} \left| W\left(\frac{i}{N}, \frac{j}{N}\right) - W(x, y) \right| dx dy.
 \end{aligned}$$

As  $W$  is continuous on the compact  $[0, 1]^2$  in this scenario (2), it is uniformly continuous due to Heine-Cantor theorem thus for any  $\varepsilon > 0$ , there exists  $\eta > 0$  such that  $|x - x'| + |y - y'| \leq \eta \Rightarrow |W(x, y) - W(x', y')| < \varepsilon$ . For  $N$  large enough,  $\frac{1}{N} < \eta$  and then (3.8.4) holds as  $d_{\square}(W^{\mathcal{G}_N^{(2)}}, W) \leq \sum_{i,j=1}^N \int_{\frac{i-1}{N}}^{\frac{i}{N}} \int_{\frac{j-1}{N}}^{\frac{j}{N}} \varepsilon dx dy = \varepsilon$ .

**Proof of (3.8.5)** Recall that

$$\|W^{\mathcal{G}_N^{(2)}} - W\|_{\infty \rightarrow \infty, \nu} = \sup_{\|g\|_{\infty} \leq 1} \sup_{u \in [0,1]} \left| \int_0^1 (W^{\mathcal{G}_N^{(2)}}(u, v) - W(u, v)) g(v) dv \right|.$$

As done for (3.8.4), we use the uniform continuity of  $W$ : for any  $\varepsilon > 0$ , we take  $\eta > 0$  such that  $|x - x'| + |y - y'| \leq \eta \Rightarrow |W(x, y) - W(x', y')| < \varepsilon$ . Fix  $g$

such that  $\|g\|_\infty \leq 1$  and  $u \in ]0, 1]$ , for any  $N$  there exists a unique  $i$  such that  $u \in B_{N,i} = \left( \frac{i-1}{N}, \frac{i}{N} \right]$ . For  $N$  large enough,  $\frac{2}{N} < \eta$  and we have then

$$\begin{aligned} \int_0^1 \left( W^{\mathcal{G}_N^{(2)}}(u, v) - W(u, v) \right) g(v) dv &= \sum_{j=1}^N \int_{B_j} \left( W \left( \frac{i}{N}, \frac{j}{N} \right) - W(u, v) \right) g(v) dv \\ &\leq \sum_{j=1}^N \int_{\frac{j-1}{N}}^{\frac{j}{N}} \left| W \left( \frac{i}{N}, \frac{j}{N} \right) - W(u, v) \right| |g(v)| dv \leq \varepsilon, \end{aligned}$$

independently from the choices of  $g$  and  $u$ : we have shown that in this Scenario,  $\|W^{\mathcal{G}_N^{(2)}} - W\|_{\infty \rightarrow \infty, \nu} \xrightarrow{N \rightarrow \infty} 0$ .

**Proof of (3.8.7) and (3.8.6)** As  $W$  is continuous on  $[0, 1]^2$  and  $(s, y) \mapsto \gamma(s, y) = \int_0^s h(s-u) \lambda_u(y) du$  is also continuous on  $[0, T] \times [0, 1]$  as a convolution between  $h$  locally integrable and  $\lambda$  continuous, the application  $(x, y, s) \mapsto F(x, y, s) = W(x, y) \gamma(s, y)$  is continuous on the compact set  $K = [0, 1] \times [0, 1] \times [0, T]$ , it is uniformly continuous due to Heine-Cantor theorem. Then, for  $\varepsilon > 0$ , there exists  $\eta > 0$  such that for any  $(x, y, s)$  and  $(x', y', s')$  in  $K$ ,  $|x - x'| + |y - y'| + |s - s'| \leq \eta \Rightarrow |F(x, y, s) - F(x', y', s')| < \varepsilon$ . For  $N$  large enough,  $\frac{1}{N} < \eta$  and we have then

$$\begin{aligned} A_{i,T,5}^{(N)} &= \int_0^T \left| \sum_{j=1}^N \int_{\frac{j-1}{N}}^{\frac{j}{N}} F(x_i, x_j, s) dy - \sum_{j=1}^N \int_{\frac{j-1}{N}}^{\frac{j}{N}} F(x_i, y, s) dy \right| ds \\ &\leq \int_0^T \sum_{j=1}^N \int_{\frac{j-1}{N}}^{\frac{j}{N}} |F(x_i, x_j, s) - F(x_i, y, s)| dy ds \leq T\varepsilon. \end{aligned} \quad (3.8.44)$$

Summing on  $i$  or taking the supremum, (3.8.7) and (3.8.6) follow.

## 3.9 The empirical measure and the spatial profile (proofs)

### 3.9.1 Proof of Theorem 3.3.15

We prove the convergence of  $\mathbf{E} [d_{BL}(\mu_N, \mu_\infty)] \xrightarrow{N \rightarrow \infty} 0$ . Some of the following arguments come from [Chevallier et al. \(2019\)](#). We consider  $\mathbb{D}([0, T], \mathbb{N})$  with the distance  $d_0$  introduced in ([Billingsley, 1968](#), §14) which makes it complete, and we have for any  $\eta, \zeta$  in  $\mathbb{D}([0, T], \mathbb{N})$ ,  $d_0(\eta, \zeta) \leq \sup_{t \leq T} |\eta(t) - \zeta(t)|$ . Recall that  $d_{BL}(\mu_N, \mu_\infty) = \sup_{\|\phi\|_{BL} \leq 1} \left| \int \phi d\mu_N - \int \phi d\mu_\infty \right|$ . We start by proving that for any  $\phi$  fixed,  $\mathbf{E} \left| \int \phi (d\mu_N - d\mu_\infty) \right| \xrightarrow{N \rightarrow \infty} 0$ . By an argument of compactness, we show how it implies (3.3.20).

*Step 1 - Convergence when  $\phi$  is fixed.* We fix  $\phi$  a real-valued function on  $S$  (recall that  $S := \mathbb{D}([0, T], \mathbb{N}) \times I$ ) such that  $\|\phi\|_{BL} \leq 1$ . Then with the coupling

introduced in Definition 3.3.1:

$$\begin{aligned} \mathbf{E} \left| \int \phi(d\mu_N - d\mu_\infty) \right| &= \mathbf{E} \left| \frac{1}{N} \sum_{i=1}^N \phi \left( Z_i^{(N)}, x_i \right) - \int \phi(\eta, x) P_{[0,T],\infty}(d\eta|x) \nu(dx) \right| \\ &\leq \mathbf{E} \left| \frac{1}{N} \sum_{i=1}^N \int \phi(\eta, x_i) P_{[0,T],\infty}(d\eta|x_i) - \int \phi(\eta, x) P_{[0,T],\infty}(d\eta|x) \nu(dx) \right| \\ &\quad + \mathbf{E} \left| \frac{1}{N} \sum_{i=1}^N \left( \phi(\bar{Z}_i, x_i) - \int \phi(\eta, x_i) P_{[0,T],\infty}(d\eta|x_i) \right) \right| \\ &\quad + \mathbf{E} \left| \frac{1}{N} \sum_{i=1}^N \left( \phi \left( Z_i^{(N)}, x_i \right) - \phi(\bar{Z}_i, x_i) \right) \right| := A + B + C. \end{aligned}$$

The term  $A$  is treated easily with Theorems 3.3.10 or 3.3.12: as  $\phi$  is Lipschitz continuous and  $\|\phi\|_L \leq 1$ ,

$$A \leq \frac{1}{N} \sum_{i=1}^N \mathbf{E} \left[ d_0 \left( Z_i^{(N)}, \bar{Z}_i \right) \right] \leq \frac{1}{N} \sum_{i=1}^N \mathbf{E} \left[ \sup_{t \leq T} \left| Z_i^{(N)}(t) - \bar{Z}_i(t) \right| \right] \xrightarrow{N \rightarrow \infty} 0.$$

To treat  $B$ , we set for each  $i \in \llbracket 1, N \rrbracket$   $G_i := \phi(\bar{Z}_i, x_i)$ , it is a random variable with expectation  $\int \phi(\eta, x_i) P_{[0,T],\infty}(d\eta|x_i)$ . We have then applying Lemma 3.A.2,  $B \leq \frac{1}{N} \sqrt{\sum_{i=1}^N \text{Var}(G_i)}$ . To calculate  $\text{Var}(G_i)$ , let  $(\tilde{Z}_i(t))_{0 \leq t \leq T}$  be an independent copy of  $(\bar{Z}_i(t))_{0 \leq t \leq T}$  and set  $\tilde{G}_i := g(\tilde{Z}_i, x_i)$ , then denoting by  $\tilde{E}$  the expectation taken with respect to  $\tilde{G}_i$ , we have

$$\text{Var}(G_i) = \mathbf{E} \left[ (G_i - \mathbf{E}[G_i])^2 \right] = \mathbf{E} \left[ \tilde{E} \left[ G_i - \tilde{G}_i \right]^2 \right] \leq \mathbf{E} \left[ \tilde{E} \left[ (G_i - \tilde{G}_i)^2 \right] \right]$$

by Jensen's inequality. We have, as  $\|g\|_L \leq 1$ ,

$$\begin{aligned} \tilde{E} \left[ (G_i - \tilde{G}_i)^2 \right] &\leq \tilde{E} \left[ d_0 \left( \bar{Z}_i, \tilde{Z}_i \right)^2 \right] \leq \tilde{E} \left[ \left( \sup_{0 \leq t \leq T} \left| \bar{Z}_i(t) - \tilde{Z}_i(t) \right| \right)^2 \right] \\ &\leq 2\bar{Z}_i(T)^2 + 2\tilde{E} \left[ \tilde{Z}_i(T)^2 \right] \end{aligned}$$

as the processes are increasing. Thus we obtain  $\text{Var}(G_i) \leq 4\mathbf{E} \left[ \bar{Z}_i(T)^2 \right]$ . As  $Z_i(T)$  is a Poisson random variable with rate  $\int_0^T \lambda_t(x_i) dt$ ,

$$\mathbf{E} \left[ \tilde{Z}_i(T)^2 \right] = \text{Var} \left( \tilde{Z}_i(T) \right) + \left( \mathbf{E} \left[ \tilde{Z}_i(T) \right] \right)^2 = \int_0^T \lambda_t(x_i) dt + \left( \int_0^T \lambda_t(x_i) dt \right)^2$$

which is finite as  $\lambda$  is bounded (Theorem 3.2.7). We have then shown that  $B \xrightarrow{N \rightarrow \infty} 0$ . To treat  $C$ , note that it can be rewritten as

$$C = \left| \int \int \phi(\eta, x) P_{[0,T],\infty}(d\eta|x) \left( \nu^{(N)}(dx) - \nu(dx) \right) \right|.$$

We denote by  $h$  the bounded function  $h(x) = \int \phi(\eta, x) P_{[0,T],\infty}(d\eta|x)$ . Under Scenario (1),  $C = \left| \frac{1}{N} \sum_{i=1}^N h(x_i) - \int_I h(x) \nu(dx) \right| \xrightarrow[N \rightarrow \infty]{} 0$  by the Law of Large Numbers. Under Scenario (2),  $C = \left| \frac{1}{N} \sum_{i=1}^N h\left(\frac{i}{N}\right) - \int_0^1 h(x) dx \right|$  and we recognise a Riemann sum: it suffices to show that  $h$  is continuous to have  $C \xrightarrow[N \rightarrow \infty]{} 0$ . Fix  $x$  in  $I$  and consider a sequence  $(x_n)$  such that  $x_n \xrightarrow[n \rightarrow \infty]{} x$ . We have

$$\begin{aligned} |h(x_n) - h(x)| &\leq \int |\phi(\eta, x_n) - \phi(\eta, x)| P_{[0,T],\infty}(d\eta|x_n) \\ &\quad + \left| \int \phi(\eta, x) (P_{[0,T],\infty}(d\eta|x) - P_{[0,T],\infty}(d\eta|x_n)) \right|. \end{aligned}$$

We deal with the first term: by the Lipschitz continuity of  $\phi$  and the fact that  $P_{[0,T],\infty}(\cdot|x_n)$  is a probability measure, we have

$$\int |\phi(\eta, x_n) - \phi(\eta, x)| P_{[0,T],\infty}(d\eta|x_n) \leq \|x - x_n\| \xrightarrow[n \rightarrow \infty]{} 0.$$

As  $x$  is fixed, to have the second term converging to 0, we show that for any function  $\psi$  with Lipschitz constant  $\|\psi\|_L \leq 1$  defined on  $\mathbb{D}([0, T], \mathbb{N})$ , the function  $\rho(y) := \int \psi(\eta) P_{[0,T],\infty}(d\eta|y)$  is continuous on  $I$ : let  $\pi$  be a random Poisson measure with intensity  $dsdz$  on  $\mathbb{R}_+ \times \mathbb{R}_+$ , and for each  $y \in I$  construct a Poisson point process  $\bar{Z}^y$  on  $[0, T]$  with intensity  $(\lambda_t(y))$  by taking  $\bar{Z}^y(t) = \int_0^t \int_0^\infty \mathbf{1}_{z \leq \lambda_s(y)} \pi(ds, dz)$ . Then, as  $\psi$  is Lipschitz continuous,

$$\begin{aligned} |\rho(x) - \rho(x_n)| &= |\mathbf{E}[\psi(\bar{Z}^{x_n}) - \psi(\bar{Z}^x)]| \leq \mathbf{E}[d_0(\bar{Z}^{x_n}, \bar{Z}^x)] \\ &\leq \mathbf{E} \left[ \sup_{0 \leq t \leq T} |\bar{Z}^{x_n}(t) - \bar{Z}^x(t)| \right] \leq \mathbf{E} \left[ \int_0^t |d(\bar{Z}^{x_n}(s) - \bar{Z}^x(s))| \right] \\ &\leq \mathbf{E} \left[ \int_0^t \int_0^\infty |\mathbf{1}_{z \leq \lambda_s(x_n)} - \mathbf{1}_{z \leq \lambda_s(x)}| \pi(ds, dz) \right] \\ &\leq \int_0^t |\lambda_s(x_n) - \lambda_s(x)| ds \leq T \|x - x_n\|^\vartheta \xrightarrow[n \rightarrow \infty]{} 0, \end{aligned}$$

with (3.2.7). Then  $\rho$  is indeed continuous on  $I$ , and so is  $h$  hence  $C \xrightarrow[N \rightarrow \infty]{} 0$ . We have shown that for any function  $\phi$  on  $S$  such that  $\|\phi\|_{BL} \leq 1$ , we have

$$\mathbf{E} \left| \int \phi(d\mu_N - d\mu_\infty) \right| \xrightarrow[N \rightarrow \infty]{} 0. \quad (3.9.1)$$

*Step 2 - Approximation of any  $\phi$  by a finite set of functions and conclusion.* To derive (3.3.20), we use an argument from (Luçon, 2020, Lemma 4.5) and (Dudley, 2002, Theorem 11.3.3). For all  $\varepsilon > 0$ , there exists a compact set  $K \subset S$  with  $\mu_\infty(K) > 1 - \varepsilon$ . The set of functions  $B := \{\phi|_K, \|\phi\|_{BL} \leq 1\}$ , restricted to  $K$  is a compact set by Arzela-Ascoli Theorem, hence there exists  $k \geq 1$  and  $k$  functions in  $B$ ,  $\phi_1, \dots, \phi_k$ , such that for any  $\phi$  satisfying  $\|\phi\|_{BL} \leq 1$ , there exists  $j \leq k$  that

verifies  $\sup_{y \in K} |\phi(y) - \phi_j(y)| \leq \varepsilon$ . We denote by  $K^\varepsilon := \{z \in S, d_S(z, K) < \varepsilon\}$ . Then  $\sup_{z \in K^\varepsilon} |\phi(z) - \phi_j(z)| < 3\varepsilon$  as for any  $z \in K^\varepsilon$ , we can find  $y_z \in K$  such that  $d_S(z, y_z) < \varepsilon$  and

$$\begin{aligned} |\phi(z) - \phi_j(z)| &\leq |\phi(z) - \phi(y_z)| + |\phi(y_z) - \phi_j(y_z)| + |\phi_j(y_z) - \phi_j(z)| \\ &\leq \|\phi\|_L d_S(z, y_z) + \varepsilon + \|\phi_j\|_L d_S(z, y_z) \leq 3\varepsilon. \end{aligned}$$

We introduce the function  $g(z) = \max\left(0, 1 - \frac{d_S(z, K)}{\varepsilon}\right)$  on  $S$ . Note that  $\mathbf{1}_K \leq g \leq \mathbf{1}_{K^\varepsilon}$  and  $g$  is bounded and Lipschitz continuous. Then, integrating on  $\mu_N$ , we obtain  $\mu_N(K^\varepsilon) \geq \int g d\mu_N$ . We put together all the previous bounds to have, for any  $\phi$  such that  $\|\phi\|_{BL} \leq 1$ :

$$\begin{aligned} \left| \int \phi (d\mu_\infty - d\mu_N) \right| &\leq \int |\phi - \phi_j| (d\mu_\infty + d\mu_N) + \left| \int \phi_j (d\mu_\infty - d\mu_N) \right| \\ &\leq \int_{K^\varepsilon} |\phi - \phi_j| (d\mu_\infty + d\mu_N) + \int_{S-K^\varepsilon} |\phi - \phi_j| (d\mu_\infty + d\mu_N) + \left| \int \phi_j (d\mu_\infty - d\mu_N) \right| \\ &\leq 3\varepsilon \cdot 2 + 2\mu_\infty (S - K^\varepsilon) + 2\mu_N (S - K^\varepsilon) + \left| \int \phi_j (d\mu_\infty - d\mu_N) \right|. \end{aligned}$$

Hence, taking the supremum on such function  $\phi$  we obtain

$$\sup_{\phi, \|\phi\|_{BL} \leq 1} \left| \int \phi (d\mu_N - d\mu_\infty) \right| \leq 8\varepsilon + 2 \left( 1 - \int g d\mu_N \right) + \max_{1 \leq j \leq k} \left| \int \phi_j (d\mu_\infty - d\mu_N) \right|.$$

Using (3.9.1), for  $N$  large enough  $\mathbb{E}[\int g d\mu_N] > \int g d\mu_\infty - \varepsilon$ . As  $\int g d\mu_\infty \geq \mu_\infty(K) \geq 1 - \varepsilon$ , we have  $\mathbb{E}[\int g d\mu_N] > 1 - 2\varepsilon$  and then  $\mathbf{E}[d_{BL}(\mu_N, \mu_\infty)] \leq 12\varepsilon + \mathbf{E}[\max_{1 \leq j \leq k} |\int \phi_j (d\mu_\infty - d\mu_N)|]$ . Using (3.9.1), as there is a finite number of functions considered, we have that  $\mathbf{E}[\max_{1 \leq j \leq k} |\int \phi_j (d\mu_\infty - d\mu_N)|] \xrightarrow{N \rightarrow \infty} 0$  which concludes the proof of (3.3.20).  $\square$

### 3.9.2 Proof of Proposition 3.3.17

We show the convergence of the spatial profile  $U_N$ , when the positions are regularly distributed on  $[0, 1]$  and  $W$  is continuous. We have

$$\begin{aligned} \mathbf{E} \left[ \int_0^T \int_0^1 |U_N(t, x) - u_t(x)| dx dt \right] &\leq \mathbf{E} \left[ \int_0^T \frac{1}{N} \sum_{i=1}^N |U_{i,N}(t) - u_t(x_i)| dt \right] \\ &\quad + \int_0^T \int_0^1 \left| \sum_{i=1}^N \mathbf{1}_{(\frac{i-1}{N}, \frac{i}{N}]}(x) (u_t(x_i) - u_t(x)) \right| dx dt. \end{aligned}$$

The first term is dealt with the proof of Theorem 3.3.10: recall (3.8.9), we recognise

$$\int_0^T \mathbf{E} [|U_{i,N}(t) - u(t, x_i)|] dt \leq \left( \sum_{k=1}^5 A_{i,T,k}^{(N)} \right),$$

and we have showed that  $\frac{1}{N} \sum_{i=1}^N A_{i,T,k}^{(N)} \xrightarrow[N \rightarrow \infty]{} 0$   $\mathbb{P}$ -almost surely for each  $k = 1, \dots, 5$ .

We then have  $\mathbf{E} \left[ \int_0^T \frac{1}{N} \sum_{i=1}^N |U_{i,N}(t) - u_t(x_i)| dt \right] \xrightarrow[N \rightarrow \infty]{} 0$   $\mathbb{P}$ -almost surely. The other term is treated easily: as  $(t, x) \mapsto u_t(x)$  is continuous on the compact set  $[0, T] \times [0, 1]$  it is uniformly continuous. Fix  $\varepsilon > 0$ , then there exists  $\eta > 0$  such that if  $\|t - t'\| + \|x - x'\| \leq \eta$ ,  $|u(t, x) - u(t', x')| \leq \frac{\varepsilon}{T}$ . We have then for  $N$  large enough (such that  $\frac{1}{N} \leq \eta$ ):

$$\begin{aligned} \int_0^T \int_0^1 \left| \sum_{i=1}^N \mathbf{1}_{\left(\frac{i-1}{N}, \frac{i}{N}\right]}(x) (u_t(x_i) - u_t(x)) \right| dx dt &= \int_0^T \sum_{i=1}^N \int_{\frac{i-1}{N}}^{\frac{i}{N}} |u_t(x_i) - u_t(x)| dx dt \\ &\leq \int_0^T \frac{\varepsilon}{T} dt = \varepsilon, \end{aligned}$$

which concludes the proof.  $\square$

## 3.10 Behavior in large time limit - Linear case (proofs)

### 3.10.1 Proof of Theorem 3.4.4

We show that in the subcritical case,  $\lambda_t(x)$  has a large time limit given by (3.4.8). Assumption (3.4.7) implies the existence of some  $n_0$  such that  $\|h\|_1^{n_0} \|T_W^{n_0}\| < 1$ .

*Step 1* - We show existence and uniqueness of  $\ell$  by applying Banach fixed-point Theorem. We consider the map defined on  $\mathcal{C}_b(I, \mathbb{R})$  (the set of bounded continuous functions defined on  $I$ ):

$$\begin{aligned} F : g &\longmapsto F(g) \text{ such that for all } x \in I, \\ F(g)(x) &= v_\infty(x) + \|h\|_1 \int_I W(x, y) g(y) \nu(dy). \end{aligned}$$

As  $v_\infty$  is bounded on  $I$ ,  $F(g)$  is bounded for any  $g \in \mathcal{C}_b(I, \mathbb{R})$  by  $\|v_\infty\|_\infty + \|h\|_1 \|g\|_\infty C_W^{(1)} < \infty$ . We check now that for any  $g$ ,  $F(g)$  is continuous. Let  $(x, z) \in I \times I$ . We have as  $v_\infty$  is Lipschitz continuous and using (3.2.5), for any  $g \in \mathcal{C}_b(I, \mathbb{R})$ :

$$\begin{aligned} |F(g)(x) - F(g)(z)| &\leq |v_\infty(x) - v_\infty(z)| + \|h\|_1 \left| \int_I (W(x, y) - W(z, y)) g(y) \nu(dy) \right| \\ &\leq \|v_\infty\|_L \|x - z\| + \|h\|_1 \|g\|_\infty \|x - z\|^\vartheta. \end{aligned}$$

We have then shown the existence of a constant  $C_g$  independent of the choice of  $(x, z)$  such that  $|F(g)(x) - F(g)(z)| \leq C_g \phi(\|x - z\|)$ . Hence,  $\mathcal{C}_b(I, \mathbb{R})$  is stable by  $F$ .



We are going to prove that  $F$  admits a **unique fixed point**, which is  $\ell$  satisfying (3.4.8). To do it, we show that some iteration of  $F$  is contractive, and then the Banach fixed-point Theorem gives the result. Let  $g$  and  $\tilde{g}$  be two functions in  $\mathcal{C}_b(I, \mathbb{R})$ . As  $Fg = v_\infty + \|h\|_1 T_W g$ , we have immediately that  $F^{n_0} g = \sum_{k=0}^{n_0-1} \|h\|_1^k T_W^k v_\infty + \|h\|_1^{n_0} T_W^{n_0} g$ . Then

$$\|F^{n_0} g - F^{n_0} \tilde{g}\| = \|h\|_1^{n_0} \|T_W^{n_0}(g - \tilde{g})\| \leq \|h\|_1^{n_0} \|T_W^{n_0}\| \|g - \tilde{g}\|_\infty.$$

As  $n_0$  is chosen such that  $\|h\|_1^{n_0} \|T_W^{n_0}\| < 1$ ,  $F^{n_0}$  is contractive, thus has a unique fixed point which is also the unique fixed point of  $F$  in  $\mathcal{C}_b(I, \mathbb{R})$  that we call  $\ell$ , solution to (3.4.9). Note that such a  $\ell$  is necessarily nonnegative, as the iterative map  $F$  preserves positivity.

*Step 2* - Let us show that under the present hypotheses,  $\sup_{t \geq 0} \sup_{x \in I} |\lambda_t(x)| < \infty$ . As  $\lambda_t(x) = v_t(x) + h * (T_W \lambda_t)(x)$ ,  $T_W \lambda_t(x) = T_W v_t(x) + h * T_W^2 \lambda_t(x)$  and the iteration gives  $\lambda_t(x) = (\sum_{k=0}^{n_0-1} h^{*k} * T_W^k v_t)(x) + h^{*n_0} * T_W^{n_0} \lambda_t(x)$  for any  $(t, x) \in \mathbb{R}_+ \times I$ , hence  $\|\lambda_t\|_\infty \leq C + \|h\|_1^{n_0} \|T_W^{n_0}\| \|\lambda_t\|_\infty$  with  $C = C(v, h, W)$  a positive constant.

As we are in the subcritical case, it gives then  $\sup_{t \geq 0} \|\lambda_t\|_\infty \leq \frac{C(v, h, W)}{1 - \|h\|_1^{n_0} \|T_W^{n_0}\|} < \infty$ . As  $(t, x) \mapsto \lambda_t(x)$  is then continuous and bounded on  $\mathbb{R}_+ \times I$ , we can define its (temporal) Laplace transform: for any  $x \in I$  and  $z > 0$ , let

$$\Lambda(z, x) := \int_0^\infty e^{-tz} \lambda_t(x) dt. \quad (3.10.1)$$

Let us study  $z\Lambda(z, x)$ . We have, for any  $x \in I$  and  $z > 0$ ,

$$z\Lambda(z, x) = \int_0^\infty z e^{-tz} \lambda_t(x) dt = \lambda_0(x) + \int_0^\infty e^{-tz} \frac{\partial \lambda_t}{\partial t}(x) dt.$$

Suppose that we are able to show that  $I(x) := \int_0^\infty \left| \frac{\partial \lambda_t}{\partial t}(x) \right| < \infty$  for some  $x$ .

Then, by dominated convergence theorem,  $\int_0^\infty e^{-tz} \frac{\partial \lambda_t}{\partial t}(x) dt$  converges as  $z \rightarrow 0$

to the finite limit  $\int_0^\infty \frac{\partial \lambda_t}{\partial t}(x) dt$ . This implies in particular that  $\lambda_t(x)$  has a finite limit as  $t \rightarrow \infty$ , and we have in this case  $\lim_{z \rightarrow 0} z\Lambda(z, x) = \lim_{t \rightarrow \infty} \lambda_t(x)$ . We have, by

integrating by parts that  $\frac{\partial \lambda_t}{\partial t}(x)$  is equal to

$$\begin{aligned} & \frac{\partial v_t}{\partial t}(x) + h(0) \int_I W(x, y) \lambda_t(y) \nu(dy) + \int_0^t \int_I W(x, y) h'(t-s) \lambda_s(y) \nu(dy) ds \\ &= \frac{\partial v_t}{\partial t}(x) + \int_I W(x, y) h(t) \lambda_0(y) \nu(dy) + \int_0^t \int_I W(x, y) h(t-s) \frac{\partial \lambda_s}{\partial s}(y) \nu(dy) ds, \end{aligned}$$

where we used Theorem 3.2.7 for the regularity of  $\frac{\partial \lambda_s}{\partial s}$ . We also know from Theorem 3.2.7 that  $(t, x) \rightarrow \frac{\partial \lambda_t}{\partial t}(x)$  is bounded on  $[0, T] \times I$  for any  $T > 0$ , which implies

that for any  $A > 0$ ,  $\sup_{x \in I} \int_0^A \left| \frac{\partial \lambda_t}{\partial t}(x) \right| dt < \infty$ . Integrating on  $[0, A]$ , we have using (3.4.2)

$$\begin{aligned} \int_0^A \left| \frac{\partial \lambda_t}{\partial t}(x) \right| dt &\leq C_v + \|h\|_1 \|\lambda\|_\infty D(x) \\ &\quad + \int_0^A \int_0^t \int_I W(x, y) h(t-s) \left| \frac{\partial \lambda_s}{\partial s}(y) \right| \nu(dy) ds dt. \end{aligned}$$

Yet with a change in the bounds of the integrals ( $0 \leq s \leq t \leq A$ ), we have

$$\begin{aligned} &\int_0^A \int_0^t \int_I W(x, y) h(t-s) \left| \frac{\partial \lambda_s}{\partial s}(y) \right| \nu(dy) ds dt \\ &= \int_I W(x, y) \int_0^A \int_0^t h(t-s) \left| \frac{\partial \lambda_s}{\partial s}(y) \right| ds dt \nu(dy) \\ &= \int_I W(x, y) \int_0^A \left( \int_s^A h(t-s) dt \right) \left| \frac{\partial \lambda_s}{\partial s}(y) \right| ds \nu(dy) \\ &\leq \|h\|_1 \int_I W(x, y) \int_0^A \left| \frac{\partial \lambda_s}{\partial s}(y) \right| ds \nu(dy). \end{aligned}$$

Setting  $I_A(x) := \int_0^A \left| \frac{\partial \lambda_s}{\partial s}(x) \right| ds$  and  $C = C_v + \|h\|_1 \|\lambda\|_\infty C_W^{(1)}$ , we have shown that  $I_A(x) \leq C + \|h\|_1 (T_W I_A)(x)$  and by iteration

$$I_A(x) \leq \sum_{k=0}^{n_0-1} C \|h\|_1^k C_W^{(1)k} + \|h\|_1^{n_0} T_W^{n_0} I_A(x) \leq \sum_{k=0}^{n_0-1} C \|h\|_1^k C_W^{(1)k} + \|h\|_1^{n_0} \|T_W^{n_0}\| \|I_A\|_\infty,$$

and then  $\|I_A\|_\infty \leq \frac{C(h, v, W)}{1 - \|h\|_1^{n_0} \|T_W^{n_0}\|} = C'$ , with  $C'$  a positive constant independent of  $A$ . We can then let  $A \rightarrow \infty$  to obtain  $\sup_{x \in I} I(x) < \infty$ . Hence by dominated convergence  $\lim_{z \rightarrow 0} z\Lambda(z, x)$  exists for any  $x \in I$  and is equal to  $\lim_{t \rightarrow \infty} \lambda_t(x) =: \ell(x)$  that can now be defined. Coming back to the definition on  $\Lambda$ , we do the same for  $v$  and define for any  $x \in I$  and  $z > 0$   $V(z, x) := \int_0^\infty e^{-tz} v_t(x) dt$ . As  $v_t(x) \xrightarrow[t \rightarrow \infty]{} v_\infty(x)$ , note that  $\lim_{z \rightarrow 0} zV(z, x) = v_\infty(x)$ . As  $h$  is integrable in this framework, we can also define its Laplace transform for any  $z \geq 0$  by  $H(z) := \int_0^\infty e^{-tz} h(t) dt$ , with  $H(0) = \|h\|_1$ . Using the fact that the Laplace transform of a convolution is the product of the Laplace transforms, we have for any  $x \in I$  and  $z > 0$

$$z\Lambda(z, x) = zV(z, x) + H(z) \int_I W(x, y) z\Lambda(z, y) \nu(dy). \quad (3.10.2)$$

Letting  $z \rightarrow 0$  in (3.10.2), we obtain that  $\ell$  is solution of the equation (3.4.8).  $\square$

### 3.10.2 Proof of Propositions 3.4.7 and 3.4.8

*Proof of Proposition 3.4.7.* The boundedness of  $T_W$  on  $L^2(I)$  follows from (3.4.15) and Cauchy-Schwarz inequality: for  $g \in L^2(I)$

$$\begin{aligned} \|T_W g\|_2^2 &= \int \left( \int W(x, y) g(y) \nu(dy) \right)^2 \nu(dx) \\ &\leq \int \left( \int W(x, y)^2 \nu(dy) \right) \left( \int g(y)^2 \nu(dy) \right) \nu(dx) \leq \|g\|_2^2 C_W^{(2)}. \end{aligned}$$

It is standard to see that  $T_W$  is compact on  $L^2(I)$  and self-adjoint, by (3.4.16), so that the same result holds readily for  $T_W^p$  for all  $p \geq 1$ . The fact that the spectrum of  $T_W^p$  is made of a countable set of eigenvalues with no other accumulation points than 0 is a mere application of the spectral theorem for compact operators. Let us now prove (3.4.17): first note that it suffices to prove that  $r_2(T_W^2) = r_\infty(T_W^2)$ . Indeed, for any continuous operator  $T$  with spectral radius  $r(T)$ , for all  $p \geq 1$ ,  $r(T^p)^{\frac{1}{p}} = \left( \lim_{n \rightarrow \infty} \|T^{pn}\|^{\frac{1}{n}} \right)^{\frac{1}{p}} = \lim_{n \rightarrow \infty} \|T^{pn}\|^{\frac{1}{pn}} = r(T)$ , so that  $r(T^p) = r(T)^p$ . Hence  $r_2(T_W^2) = r_\infty(T_W^2)$  gives  $r_2(T_W) = r_\infty(T_W)$  and (3.4.17) follows. We prove that  $r_2(T_W^2) = r_\infty(T_W^2)$  by proving that they have the same spectrum. To do so, first note that  $T_W^2 : L^\infty(I) \rightarrow L^\infty(I)$  is compact: consider  $(f_n)_n$  a bounded sequence of  $L^\infty(I)$ . It is then also bounded in  $L^2(I)$ , and as  $T_W : L^2(I) \rightarrow L^2(I)$  is compact, there exists a subsequence  $(f_{\phi(n)})$  such that  $T_W f_{\phi(n)}$  converges in  $L^2(I)$  to a certain  $g$ . Then for any  $x \in I$ ,

$$\begin{aligned} |T_W^2 f_{\phi(n)} - T_W g|(x) &\leq \int_I W(x, y) |T_W f_{\phi(n)}(y) - g(y)| dy \\ &\leq \sqrt{C_W^{(2)}} \|T_W f_{\phi(n)} - g\|_2 \xrightarrow[n \rightarrow \infty]{} 0, \end{aligned}$$

thus  $T_W^2 : L^\infty(I) \rightarrow L^\infty(I)$  is compact. Hence, if one denotes by  $\sigma_\infty(T_W^2)$  and  $\sigma_2(T_W^2)$  the corresponding spectrum of  $T_W^2$  (in  $L^\infty(I)$  and  $L^2(I)$  respectively), we have that each nonzero element of  $\sigma_\infty(T_W^2)$  and  $\sigma_2(T_W^2)$  is an eigenvalue of  $T_W^2$ : let  $\mu \in \sigma_2(T_W^2) \setminus \{0\}$ , there exists  $g \in L^2(I)$  such that  $\mu g = T_W^2 g$ . As

$$|T_W^2 g(x)| = \left| \int_I W(x, y) \int_I W(y, z) g(z) \nu(dz) \nu(dy) \right| \leq C_W^{(1)} \sqrt{C_W^{(2)}} \|g\|_2 < \infty,$$

$g = \frac{1}{\mu} T_W^2 g \in L^\infty(I)$  and  $\mu \in \sigma_\infty(T_W^2)$ . Conversely, let  $\mu \in \sigma_\infty(T_W^2) \setminus \{0\}$ , there exists  $g \in L^\infty(I)$  such that  $\mu g = T_W^2 g$ . As  $L^\infty(I) \subset L^2(I)$ ,  $\mu \in \sigma_2(T_W^2)$ . Hence  $r_2(T_W^2) = r_\infty(T_W^2)$  and (3.4.17) follows.

Let us now prove the second part of Proposition 3.4.7: this is essentially a reformulation of the Jentzsh-Krein-Rutman Theorem (see Theorem 3.A.8): under assumption (3.4.16), the spectral radius  $r_2(T_W^k)$  is an eigenvalue of  $T_W^k$  with a unique normalized eigenfunction  $h_0$  such that  $h_0 > 0$ ,  $\nu$  a.e. on  $I$  and every other eigenvalue  $\mu$  of  $T_W^k$  has modulus  $|\mu| < r_2(T_W^k)$ . It remains to prove that  $h_0^{(k)}$  is in fact continuous and bounded. As  $\|T_W^k h_0^{(k)}\|_\infty \leq \left( C_W^{(1)} \right)^{k-1} \sqrt{C_W^{(2)}} \|h_0^{(k)}\|_2$  (using Cauchy-Schwarz inequality) and  $h_0 = \frac{1}{r(T_W^k)} T_W^k h_0^{(k)}$ ,  $h_0^{(k)}$  is bounded. Condition

(3.2.5) implies that  $T_W h_0^{(k)}$  is continuous on  $I$ , hence  $h_0^{(k)}$  is a positive continuous function on  $I$ .  $\square$

*Proof of Proposition 3.4.8.* Consider for the moment the case  $k = 1$  (see (3.4.14)). One benefit of working in  $L^2(I)$  instead of  $L^\infty(I)$  is to take advantage of the Hilbert structure associated to  $T_W$ : we know from the spectral theorem, that we can complete  $h_0$  in an Hilbert orthonormal basis  $(h_0, h_1, \dots)$  of eigenvectors in  $L^2(I)$  associated to the eigenvalues  $(\mu_0 = r_\infty, \mu_1, \mu_2, \dots)$  with  $\sup_{k \geq 1} |\mu_k| =: \tilde{r}(T_W) < r_\infty$ . We denote by  $P_0$  the projection on  $\text{Vect}(h_0)$  and  $P_1 = Id - P_0$ : for any  $g \in L^2(I)$ ,  $P_0 g = \langle g, h_0 \rangle h_0 =: p_0(g) h_0$  (with  $p_0(g) \in \mathbb{R}$ ) and  $P_1 g = \sum_{n \geq 1} \langle g, h_n \rangle h_n$ . The strategy of proof of Proposition 3.4.8 is then to analyse separately the dynamics of  $P_0 \lambda$  and  $P_1 \lambda$  for  $\lambda$  solution to (3.4.1). Concerning  $P_0 \lambda$ , as  $P_0$  projects onto  $h_0$ , eigenfunction associated to the dominant eigenvalue  $r_2(T_W)$ , its analysis reduces to a simple one-dimensional linear convolution equation, whose behavior in large time has been analysed in details (see (Delattre *et al.*, 2016a, Lemma 26) or (Feller, 1941, Th 4)). The second step is to show that the contribution of  $P_1 \lambda$  remains of lower order as  $t \rightarrow \infty$ .

We focus first on the dynamics of  $P_0 \lambda$ . Using (3.4.1), as  $T$  and  $P_0$  commute,

$$\begin{aligned} p_0(\lambda_t) h_0 &= P_0 \lambda_t = P_0 v_t + \int_0^t h(t-s) T_W P_0 \lambda_s ds \\ &= \left( p_0(v_t) + r_\infty \int_0^t h(t-s) p_0(\lambda_s) ds \right) h_0. \end{aligned}$$

As  $h_0(x) > 0$  everywhere (since  $h_0$  is continuous), we obtain that  $p_0(\lambda_t)$  solves the convolution equation in  $\mathbb{R}$

$$p_0(\lambda_t) = p_0(v_t) + r_\infty \int_0^t h(t-s) p_0(\lambda_s) ds. \quad (3.10.3)$$

Theorem 3.A.9 gives then that  $p_0(\lambda_t) \sim_{t \rightarrow \infty} C e^{\sigma_r t}$  where  $C > 0$  depends on the parameter functions and  $\sigma_r > 0$  verifies  $r_\infty \int_0^\infty e^{-\sigma_r t} h(t) dt = 1$ . We focus now on the other projection,  $P_1 \lambda$ . We project on the rest of the space and take the norm  $L^2(I)$ :  $\|P_1 \lambda_t\|_2 \leq \|P_1 v_t\|_2 + \int_0^t h(t-s) \|T_W P_1 \lambda_s\|_2 ds$ . As

$$T_W P_1 \lambda_s = T_W \left( \sum_{n \geq 1} \langle P_1 \lambda_s, h_n \rangle h_n \right) = \sum_{n \geq 1} \langle P_1 \lambda_s, h_n \rangle \mu_n h_n,$$

we have  $\|T_W P_1 \lambda_s\|_2^2 = \sum_{n \geq 1} |\langle P_1 \lambda_s, h_n \rangle|^2 |\mu_n|^2 \leq \tilde{r}(T_W)^2 \|P_1 \lambda(s, \cdot)\|_2^2$  so that

$$\|P_1 \lambda_t\|_2 \leq \|P_1 v_t\|_2 + \tilde{r}(T_W) \int_0^t h(t-s) \|P_1 \lambda_s\|_2 ds. \quad (3.10.4)$$

If we define  $\alpha(t) = \|P_1 \lambda_t\|_2$ , we see that  $\alpha$  satisfies the convolution inequality

$$\alpha(t) \leq \|P_1 v_t\|_2 + \tilde{r}(T_W) \int_0^t h(t-s) \alpha(s) ds,$$

hence we can compare it to  $\beta_{\tilde{r}}(t)$  solution of the convolution equality

$$\beta(t) = \|P_1 u_0(t, \cdot)\|_2 + 1 + \tilde{r}(T_W) \int_0^t h(t-s)\beta(s)ds$$

with Lemma 3.A.4:  $\alpha(t) \leq \beta(t)$  for all  $t \geq 0$ , that is  $\|P_1 \lambda_t\|_2 \leq \beta_{\tilde{r}}(t)$  for all  $t \geq 0$ . We want now to show that  $\beta_{\tilde{r}}(t) = o(e^{\sigma_{\tilde{r}} t})$  when  $t \rightarrow \infty$ . First suppose that we are in the case  $\|h\|_1 \tilde{r}(T_W) > 1$ . We apply (as done for  $P_0$ ) Theorem 3.A.9 and obtain  $\beta_{\tilde{r}}(t) \sim_{t \rightarrow \infty} \tilde{C} e^{\sigma_{\tilde{r}} t}$  where  $\tilde{C} > 0$  depends on the parameter functions and  $\sigma_{\tilde{r}} > 0$  verifies  $\tilde{r}(T_W) \int_0^\infty e^{-\sigma_{\tilde{r}} t} h(t) dt = 1$ . In this case,  $\sigma_{\tilde{r}} < \sigma_r$  as  $\tilde{r}(T_W) < r_\infty$ , and  $\beta_{\tilde{r}}(t) = o(e^{\sigma_r t})$  follows. Suppose now that we are in the case  $\|h\|_1 \tilde{r}(T_W) \leq 1$ . As  $\|h\|_1 r_\infty > 1$ , we can find  $\bar{r}$  such that  $\tilde{r}(T_W) < \bar{r} < r_\infty$  and  $\|h\|_1 \bar{r} > 1$ . Then, considering  $\delta$  satisfying  $\delta_{\bar{r}}(t) = \|P_1 v_t\|_2 + 2 + \bar{r} \int_0^t h(t-s)\delta_{\bar{r}}(s)ds$ , as done before Lemma 3.A.4 gives  $\beta_{\tilde{r}}(t) \leq \delta_{\bar{r}}(t)$  and Theorem 3.A.9 gives  $\delta_{\bar{r}}(t) \sim_{t \rightarrow \infty} \bar{C} e^{\sigma_{\bar{r}} t}$  where  $\bar{C} > 0$  depends on the parameter functions and  $\sigma_{\bar{r}} > 0$  verifies  $\bar{r} \int_0^\infty e^{-\sigma_{\bar{r}} t} h(t) dt = 1$ . We have then that  $\beta_{\tilde{r}}(t) \leq \delta_{\bar{r}}(t) \sim_{t \rightarrow \infty} \bar{C} e^{\sigma_{\bar{r}} t} = o(e^{\sigma_r t})$ . In any case, we obtain  $\|P_1 \lambda_t\|_2 = o(e^{\sigma_r t})$ , and as Parseval equality gives

$$\|\lambda_t\|_2^2 = \|P_0 \lambda_t\|_2^2 + \|P_1 \lambda_t\|_2^2,$$

it implies that  $\|\lambda_t\|_2 \sim_{t \rightarrow \infty} C e^{\sigma_r t} \xrightarrow[t \rightarrow \infty]{} +\infty$ , with  $C$  a positive constant, hence the result.

*Case  $k > 1$ .* We deal with  $k = 2$  and leave the generalisation to the reader. Hypothesis 3.4.6 (3.4.14) is then that the kernel of  $T_W^2$  is positive. As  $\lambda_t(x) = v_t(x) + \int_0^t h(t-s)T_W \lambda_s(x)ds$ , we have  $T_W \lambda_t(x) = T_W v_t(x) + \int_0^t h(t-s)T_W^2 \lambda_s(x)ds$  and

$$\begin{aligned} \lambda_t(x) &= v_t(x) + \int_0^t h(t-s)T_W v_s(x)ds + \int_0^t h(t-s) \int_0^s h(s-u)T_W^2 \lambda_u(x)duds \\ &= \tilde{v}_t(x) + \int_0^t \tilde{h}(t-s)T_W^2 \lambda_s(x)ds. \end{aligned} \quad (3.10.5)$$

with  $\tilde{h} = h * h$  and  $\tilde{v}_t(x) = v_t(x) + \int_0^t h(t-s)T_W v_s(x)ds$ . As  $\|\tilde{h}\|_1 = \int_0^\infty \int_0^t h(t-s)h(s)dsdt = \int_0^\infty h(s) \int_s^\infty h(t-s)dt ds = \|h\|_1^2$  and  $\sqrt{r(T_W^2)} = r(T_W)$ , the condition (3.4.13) implies  $\|\tilde{h}\|_1 r(T_W^2) = \|h\|_1^2 r(T_W)^2 > 1$ . Then, we can apply the previous case ( $k = 1$ ) on  $\lambda$  satisfying (3.10.5).  $\square$

## 3.A Useful results

We gather here useful results that we use in the paper.

### 3.A.1 Various technical lemmas

The proof of the Lemmas 3.A.1 and 3.A.3 can be found respectively in Lemmas 22 and 23 of Delattre *et al.* (2016a).

**Lemma 3.A.1.** *Let  $\phi : [0, \infty[ \rightarrow \mathbb{R}$  be locally integrable and  $\alpha : [0, \infty[ \rightarrow \mathbb{R}$  with finite variations on compact intervals such that  $\alpha(0) = 0$ . Then for all  $t \geq 0$ , we have*

$$\int_0^t \int_0^{s^-} \phi(s-u) d\alpha(u) ds = \int_0^t \int_0^s \phi(s-u) d\alpha(u) ds = \int_0^t \phi(t-s) \alpha(s) ds.$$

**Lemma 3.A.2.** *Let  $(X_i)_{1 \leq i \leq N}$  be a family of  $N$  independent random variables. Then*

$$\mathbf{E} \left[ \left| \frac{1}{N} \sum_{i=1}^N (X_i - \mathbf{E}[X_i]) \right| \right] \leq \frac{1}{N} \sqrt{\sum_{i=1}^N \text{Var}(X_i)}.$$

*Proof.* We set  $Y := \frac{1}{N} \sum_{i=1}^N X_i$ , then  $\mathbf{E}[Y] = \frac{1}{N} \sum_{i=1}^N \mathbf{E}[X_i]$  and  $\text{Var}(Y) = \frac{1}{N^2} \sum_{i=1}^N \text{Var}(X_i)$  by independence. We have using Jensen's inequality that

$$\mathbf{E} \left[ \left| \frac{1}{N} \sum_{i=1}^N (X_i - \mathbf{E}[X_i]) \right| \right] = \mathbf{E} [|Y - \mathbf{E}[Y]|] \leq \sqrt{\mathbf{E} [(Y - \mathbf{E}[Y])^2]} = \sqrt{\text{Var}(Y)},$$

and the result follows with the expression of  $\text{Var}(Y)$ .  $\square$

**Lemma 3.A.3.** *Let  $\phi : [0, \infty[ \rightarrow [0, \infty[$  be a locally integrable function and  $g : [0, \infty[ \rightarrow [0, \infty[$  a locally bounded function.*

*i) Let  $u$  be a locally bounded nonnegative function such that for all  $t \geq 0$ :  $u(t) \leq g(t) + \int_0^t \phi(t-s)u(s)ds$ . Then for all  $T \geq 0$  there exists  $C_T$  (depending on  $T$  and  $\phi$ ) verifying  $\sup_{[0, T]} u(t) \leq C_T \sup_{[0, T]} g(t)$ .*

*ii) Let  $(u_n)$  be a sequence of locally bounded nonnegative functions such that for all  $t \geq 0$  and  $n \geq 0$ :  $u_{n+1}(t) \leq \int_0^t \phi(t-s)u_n(s)ds$ . Then for all  $T \geq 0$  there exists  $C_T$  (depending on  $T$ ,  $\phi$  and  $u_0$ ) verifying  $\sup_{[0, T]} \sum_{n \geq 0} u_n(t) \leq C_T$ .*

*iii) Let  $(u_n)$  be a sequence of locally bounded nonnegative functions such that for all  $t \geq 0$  and  $n \geq 0$ :  $u_{n+1}(t) \leq g(t) + \int_0^t \phi(t-s)u_n(s)ds$ . Then for all  $T \geq 0$  there exists  $C_T$  (depending on  $T$ ,  $\phi$ ,  $u_0$  and  $g$ ) verifying  $\sup_{[0, T]} \sup_{n \geq 0} u_n(t) \leq C_T$ .*

**Lemma 3.A.4.** *Let  $r > 0$ ,  $h$  be a nonnegative locally integrable function,  $u$ ,  $\alpha$  and  $\beta$  be locally bounded nonnegative continuous functions such that for all  $t \geq 0$ :*

$$\begin{aligned} \alpha(t) &\leq u(t) + r \int_0^t h(t-s)\alpha(s)ds, \\ \beta(t) &= u(t) + 1 + r \int_0^t h(t-s)\beta(s)ds. \end{aligned}$$

*Then  $\alpha(t) \leq \beta(t)$  for all  $t \geq 0$ .*

*Proof.* Let  $t^* = \inf \{s > 0, \alpha(s) > \beta(s)\}$ . Note that  $t^* > 0$  as  $\beta(0) - \alpha(0) \geq 1$ . Suppose  $t^* < \infty$ , then  $\beta(t^*) - \alpha(t^*) \geq 1 + \int_0^{t^*} h(t-s)(\beta(s) - \alpha(s)) ds \geq 1$  which is impossible, then necessarily  $t^* = +\infty$  and  $\alpha(t) \leq \beta(t)$  for all  $t \geq 0$ .  $\square$

**Lemma 3.A.5.** *Let  $u$  and  $h$  be locally square integrable functions,  $u$  nonnegative,  $T > 0$  and  $\alpha, \beta$  two constants. Assume that for any  $t \in [0, T]$ ,  $u(t) \leq \alpha \int_0^t h(t-s)u(s)ds + \beta$ . Then  $u$  satisfies the following Grönwall's inequality:*

$$u(T) \leq \sqrt{2}\beta \exp(\alpha^2 \|h\|_{T,2}^2 T).$$

*Proof.* Using Cauchy-Schwarz inequality,  $u(t)^2 \leq 2\alpha^2 \|h\|_{T,2}^2 \int_0^t u(s)^2 ds + 2\beta^2$ . We conclude by applying standard Grönwall lemma to  $u^2$  and taking the square root (since  $u \geq 0$ ).  $\square$

**Lemma 3.A.6.** *Fix  $N > 1$  and  $(Y_l)_{l=1,\dots,n}$  real valued random variables defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . Suppose that there exists  $\nu > 0$  such that, almost surely, for all  $l = 1, \dots, n-1$ ,  $Y_l \leq 1$ ,  $\mathbb{E}[Y_{l+1} | Y_l] = 0$  and  $\mathbb{E}[Y_{l+1}^2 | Y_l] \leq \nu$ . Then  $\mathbb{P}(n^{-1}(Y_1 + \dots + Y_n) \geq x) \leq \exp\left(-n \frac{x^2}{2\nu} B\left(\frac{x}{\nu}\right)\right)$  for all  $x \geq 0$ , where*

$$B(u) := u^{-2}((1+u)\log(1+u) - u). \quad (3.A.1)$$

*Proof.* A direct application of (Dembo & Zeitouni, 1998, Corollary 2.4.7) gives that

$$\mathbb{P}(n^{-1}(Y_1 + \dots + Y_n) \geq x) \leq \exp\left(-nH\left(\frac{x+\nu}{1+\nu} \middle| \frac{\nu}{1+\nu}\right)\right),$$

where  $H(p|q) := p \log(p/q) + (1-p) \log((1-p)/(1-q))$  for  $p, q \in [0, 1]$ . Then, the inequality  $H\left(\frac{x+\nu}{1+\nu} \middle| \frac{\nu}{1+\nu}\right) \geq \frac{x^2}{2\nu} B\left(\frac{x}{\nu}\right)$  (see (Dembo & Zeitouni, 1998, Exercise 2.4.21)) gives the result.  $\square$

**Lemma 3.A.7.** *Fix  $N \geq 1$ ,  $(p_1, \dots, p_N)$  in  $[0, 1]$  and a sequence  $(d_1, \dots, d_N)$  such that  $|d_l| \leq 1$  for any  $l \in \llbracket 1, N \rrbracket$ . Suppose that there exists  $\kappa_N > 0$  and  $w_N \in ]0, 1]$  such that  $p_l \leq w_N$  for any  $l \in \llbracket 1, N \rrbracket$ . Then, setting  $\varepsilon_n := 32 \frac{\kappa_N^2 w_N}{N} \log(N)$  for  $(U_1, \dots, U_N)$  independent random variables with  $U_l \sim \mathcal{B}(p_l)$ , we have*

$$\mathbb{P}\left(\left|\frac{\kappa_N}{N} \sum_{l=1}^N (U_l - p_l) d_l\right| > \varepsilon_N\right) \leq 2 \exp\left(-16 \log(N) B\left(4\sqrt{2} \left(\frac{\log(N)}{N w_N}\right)^{\frac{1}{2}}\right)\right) \quad (3.A.2)$$

with  $B$  defined in (3.A.1).

*Proof.* This is a simple corollary of Lemma 3.A.6 applied to  $Y_l := (U_l - p_l)d_l$ .  $\square$

### 3.A.2 Spectral theory

We include here advanced results of spectral analysis that are used in the Chapter.

### 3.A.3 Jentzsch/Krein–Rutman Theorem

The following theorem can be found in (Schaefer, 1974, Theorem 6.6) or in (Zerner, 1987, Theorem 1).

**Theorem 3.A.8.** *Let  $E := L^p(\mu)$ , where  $1 \leq p \leq +\infty$  and  $(X, \Sigma, \mu)$  is a  $\sigma$ -finite measure space. Suppose  $T \in \mathcal{L}(E)$  is an operator given by a  $(\Sigma \times \Sigma)$ -measurable kernel  $K \geq 0$ , satisfying these two assumptions:*

- i) Some power of  $T$  is compact.*
- ii)  $S \in \Sigma$  and  $\mu(S) > 0, \mu(X \setminus S) > 0$  implies*

$$\int_{X \setminus S} \int_S K(s, t) d\mu(s) d\mu(t) > 0.$$

*Then  $r(T) > 0$  is an eigenvalue of  $T$  with a unique normalized eigenfunction  $f$  satisfying  $f(s) > 0$   $\mu$  a.e. Moreover, if  $K(s, t) > 0$   $\mu \otimes \mu$  a.e. then every other eigenvalue  $\lambda$  of  $T$  has modulus  $|\lambda| < r(T)$ .*

### 3.A.4 Renewal theory

The following theorem can be found in (Feller, 1941, Theorem 4). This article studies the behavior of solutions of the integral equation

$$u(t) = g(t) + \int_0^t u(t-x)f(x)dx, \tag{3.A.3}$$

where  $f$  and  $g$  are measurable, nonnegative and bounded in every finite interval  $[0, T]$ .

**Theorem 3.A.9.** *Suppose  $\int_0^\infty f(t)dt > 1$ ,  $\int_0^\infty g(t)dt = b < \infty$ . Suppose moreover that there exists an integer  $n \geq 2$  such that the moments  $m_k = \int_0^\infty t^k f(t)dt$ ,  $k = 1, 2, \dots, n$ , are finite and that the functions  $f(t)$ ,  $tf(t)$ ,  $t^2f(t)$ ,  $\dots$ ,  $t^{n-2}f(t)$  are of bounded total variation over  $(0, \infty)$ . Suppose finally that*

$$\lim_{t \rightarrow \infty} t^{n-2}g(t) = 0 \text{ and } \lim_{t \rightarrow \infty} t^{n-2} \int_t^\infty g(x)dx = 0.$$

*Then there is a unique  $\alpha_0 > 0$  such that  $\mathcal{L}_f(\alpha_0) = 1$  and for some constant  $C$  depending on  $(\mu, f, g)$ :*

$$u(t) \sim_{t \rightarrow \infty} Ce^{\alpha_0 t}, \text{ where } u \text{ solves (3.A.3).}$$

## 3.B About matrices, graphs and graphon

Here we give a small introduction to the illustration of graphs by graphon, that gathers results from Lovász (2012).



For a system of  $N$  particles, we set  $\xi^{(N)}$  a  $N \times N$  matrix representing the graph of their interaction, with values in  $\{0, 1\}$ : if the edge from the  $j$ -th particle to the  $i$ -th exists, then  $\xi_{ij}^{(N)} = 1$  and 0 otherwise.

We can consider different norms on matrices.

**Definition 3.B.1.** Let  $A = (a_{ij})$  be a  $N \times N$  matrix. Its cut-norm is defined as

$$\|A\|_{\square} := \frac{1}{N^2} \max_{S, T \subset \{1, \dots, N\}} \left| \sum_{i \in S, j \in T} a_{ij} \right|, \quad (3.B.1)$$

and we define also

$$\|A\|_{\infty \rightarrow 1} := \sup_{s_i, t_j \in \{\pm 1\}} \sum_{i, j=1}^N a_{ij} s_i t_j. \quad (3.B.2)$$

**Lemma 3.B.2.** *The norms  $\|\cdot\|_{\square}$  and  $\|\cdot\|_{\infty \rightarrow 1}$  are equivalent: if  $A$  is a  $N \times N$  matrix, then*

$$N^2 \|A\|_{\square} \leq \|A\|_{\infty \rightarrow 1} \leq 4N^2 \|A\|_{\square}. \quad (3.B.3)$$

*Proof from Alon & Naor (2006).* Let  $A = \{a_{ij}\}_{1 \leq i, j \leq N}$  and  $x_i, y_i$  in  $\{\pm 1\}$  for  $1 \leq i \leq N$ . Denote by  $I_+ = \{i, x_i = 1\}$ ,  $I_- = \{i, x_i = -1\}$ ,  $J_+ = \{j, y_j = 1\}$  and  $J_- = \{j, y_j = -1\}$ , we have

$$\begin{aligned} \sum_{i, j=1}^N a_{ij} x_i y_j &= \sum_{i \in I_+, j \in J_+} a_{ij} + \sum_{i \in I_-, j \in J_-} a_{ij} - \sum_{i \in I_-, j \in J_+} a_{ij} - \sum_{i \in I_+, j \in J_-} a_{ij} \\ &\leq \left| \sum_{i \in I_+, j \in J_+} a_{ij} \right| + \left| \sum_{i \in I_-, j \in J_-} a_{ij} \right| + \left| \sum_{i \in I_-, j \in J_+} a_{ij} \right| + \left| \sum_{i \in I_+, j \in J_-} a_{ij} \right| \\ &\leq 4N^2 \|A\|_{\square}. \end{aligned}$$

Taking the sup on  $\{x_i, y_j \in \{\pm 1\}\}$ , we obtain

$$\|A\|_{\infty \rightarrow 1} \leq 4N^2 \|A\|_{\square}.$$

To show the other inequality, suppose that  $N^2 \|A\|_{\square} = \sum_{i \in S, j \in T} a_{ij}$  (or consider  $-A$  if  $N^2 \|A\|_{\square} = -\sum_{i \in S, j \in T} a_{ij}$ ). Then, for every  $i$  and  $j$  in  $\{1, \dots, N\}$ , let  $x_i = \mathbf{1}_{i \in S} - \mathbf{1}_{i \notin S}$  and  $y_j = \mathbf{1}_{j \in T} - \mathbf{1}_{j \notin T}$ . We then have  $\left(\frac{1+x_i}{2}\right) \left(\frac{1+y_j}{2}\right) = \mathbf{1}_{i \in S, j \in T}$  and can write

$$\begin{aligned} N^2 \|A\|_{\square} &= \sum_{i \in S, j \in T} a_{ij} = \sum_{i, j=1}^N a_{ij} \left(\frac{1+x_i}{2}\right) \left(\frac{1+y_j}{2}\right) \\ &= \frac{1}{4} \sum_{i, j=1}^N a_{ij} + \frac{1}{4} \sum_{i, j=1}^N a_{ij} y_j + \frac{1}{4} \sum_{i, j=1}^N a_{ij} x_i + \frac{1}{4} \sum_{i, j=1}^N a_{ij} x_i y_j. \end{aligned}$$

Taking the sup on  $\{x_i, y_j \in \{\pm 1\}\}$  on each term, we obtain  $N^2 \|A\|_{\square} \leq \|A\|_{\infty \rightarrow 1}$ .  $\square$

To each  $N \times N$  matrix  $A = (a_{ij})$  with real values, we can associate a graphon function  $W_A$  defined on  $[0, 1]^2$  by, for every  $(u, v) \in [0, 1]^2$  :

$$W_A(u, v) = \sum_{i,j=1}^N a_{ij} \mathbf{1}_{\left(\frac{i-1}{N}, \frac{i}{N}\right] \times \left(\frac{j-1}{N}, \frac{j}{N}\right]}(u, v).$$

Note that we have then

$$\frac{1}{N^2} \sum_{i,j=1}^N a_{ij} = \int_0^1 \int_0^1 W_A(u, v) du dv.$$

This association between a graph of  $N$  particles and a graphon allows us to represent a graph (with ordered vertices) by a function, and to study sequences of graphs eventually defined on different numbers of particles. Particularly, we can link the study of a sequence  $(A_N)$  of graphs such that  $A_N$  is a graph on  $N$  particles to the study of functions  $W_{A_N}$  defined on  $[0, 1]^2$ . We can see in Figure 2.6 an illustration to this convergence.

In an analogous manner as Definition 3.B.1, we can define the following norms.

**Definition 3.B.3.** Let  $W$  be a function defined on  $[0, 1]^2$  with values in  $\mathbb{R}$ . We define

$$\|W\|_{\square} := \sup_{S, T \subset [0, 1]^2} \left| \int_{S \times T} W(x, y) dx dy \right|$$

and

$$\begin{aligned} \|W\|_{\infty \rightarrow 1} &:= \sup_{\|g\|_{\infty} \leq 1} \int_{[0, 1]} \left| \int_{[0, 1]} W(x, y) g(y) dy \right| dx \\ &= \sup_{\|f\|_{\infty}, \|g\|_{\infty} \leq 1} \left| \int \int W(x, y) f(x) g(y) dx dy \right|. \end{aligned}$$

**Notation 3.B.4.** We define the cut-distance between two functions by

$$d_{\square}(A, B) = \|A - B\|_{\square}.$$

**Lemma 3.B.5.**  $\|\cdot\|_{\square}$  and  $\|\cdot\|_{\infty \rightarrow 1}$  are equivalent: if  $W$  is a function defined on  $[0, 1]^2$  with values in  $\mathbb{R}$ , then

$$\|W\|_{\square} \leq \|W\|_{\infty \rightarrow 1} \leq 4\|W\|_{\square}.$$

The proof can be find in (Lovász, 2012, Lemma 8.11).

*Remark.* We can also extend these norms and use them not on  $[0, 1]^2$  but on any measure space  $I$  with probability measure  $\nu$  by defining

$$\|W\|_{\square, \nu} := \sup_{S, T \subset I} \left| \int_{S \times T} W(x, y) \nu(dx) \nu(dy) \right|$$

and

$$\|W\|_{\infty \rightarrow 1, \nu} := \sup_{\|g\|_{\infty} \leq 1} \int_I \left| \int_I W(x, y) g(y) \nu(dy) \right| \nu(dx).$$



# Chapter 4

## Long-term stability of interacting Hawkes processes on random graphs

*This chapter consists in a modified version of my article [Agathe-Nerine \(2023a\)](#) (expanded with some computations), published in *Electronic Journal of Probability*, with minor wording and notation changes for harmonization purposes.*

In this Chapter, we study our main model (2.3.7), but we give a more general writing as we consider

$$\lambda_{N,i}(t) = F(X_{N,i}(t-), \eta_i(x_i)),$$

where  $\eta$  represents deterministic contributions to the intensity and  $X_{N,i}$  is given in (2.3.8). This choice was made as we focus here more on the stochastic contribution  $X_{N,i}$  in the intensity, on its asymptotic behavior. It is straightforward to go back to (2.3.7) by considering  $\eta = (\mu, v)$  and  $F(X, \eta) = \mu + f(v + X)$ . See Remark 4.1.3 below for more details.

We also restrict to the case  $I = [0, 1]$  and the neurons are evenly located on  $I$ , that is for  $1 \leq i \leq N$ ,  $x_i = \frac{i}{N}$ . The interaction term is chosen as  $w_{ij}^{(N)} = \frac{\xi_{ij}^{(N)}}{\rho_N}$ , where  $\rho_N$  is a dilution parameter and  $\xi_{ij}^{(N)}$  follows a Bernoulli distribution with parameter  $\rho_N W(x_i, x_j)$ ,  $W$  being a macroscopic interaction kernel.

### Abstract

We consider a population of Hawkes processes modeling the activity of  $N$  interacting neurons. The neurons are regularly positioned on the segment  $[0, 1]$ , and the connectivity between neurons is given by a random possibly diluted and inhomogeneous graph where the probability of presence of each edge depends on the spatial position of its vertices through a spatial kernel. The main result of the paper concerns the long-time stability of the synaptic current of the population, as  $N \rightarrow \infty$ , in the subcritical regime in case the synaptic memory kernel is exponential, up to time horizons that are polynomial in  $N$ .

## Contents

---

4.1	Introduction . . . . .	<b>124</b>
4.1.1	Hawkes processes in neuroscience . . . . .	124
4.1.2	Notation . . . . .	127
4.1.3	The model . . . . .	127
4.2	Hypotheses and main results . . . . .	<b>130</b>
4.2.1	Hypotheses . . . . .	130
4.2.2	Main results . . . . .	132
4.2.3	Examples and extensions . . . . .	134
4.2.4	Link with the literature . . . . .	138
4.2.5	Strategy of proof and organization of the paper . . . . .	140
4.3	Asymptotic behavior of $(X_t)$ . . . . .	<b>141</b>
4.3.1	Estimates on the operator $\mathcal{L}$ . . . . .	141
4.3.2	About the large time behavior of $X_t$ . . . . .	142
4.4	Large time behavior of $X_N(t)$ . . . . .	<b>144</b>
4.4.1	Mild formulation . . . . .	144
4.4.2	Proof of the large time behaviour . . . . .	146
4.5	Noise perturbation (proofs) . . . . .	<b>149</b>
4.5.1	Control of the moments of the process $Z_{N,i}$ . . . . .	150
4.5.2	Proof of Proposition 4.4.2 . . . . .	151
4.6	Drift term (proofs) . . . . .	<b>155</b>
4.6.1	Notation . . . . .	156
4.6.2	Preliminary results . . . . .	156
4.6.3	Proof of Proposition 4.4.3 . . . . .	165
4.7	About the finite time behavior (proofs) . . . . .	<b>165</b>
4.7.1	Main technical results . . . . .	165
4.7.2	Proofs of Propositions 4.7.1 and 4.7.2 . . . . .	166
4.A	Auxiliary results . . . . .	<b>168</b>
4.A.1	Concentration results . . . . .	168
4.A.2	Other technical results . . . . .	170

---

## 4.1 Introduction

### 4.1.1 Hawkes processes in neuroscience

In the present paper we study the large time behavior of a population of interacting and spiking neurons, as the size of the population  $N$  tends to infinity. We model

the activity of a neuron by a point process where each point represents the time of a spike:  $Z_{N,i}(t)$  counts the number of spikes during the time interval  $[0, t]$  of the  $i$ th neuron of the population. Its intensity at time  $t$  conditioned on the past  $[0, t)$  is given by  $\lambda_{N,i}(t)$ , in the sense that

$$\mathbf{P}(Z_{N,i} \text{ jumps between } (t, t+dt) | \mathcal{F}_t) = \lambda_{N,i}(t)dt,$$

where  $\mathcal{F}_t := \sigma(Z_{N,i}(s), s \leq t, 1 \leq i \leq N)$ .

For the choice of  $\lambda_{N,i}$ , we want to account for the dependence of the activity of a neuron on the past of the whole population : the spike of one neuron can trigger others' spikes. *Hawkes processes* are then a natural choice to emphasize this interdependency. A generic choice is

$$\lambda_{N,i}(t) = \mu_t(x_i) + f \left( v_t(x_i) + \frac{1}{N} \sum_{j=1}^N w_{ij}^{(N)} \int_0^{t-} h(t-s) dZ_{N,j}(s) \right). \quad (4.1.1)$$

Here, with the  $i$ th neuron at position  $x_i = \frac{i}{N} \in I := [0, 1]$ ,  $f : \mathbb{R} \rightarrow \mathbb{R}_+$  represents the (possible) non linear synaptic integration,  $\mu_t : I \rightarrow \mathbb{R}_+$  a spontaneous activity of the neuron at time  $t$ ,  $v_t : I \rightarrow \mathbb{R}$  a past activity and  $h : \mathbb{R}_+ \rightarrow \mathbb{R}$  a memory function which models how a past jump of the system affects the present intensity. The term  $w_{ij}^{(N)}$  represents the random inhomogeneous interaction between neurons  $i$  and  $j$ , that will be modeled here in terms of the realization of a random graph.

Since the seminal works of [Hawkes \(1971\)](#); [Hawkes & Oakes \(1974\)](#), there has been a renewed interest in the use of Hawkes processes, especially in neuroscience. A common simplified framework is to consider an interaction on the complete graph, that is taking  $w_{ij}^{(N)} = 1$  in (4.1.1), as done in [Delattre et al. \(2016a\)](#). In this case, a very simple instance of (4.1.1) concerns the so called *linear case*, when  $f(x) = x$ ,  $\mu_t(x) = \mu$  and  $v = 0$ , that is  $\lambda_{N,i}(t) = \lambda_N(t) = \mu + \frac{1}{N} \sum_{j=1}^N \int_0^{t-} h(t-s) dZ_{N,j}(s)$ , with  $h \geq 0$  (see [Delattre et al. \(2016a\)](#)). The biological evidence [Bosking et al. \(1997\)](#); [Mountcastle \(1997\)](#) of a spatial organisation of neurons in the brain has led to more elaborate Hawkes models with spatial interaction, possibly including inhibition (see [Touboul \(2014\)](#); [Ditlevsen & Löcherbach \(2017\)](#); [Chevallier et al. \(2019\)](#)). This would correspond in (4.1.1) to take  $w_{ij}^{(N)} = W(x_i, x_j)$ , where  $W$  is a macroscopic interaction kernel, usual examples being the exponential distribution on  $\mathbb{R}$ ,  $W(x, y) = \frac{1}{2\sigma} \exp\left(-\frac{|x-y|}{\sigma}\right)$  or the ‘‘Mexican hat’’ distribution  $W(x, y) = e^{-|x-y|} - Ae^{-\frac{|x-y|}{\sigma}}$ ,  $A \in \mathbb{R}$ ,  $\sigma > 0$ . The macroscopic limit of the multivariate Hawkes process (4.1.1) is then given by a family of spatially extended inhomogeneous Poisson processes whose intensities  $(\lambda_t(x))_{x \in I}$  solve the convolution equation

$$\lambda_t(x) = \mu_t(x) + f \left( v_t(x) + \int_I W(x, y) \int_0^t h(t-s) \lambda_s(y) ds dy \right). \quad (4.1.2)$$

A crucial example is the exponential case, that is when  $h(t) = e^{-\alpha t}$  for some  $\alpha > 0$ . In this case, the Hawkes process with intensity (4.1.1) is Markovian (see [Ditlevsen & Löcherbach \(2017\)](#)). Denoting in (4.1.2)  $u_t(x) := v_t(x) + \int_I W(x, y) \int_0^t h(t-s) \lambda_s(y) ds dy$  as the potential of a neuron (the synaptic current) localised in  $x$  at time

$t$  (so that (4.1.2) becomes  $\lambda_t(x) = f(u_t(x))$ ), an easy computation (see [Chevallier et al. \(2019\)](#)) gives that, when  $v_t(x) = e^{-\alpha t}v_0(x)$  for some  $v_0$ ,  $u$  solves the *Neural Field Equation* (NFE)

$$\frac{\partial u_t(x)}{\partial t} = -\alpha u_t(x) + \int_I W(x, y)f(u_t(y))dy + I_t(x), \quad (4.1.3)$$

with source term  $I_t(x) := \int_I W(x, y)\mu_t(y)dy$ . Equation (4.1.3) has been extensively studied in the literature, mostly from a phenomenological perspective ([Wilson & Cowan, 1972](#); [Amari, 1977](#)), and is an important example of macroscopic neural dynamics with non-local interactions (we refer to [Bressloff \(2012\)](#) for an extensive review on the subject).

In a previous work [Agathe-Nerine \(2022\)](#) (that is Chapter 3 of the present thesis), we give a microscopic interpretation of the macroscopic kernel  $W$  in terms of an inhomogeneous graph of interaction. We consider  $w_{ij}^{(N)} = \xi_{ij}^{(N)}\kappa_i$  in (4.1.1), where  $(\xi_{ij}^{(N)})_{1 \leq i, j \leq N}$  is a collection of independent Bernoulli variables, with individual parameter  $W(x_i, x_j)$ : the probability that two neurons are connected depends on their spatial positions. The term  $\kappa_i$  is a suitable local renormalisation parameter, to ensure that the interaction remains of order 1. This modeling constitutes a further difficulty in the analysis as we are no longer in a mean-field framework: contrary to the case  $w_{ij}^{(N)} = 1$ , the interaction (4.1.1) is no longer a functional of the empirical measure of the particles  $(Z_{N,1}, \dots, Z_{N,N})$ . A recent interest has been shown to similar issues in the case of diffusions interacting on random graphs (first in the homogeneous Erdős-Rényi case ([Delattre et al., 2016b](#); [Coppini et al., 2019, 2022](#); [Coppini, 2022](#)), and secondly for inhomogenous random graph ([Luçon, 2020](#); [Bet et al., 2020](#); [Bayraktar et al., 2023](#))). See also [Pfaffelhuber et al. \(2022\)](#) where the interaction is random (either excitatory or inhibitory) on the complete graph with a diffusive scaling in  $1/\sqrt{N}$  when the excitation and inhibition are balanced.

A common motivation between [Agathe-Nerine \(2022\)](#) in the case of Hawkes processes and [Luçon \(2020\)](#); [Bet et al. \(2020\)](#); [Bayraktar et al. \(2023\)](#) in the case of diffusions is to understand how the inhomogeneity of the underlying graph may or may not influence the long time dynamics of the system. An issue common to all mean-field models (and their perturbations) is that there is, in general, no possibility to interchange the limits  $N \rightarrow \infty$  and  $t \rightarrow \infty$ . More precisely, restricting to Hawkes processes, a usual propagation of chaos result (see ([Delattre et al., 2016a](#), Theorem 8), ([Chevallier et al., 2019](#), Theorem 1) and ([Agathe-Nerine, 2022](#), Theorem 3.10)) may be stated as follows: for fixed  $T > 0$ , there exists some  $C(T) > 0$  such that

$$\sup_{1 \leq i \leq N} \mathbf{E} \left( \sup_{s \in [0, T]} |Z_{N,i}(s) - \bar{Z}_i(s)| \right) \leq \frac{C(T)}{\sqrt{N}}, \quad (4.1.4)$$

where  $\bar{Z}_i$  is a Poisson process with intensity  $(\lambda_t(x_i))_{t \geq 0}$  defined in (4.1.2) suitably coupled to  $Z_{N,i}$ , see the above references for details. Generically,  $C(T)$  is of the form  $\exp(CT)$ , such that (4.1.4) remains only relevant up to  $T \sim c \log N$  with  $c$  sufficiently small. In the pure mean-field linear case ( $w_{ij}^{(N)} = 1$ ,  $f(x) = x$ ),

there is a well known phase transition (Delattre *et al.*, 2016a, Theorems 10,11) when  $\|h\|_1 = \int_0^\infty h(t)dt < 1$  (*subcritical case*),  $\lambda_t \xrightarrow[t \rightarrow \infty]{} \frac{\mu}{1 - \|h\|_1}$ , whereas when  $\|h\|_1 > 1$  (*supercritical case*),  $\lambda_t \xrightarrow[t \rightarrow \infty]{} \infty$ . This phase transition has been extended to the inhomogeneous case in Agathe-Nerine (2022). In the subcritical case, one can actually improve (4.1.4) in the sense that  $C(T)$  is now linear in  $T$  so that (4.1.4) remains relevant up to  $T = o(\sqrt{N})$ . A natural question is to ask if this approximation remains valid beyond this time scale. The purpose to the present work is to address this question: we show that, in the whole generality of (4.1.1), in the subcritical regime and exponential case (see details below), the macroscopic intensity (4.1.2) converges to a finite limit when  $t \rightarrow \infty$  and that the microscopic system remains close to this limit up to polynomial times in  $N$ .

### 4.1.2 Notation

We denote by  $C_{\text{parameters}}$  a constant  $C > 0$  which only depends on the parameters inside the lower index. These constants can change from line to line or inside a same equation, we choose just to highlight the dependency they contain. When it is not relevant, we just write  $C$ . For any  $d \geq 1$ , we denote by  $|x|$  and  $x \cdot y$  the Euclidean norm and scalar product of elements  $x, y \in \mathbb{R}^d$ . For  $(E, \mathcal{A}, \mu)$  a measured space, for a function  $g$  in  $L^p(E, \mu)$  with  $p \geq 1$ , we write  $\|g\|_{E, \mu, p} := \left(\int_E |g|^p d\mu\right)^{\frac{1}{p}}$ . When  $p = 2$ , we denote by  $\langle \cdot, \cdot \rangle$  the Hermitian scalar product in  $L^2(E)$ . Without ambiguity, we may omit the subscript  $(E, \mu)$  or  $\mu$ . For a real-valued bounded function  $g$  on a space  $E$ , we write  $\|g\|_\infty := \|g\|_{E, \infty} = \sup_{x \in E} |g(x)|$ .

For  $(E, d)$  a metric space, we denote by  $\|g\|_L = \sup_{x \neq y} |g(x) - g(y)|/d(x, y)$  the Lipschitz seminorm of a real-valued function  $g$  on  $E$ . We denote by  $\mathcal{C}(E, \mathbb{R})$  the space of continuous functions from  $E$  to  $\mathbb{R}$ , and  $\mathcal{C}_b(E, \mathbb{R})$  the space of continuous bounded ones. For any  $T > 0$ , we denote by  $\mathbb{D}([0, T], E)$  the space of càdlàg (right continuous with left limits) functions defined on  $[0, T]$  and taking values in  $E$ . For any integer  $N \geq 1$ , we denote by  $\llbracket 1, N \rrbracket$  the set  $\{1, \dots, N\}$ . For any  $p \in [0, 1]$ ,  $\mathcal{B}(p)$  denotes the Bernoulli distribution with parameter  $p$ .

### 4.1.3 The model

First, let us focus on the interaction between the particles. The graph of interaction for (4.1.1) is constructed as follows:

**Definition 4.1.1.** On a common probability space  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mathbb{P})$ , we consider a family of random variables  $\xi^{(N)} = \left(\xi_{ij}^{(N)}\right)_{N \geq 1, i, j \in \llbracket 1, N \rrbracket}$  on  $\tilde{\Omega}$  such that under  $\mathbb{P}$ , for any  $N \geq 1$  and  $i, j \in \llbracket 1, N \rrbracket$ ,  $\xi^{(N)}$  is a collection of mutually independent Bernoulli random variables such that for  $1 \leq i, j \leq N$ ,  $\xi_{ij}^{(N)}$  has parameter  $W_N(\frac{i}{N}, \frac{j}{N})$ , where

$$W_N(x, y) := \rho_N W(x, y), \tag{4.1.5}$$

with  $\rho_N$  some dilution parameter and  $W : I^2 \rightarrow [0, 1]$  a macroscopic interaction kernel. We assume that the particles in (4.1.1) are connected according to the



oriented graph  $\mathcal{G}_N = (\{1, \dots, N\}, \xi^{(N)})$ . For any  $i$  and  $j$ ,  $\xi_{ij}^{(N)} = 1$  encodes for the presence of the edge  $j \rightarrow i$  and  $\xi_{ij}^{(N)} = 0$  for its absence. The interaction in (4.1.1) is fixed as

$$w_{ij}^{(N)} = \frac{\xi_{ij}^{(N)}}{\rho_N}, \quad (4.1.6)$$

so that the interaction term remains of order 1 as  $N \rightarrow \infty$ .

The class (4.1.5) of inhomogenous graphs falls into the framework of  $W$ -random graphs, see Lovász & Szegedy (2006); Borgs *et al.* (2008, 2012). One distinguishes the **dense case** when  $\lim_{N \rightarrow \infty} \rho_N = \rho > 0$  and the **diluted case** when  $\rho_N \rightarrow 0$ .

We now fix these sequences  $(\xi^{(N)})$ , and work on a filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbf{P})$  rich enough for all the following processes can be defined. We denote by  $\mathbf{E}$  the expectation under  $\mathbf{P}$  and  $\mathbb{E}$  the expectation with respect to  $\mathbb{P}$ . In the following definitions,  $N$  is fixed and the particles are regularly located on the segment  $I = [0, 1]$ . We denote by  $x_i = \frac{i}{N}$  the position of the  $i$ -th neuron in the population of size  $N$ . We also divide  $I$  in  $N$  segments of equal length, denoted by

$$B_{N,i} := \left( \frac{i-1}{N}, \frac{i}{N} \right). \quad (4.1.7)$$

We can now formally define our process of interest.

**Definition 4.1.2.** Let  $(\pi_i(ds, dz))_{1 \leq i \leq N}$  be a sequence of  $(\mathcal{F}_t)$ -adapted i.i.d. Poisson random measures on  $\mathbb{R}_+ \times \mathbb{R}_+$  with intensity measure  $dsdz$ . The multivariate counting process  $(Z_{N,1}(t), \dots, Z_{N,N}(t))_{t \geq 0}$  defined by, for all  $t \geq 0$  and  $i \in \llbracket 1, N \rrbracket$ :

$$Z_{N,i}(t) = \int_0^t \int_0^\infty \mathbf{1}_{\{z \leq \lambda_{N,i}(s)\}} \pi_i(ds, dz) \quad (4.1.8)$$

where

$$\lambda_{N,i}(t) = F(X_{N,i}(t-), \eta_t(x_i)), \quad (4.1.9)$$

and

$$X_{N,i}(t) = \sum_{j=1}^N \frac{w_{ij}^{(N)}}{N} \int_0^t h(t-s) dZ_{N,j}(s), \quad (4.1.10)$$

$\eta_t : I \rightarrow \mathbb{R}^d$  for any  $t \in [0, +\infty)$  for some  $d \geq 1$  and  $F : \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}^+$  is called a *multivariate Hawkes process* with the set of parameters  $(N, F, \xi^{(N)}, W_N, \eta, h)$ .

Our main focus is to study the quantity  $(X_{N,i})_{1 \leq i \leq N}$  defined in (4.1.10) as  $N \rightarrow \infty$ , and more precisely the random profile defined for all  $x \in I$  by:

$$X_N(t)(x) := \sum_{i=1}^N X_{N,i}(t) \mathbf{1}_{x \in B_{N,i}}, \quad (4.1.11)$$

where  $B_{N,i}$  is defined in (4.1.7).

As  $N \rightarrow \infty$ , an informal Law of Large Numbers (LLN) argument shows that the empirical mean in (4.1.10) becomes an expectation w.r.t. the candidate limit for  $Z_{N,i}$ : we can replace the sum in (4.1.10) by the integral, the microscopic interaction

term  $w_{ij}^{(N)}$  in (4.1.10) by the macroscopic term  $W(x, y)$  (where  $y$  describes the macroscopic distribution of the positions), and the past activity of the neuron  $dZ_{N,j}(s)$  by its intensity in large population. In other words, the macroscopic spatial profile will be described by

$$X_t(x) = \int_I W(x, y) \int_0^t h(t-s) \lambda_s(y) ds dy, \quad (4.1.12)$$

where the macroscopic intensity of a neuron at position  $x \in I$  denoted by  $\lambda_t(x) = F(X_t(x), \eta_t(x))$  solves

$$\lambda_t(x) = F \left( \int_I W(x, y) \int_0^t h(t-s) \lambda_s(y) ds dy, \eta_t(x) \right). \quad (4.1.13)$$

Such informal law of large number on a bounded time interval has been made rigorous under various settings, we refer for further references to [Delattre \*et al.\* \(2016a\)](#); [Chevallier \*et al.\* \(2019\)](#) and more especially to [Agathe-Nerine \(2022\)](#) which exactly incorporates the present hypotheses.

**Remark 4.1.3.** In the expression (4.1.9) of the intensity  $\lambda_{N,i}$ ,  $X_{N,i}$  given in (4.1.10) accounts for the stochastic influence of the other interacting neurons, whereas  $\eta_t$  represents the deterministic part of the intensity  $\lambda_{N,i}$ . Having in mind the generic example given in (4.1.1), a typical choice would correspond to taking  $d = 2$  with  $\eta := (\mu, v)$  and

$$F(X, \eta) = F(X, \mu, v) = \mu + f(v + X). \quad (4.1.14)$$

Once again,  $\mu$  here corresponds to the spontaneous Poisson activity of the neuron and one may see  $v$  as a deterministic part in the evolution of the membrane potential of neuron  $i$ . Note that we generalize here slightly the framework considered in [Chevallier \*et al.\* \(2019\)](#) in the sense that [Chevallier \*et al.\* \(2019\)](#) considered (4.1.14) for  $\mu \equiv 0$  and  $v_t(x) = e^{-\alpha t} v_0(x)$  for some initial membrane potential  $v_0(x)$ . In the case of (4.1.14), one retrieves the expression of the macroscopic intensity  $\lambda_t(x)$  given in (4.1.2). Typical choices of  $f$  in (4.1.14) are  $f(x) = x$  (the so-called linear model) or some sigmoid function. Note that there will be an intrinsic mathematical difficulty in dealing with the linear case in this paper, as  $f$  is not bounded in this case. As already mentioned in the introduction, for the choice of  $h(t) = e^{-\alpha t}$  and  $v_t(x) = e^{-\alpha t} v_0(x)$ , a straightforward calculation shows that  $u_t(x) := v_t(x) + X_t(x)$  solves the scalar neural field equation (4.1.3) with source term  $I_t(x) = \int_I W(x, y) \mu(t, y) dy$ .

We choose here to work with the generic expression (4.1.9) instead of (4.1.1) not only for conciseness of notation, but also to emphasize that the result does not intrinsically depend on the specific form of the function  $F$ .

**Remark 4.1.4.** We have assumed for simplicity in the current definition (4.1.10) of  $X_{N,i}(t)$  that  $X_{N,i}(0) = 0$ . Define more generally, for any  $(\vartheta_i)_{i=1, \dots, N}$ ,  $\vartheta_i \in \mathbb{R}$ ,  $X_{N,i}^{\vartheta_i}(t) = \vartheta_i + \sum_{j=1}^N \frac{w_{ij}^{(N)}}{N} \int_0^t h(t-s) dZ_{N,j}(s)$  the same process starting at  $\vartheta_i$  (here,  $\vartheta_i$  accounts for the history of the process before  $t = 0$ ). Write then the corresponding intensity (4.1.9) and process (4.1.8) as  $\lambda_{N,i}^{\vartheta_i}(t)$  and  $Z_{N,i}^{\vartheta_i}(t)$  respectively. In particular, when  $h$  is exponential (see (4.2.4)), the process  $(X_{N,i}^{\vartheta_i})_{i=1, \dots, N}$  is Markovian (see

e.g. (Ditlevsen & Löcherbach, 2017, Section 5)). The analysis of the profile  $X_N^\vartheta(t) := \sum_{i=1}^N X_{N,i}^{\vartheta_i}(t) \mathbf{1}_{B_{N,i}}$  remains the same, under the additional hypothesis that  $\sum_{i=1}^N \vartheta_i \mathbf{1}_{B_{N,i}} \xrightarrow{N \rightarrow \infty} \vartheta$  in  $L^2(I)$ . In the following, the actual dependence in the initial condition  $\vartheta$  will be dropped, whenever it is clear from the context, for simplicity of notation.

## 4.2 Hypotheses and main results

### 4.2.1 Hypotheses

**Hypothesis 4.2.1.** We assume that

- $F$  is Lipschitz continuous : there exists  $\|F\|_L$  such that for any  $x, x' \in \mathbb{R}$ ,  $\eta, \eta' \in \mathbb{R}^d$ , we have  $|F(x, \eta) - F(x', \eta')| \leq \|F\|_L (|x - x'| + |\eta - \eta'|)$ .
- $F$  is non-decreasing in the first variable, that is for any  $\eta \in \mathbb{R}^d$ , for any  $x, x' \in \mathbb{R}$  such that  $x \leq x'$ , one has  $F(x, \eta) \leq F(x', \eta)$ . Moreover, we assume that  $F$  is  $\mathcal{C}^2$  on  $\mathbb{R}^{d+1}$  with bounded derivatives. We denote by  $\partial_X F$  and  $\partial_X^2 F$  the partial derivatives of  $F$  with respect to  $X$  and (with some slight abuse of notation)  $\partial_\eta F = (\partial_{\eta_k} F)_{k=1, \dots, d}$  as the gradient of  $F$  with respect to the variable  $\eta \in \mathbb{R}^d$  as well as  $\partial_{X, \eta}^2 F = (\partial_{X, \eta_k}^2 F)_{k=1, \dots, d}$  and  $\partial_\eta^2 F = (\partial_{\eta_k, \eta_l}^2 F)_{k, l=1, \dots, d}$  the Hessian of  $F$  with respect to the variable  $\eta$ .
- $(\eta_t(x))_{t \geq 0, x \in I}$  is uniformly bounded in  $(t, x)$ . We also assume that there exists  $\eta_\infty$  Lipschitz continuous on  $I$  such that

$$\delta_t := \sup_{x \in I} |\eta_t(x) - \eta_\infty(x)| \xrightarrow{t \rightarrow \infty} 0. \quad (4.2.1)$$

- The memory kernel  $h$  is nonnegative and integrable on  $[0, +\infty)$ .
- We assume that  $W : I^2 \rightarrow [0, 1]$  is continuous. We refer nonetheless to Section 4.2.3.4 where we show that the results of the paper remain true under weaker hypotheses on  $W$ .

It has been showed in Agathe-Nerine (2022) that the process defined in (4.1.8) is well-posed, and that the large population limit intensity (4.1.13) is well defined in the following sense.

**Proposition 4.2.2.** *Under Hypothesis 4.2.1, for a fixed realization of the family  $(\pi_i)_{1 \leq i \leq N}$ , there exists a pathwise unique multivariate Hawkes process (in the sense of Definition 4.1.2) such that for any  $T < \infty$ ,  $\sup_{t \in [0, T]} \sup_{1 \leq i \leq N} \mathbf{E}[Z_{N,i}(t)] < \infty$ .*

**Proposition 4.2.3.** *Let  $T > 0$ . Under Hypothesis 4.2.1, there exists a unique solution  $\lambda$  in  $\mathcal{C}_b([0, T] \times I, \mathbb{R})$  to (4.1.13) and this solution is nonnegative.*

Both Propositions 4.2.2 and 4.2.3 can be found in Agathe-Nerine (2022) as Propositions 2.5 and 2.7 respectively, where  $F$  is chosen as  $\eta = (\mu, v)$  and  $F(x, \eta) = f(x + v)$  with  $f$  a Lipschitz function. The same proofs work for our general case  $F$ .

Proposition 4.2.3 also implies that the limiting spatial profile  $X_t$  solving (4.1.12) is well defined.

Before writing our next hypothesis, we need to introduce the following integral operator.

**Proposition 4.2.4.** *Under Hypothesis 4.2.1, the integral operator*

$$\begin{aligned} T_W : H &\longrightarrow H \\ g &\longmapsto (T_W g : x \longmapsto \int_I W(x, y)g(y)dy) \end{aligned}$$

is continuous in both cases  $H = L^\infty(I)$  and  $H = L^2(I)$ . When  $H = L^2(I)$ ,  $T_W$  is compact, its spectrum is the union of  $\{0\}$  and a discrete sequence of eigenvalues  $(\mu_n)_{n \geq 1}$  such that  $\mu_n \rightarrow 0$  as  $n \rightarrow \infty$ . Denote by  $r_\infty = r_\infty(T_W)$ , respectively  $r_2 = r_2(T_W)$  the spectral radii of  $T_W$  in  $L^\infty(I)$  and  $L^2(I)$  respectively. Moreover, we have that

$$r_2(T_W) = r_\infty(T_W). \tag{4.2.2}$$

The proof can be found in Section 4.3.1.

**Hypothesis 4.2.5.** In the whole article, we are in the subcritical case defined by

$$\|\partial_X F\|_\infty \|h\|_1 r_\infty < 1. \tag{4.2.3}$$

Note that in the complete mean-field case,  $W \equiv 1$  and  $r_\infty = 1$  so that one retrieves the usual subcritical condition as in Delattre *et al.* (2016a). In the linear case  $\eta = \mu$  and  $F(x, \eta) = \mu + x$ , (4.2.3) is exactly the subcritical condition stated in Agathe-Nerine (2022).

The aim of the paper is twofold: firstly, we state a general convergence result as  $t \rightarrow \infty$  of  $X_t$  defined in (4.1.12) (or equivalently  $\lambda_t$  in (4.1.13)), see Theorem 4.2.7. This result is valid for any general kernel  $h$  satisfying Hypothesis 4.2.1. Secondly, we address the long-term stability of the microscopic profile  $X_N$  defined in (4.1.11), see Theorem 4.2.12. Contrary to the first one, this second result is stated for the particular choice of the exponential kernel  $h$  defined as

$$h(t) = e^{-\alpha t}, \text{ with } \alpha > 0. \tag{4.2.4}$$

The parameter  $\alpha > 0$  is often called the leakage rate. The main advantage of this choice is that the process  $X_N$  then becomes Markovian (see Remark 4.1.4). This will turn out to be particularly helpful for the proof of Theorem 4.2.12. As already mentioned in the introduction, (4.2.4) is the natural framework where to observe the NFE (4.1.3) as a macroscopic limit, recall Remark 4.1.3. Note that in the exponential case (4.2.4), we have that  $\|h\|_1 = 1/\alpha$  hence the subcritical case (4.2.3) reads

$$\|\partial_X F\|_\infty r_\infty < \alpha. \tag{4.2.5}$$

For our second result (Theorem 4.2.12), we also need some hypotheses on the dilution of the graph. Recall the definition of  $\rho_N$  in Definition 4.1.1.

**Hypothesis 4.2.6.** The dilution parameter  $\rho_N \in [0, 1]$  satisfies the following dilution condition: there exists  $\tau \in (0, \frac{1}{2})$  such that

$$N^{1-2\tau} \rho_N^4 \xrightarrow{N \rightarrow \infty} \infty. \quad (4.2.6)$$

If one supposes further that  $F$  is bounded, we assume the weaker condition

$$N \rho_N^2 \xrightarrow{N \rightarrow \infty} \infty. \quad (4.2.7)$$

*Remark.* Hypothesis 4.2.6 is stronger than  $\frac{N \rho_N}{\log N} \xrightarrow{N \rightarrow \infty} \infty$ , which is a dilution condition commonly met in the literature concerning LLN results on bounded time intervals for interacting particles on random graphs: it is the same as in [Delattre et al. \(2016b\)](#); [Coppini et al. \(2019\)](#) (and slightly stronger than the optimal  $N \rho_N \rightarrow +\infty$  obtained in [Coppini et al. \(2022\)](#) in the case of diffusions and as in [Agathe-Nerine \(2022\)](#) in the case of Hawkes processes).

## 4.2.2 Main results

Our first result, Theorem 4.2.7, studies the limit as  $t \rightarrow \infty$  of the macroscopic profile  $X_t$  (as an element of  $\mathcal{C}(I)$ ) defined in (4.1.12). Our second result, Theorem 4.2.12, focuses on the large time behaviour of  $X_N(t)$  defined in (4.1.11) on any time interval of polynomial length.

### 4.2.2.1 Asymptotic behavior of $(X_t)$

Recall the definition of  $X_t$  in (4.1.12).

**Theorem 4.2.7.** *Under Hypotheses 4.2.1 and 4.2.5,*

(i) *there exists a unique continuous function  $X_\infty : I \mapsto \mathbb{R}^+$  solution of*

$$X_\infty = \|h\|_1 T_W F(X_\infty, \eta_\infty). \quad (4.2.8)$$

(ii)  *$X_t$  converges uniformly on  $I$  when  $t \rightarrow \infty$  towards  $X_\infty$ .*

The proof can be found in Section 4.3.2.

**Remark 4.2.8.** Translating the result of Theorem 4.2.7 in terms of the macroscopic intensity  $\lambda_t$  defined in (4.1.13) gives immediately that  $\lambda_t$  converges uniformly to  $\ell$  solution to

$$\ell = F(\|h\|_1 T_W \ell, \eta_\infty). \quad (4.2.9)$$

The correspondence between  $X_\infty$  and  $\ell$  (recall (4.1.12)) is simply given by  $X_\infty = \|h\|_1 T_W \ell$ .

**Remark 4.2.9.** In the particular case of an exponential memory kernel (4.2.4), as a straightforward consequence of the expression of  $X_t$  in (4.1.12) and  $X_\infty$  in (4.2.8), we have the following differential equation

$$\partial_t (X_t - X_\infty) = -\alpha (X_t - X_\infty) + T_W (F(X_t, \eta_t) - F(X_\infty, \eta_\infty)). \quad (4.2.10)$$

A simple Taylor expansion of  $X_t$  around  $X_\infty$  shows that the linearised system associated to the nonlinear (4.2.10) is then

$$\partial_t Y_t = -\alpha Y_t + T_W(GY_t), \quad (4.2.11)$$

where

$$G := \partial_X F(X_\infty, \eta_\infty). \quad (4.2.12)$$

The subcritical condition (4.2.5) translates into the existence of a spectral gap for the linear dynamics (4.2.11), which makes the stationary point  $X_\infty$  linearly stable. More precisely,

**Proposition 4.2.10.** *Assume that the memory kernel  $h$  is exponential (4.2.5). Define the linear operator*

$$\begin{aligned} \mathcal{L} : L^2(I) &\longrightarrow L^2(I) \\ g &\longmapsto \mathcal{L}(g) = -\alpha g + T_W(Gg). \end{aligned} \quad (4.2.13)$$

Then under Hypotheses 4.2.1 and 4.2.5,  $\mathcal{L}$  generates a contraction semi-group  $(e^{t\mathcal{L}})_{t \geq 0}$  on  $L^2(I)$  such that for any  $g \in L^2(I)$

$$\|e^{t\mathcal{L}}g\|_2 \leq e^{-t\gamma} \|g\|_2, \quad (4.2.14)$$

where

$$\gamma := \alpha - r_\infty \|\partial_X F\|_\infty > 0. \quad (4.2.15)$$

The proof can be found in Section 4.3.1.

#### 4.2.2.2 Long-term stability of the microscopic spatial profile

From now on, we place ourselves in the exponential case (4.2.4). We first state a convergence result of  $X_N$  towards the macroscopic  $X$  on a bounded time interval  $[0, T]$ .

**Proposition 4.2.11.** *Let  $T > 0$ . Under Hypotheses 4.2.1, 4.2.5 and 4.2.6,  $\mathbb{P}$ -a.s. for any  $\varepsilon > 0$ ,*

$$\mathbf{P} \left( \sup_{t \in [0, T]} \|X_N(t) - X_t\|_2 \geq \varepsilon \right) \xrightarrow{N \rightarrow \infty} 0. \quad (4.2.16)$$

The proof can be found in Section 4.7. Note that Proposition 4.2.11 slightly generalises (Agathe-Nerine, 2022, Proposition 3.17), that is Proposition 3.3.17 in the present thesis, where it is proven that  $\mathbf{E} \left[ \int_0^T \int_I |X_N(t)(x) - X_t(x)| dx dt \right] \xrightarrow{N \rightarrow \infty} 0$  for any  $T > 0$  (see also (Chevallier *et al.*, 2019, Cor. 2) for a similar result). Here, we are more precise as we show uniform convergence of  $X_N(t)$  in  $L^2(I)$  instead of  $L^1(I)$ .

We are now in position to state the main result of the paper: the proximity stated in Proposition 4.2.11 is not only valid on a bounded time interval, but propagates to arbitrary polynomial times in  $N\rho_N$ .

**Theorem 4.2.12.** *Choose some  $t_f > 0$  and  $m \geq 1$ . Then, under Hypotheses 4.2.1, 4.2.5 and 4.2.6,  $\mathbb{P}$ -a.s. for any  $\varepsilon > 0$ ,*

$$\mathbf{P} \left( \sup_{t \in [t_\varepsilon, (N\rho_N)^m t_f]} \|X_N(t) - X_\infty\|_2 \geq \varepsilon \right) \xrightarrow{N \rightarrow \infty} 0, \quad (4.2.17)$$

for some  $t_\varepsilon > 0$  independent on  $N$ .

The proof can be found in Section 4.4.

*Remark.* The variable  $t_\varepsilon$  in Theorem (4.2.12) represents essentially the time for the deterministic dynamics  $X_t$  to reach a neighborhood of  $X_\infty$  of size  $\varepsilon$ . The time  $t_\varepsilon$  is of order  $-\log \varepsilon / \gamma$  (where  $\gamma$  is the spectral gap (4.2.15) given by the mean-field dynamics) and diverges as  $\varepsilon \rightarrow 0$ . Using the fact that on any finite  $[0, T]$  (and in particular on  $[0, t_\varepsilon]$  for any fixed  $\varepsilon$ ),  $X_{N,t}$  converges as  $N \rightarrow \infty$  to  $X_t$  (Proposition 4.2.11), a simple triangle inequality gives that the following statement is also true: under the Hypotheses of Theorem 4.2.12,  $\mathbb{P}$ -a.s., for all  $\varepsilon > 0$ ,

$$\mathbf{P} \left( \sup_{t \in [0, (N\rho_N)^m t_f]} \|X_N(t) - X_t\|_2 \geq \varepsilon \right) \xrightarrow{N \rightarrow \infty} 0.$$

Since  $F$  is Lipschitz and  $\lambda_{N,i}(t) = F(X_{N,i}(t-), \eta_t(x_i))$  by (4.1.9), it is straightforward to derive from Theorem 4.2.12 a similar result for the profile of intensities

$$\lambda_N(t)(x) := \sum_{i=1}^N \lambda_{N,i}(t) \mathbf{1}_{x \in B_{N,i}}, \quad x \in I, \quad (4.2.18)$$

where  $B_{N,i}$  is defined in (4.1.7).

**Corollary 4.2.13.** *Recall the definition of  $\ell$  in (4.2.9). Under the same set of hypotheses of Theorem 4.2.12 and with the same notation,*

$$\mathbf{P} \left( \sup_{t \in [t_\varepsilon, (N\rho_N)^m t_f]} \|\lambda_N(t) - \ell\|_2 \geq \varepsilon \right) \xrightarrow{N \rightarrow \infty} 0. \quad (4.2.19)$$

## 4.2.3 Examples and extensions

We give here some illustrating examples of our main results.

### 4.2.3.1 Mean-field framework

To the best of the knowledge of the author, already in the simple homogeneous case of mean-field interaction, there exists no long-term stability result such as Theorem 4.2.12. We stress that our result may have an interest of its own in this case. Let us be more specific. When  $\rho_N = W_N = 1$  and  $\mu_t(x) = \mu \geq 0$ , the process introduced in Definition 4.1.2 reduces to the usual mean-field framework [Delattre et al. \(2016a\)](#):

$$Z_{N,i}(t) = \int_0^t \int_0^\infty \mathbf{1}_{\{z \leq \lambda_N(s)\}} \pi_i(ds, dz) \quad (4.2.20)$$

with  $\lambda_N(t)$  defined by

$$\lambda_N(t) = F(X_N(t-), \eta), \quad (4.2.21)$$

where

$$X_N(t) = \sum_{j=1}^N \frac{1}{N} \int_0^t h(t-s) dZ_{N,j}(s), \quad (4.2.22)$$

In this simple case, the spatial framework is no longer useful (in particular the spatial profile defined in (4.1.11) is constant in  $x$  so that the  $L^2$  framework is not relevant, one has only to work in  $\mathbb{R}$ ). The macroscopic intensity and synaptic current (respectively (4.1.13) and (4.1.12) become

$$X_t := \int_0^t h(t-s) \lambda_s ds, \quad \lambda_t := F(X_t, \eta). \quad (4.2.23)$$

The main results of the paper translate then into

**Theorem 4.2.14.** *Under Hypothesis 4.2.1 and when  $\|\partial_X F\|_\infty \|h\|_1 < 1$ , there exists a unique  $X_\infty \in \mathbb{R}_+$  solution to  $X_\infty = \|h\|_1 F(X_\infty, \eta)$ , and  $(X_t)_{t \geq 0}$  defined in (4.2.23) converges when  $t \rightarrow \infty$  towards  $X_\infty$ . Respectively,  $(\lambda_t)_{t \geq 0}$  converges towards the unique solution  $\ell \in \mathbb{R}_+$  to  $\ell = F(\|h\|_1 \ell, \eta)$ . Moreover, under the same hypotheses, in the exponential case (4.2.4), for any  $t_f > 0$  and  $m \geq 1$ ,  $\mathbb{P}$ -a.s. for any  $\varepsilon > 0$ , there exists  $t_\varepsilon > 0$  independent on  $N$  such that both*

$$\mathbf{P} \left( \sup_{t \in [t_\varepsilon, N^m t_f]} |X_N(t) - X_\infty| \geq \varepsilon \right) \text{ and } \mathbf{P} \left( \sup_{t \in [t_\varepsilon, N^m t_f]} |\lambda_N(t) - \ell| \geq \varepsilon \right) \text{ tend to } 0 \text{ as } N \rightarrow \infty.$$

**Remark 4.2.15.** The previous result applies in particular to the linear case where  $\eta = \mu$  and  $F(x, \eta) = \mu + x$ . We have then that  $\ell = \frac{\mu}{1 - \|h\|_1}$  in this case, as in Delattre *et al.* (2016a).

### 4.2.3.2 Erdős-Rényi graphs

An immediate extension of the last mean-field case concerns the case of homogeneous Erdős-Rényi graphs: choose  $W_N(x, y) = \rho_N$  for all  $x, y \in I$ . The results of our paper are valid under the dilution Hypothesis 4.2.6. It is however likely that these dilution conditions are not optimal (compare with the result of Coppini (2022) with the condition  $N\rho_N \rightarrow \infty$  in the diffusion case, but a difficulty here is that we deal with a multiplicative noise whereas it is essentially additive in Coppini (2022)).

### 4.2.3.3 Examples in the inhomogeneous case

As mentionned in Hypothesis 4.2.1, the results are valid for any  $W$  continuous, interesting examples include  $W(x, y) = 1 - \max(x, y)$ ,  $W(x, y) = 1 - xy$ , see Borgs *et al.* (2011, 2018). Note also that we do not suppose any symmetry on  $W$ . Another rich class of examples concerns the *Expected Degree Distribution* model (Chung & Lu, 2002; Ouadah *et al.*, 2019) where  $W(x, y) = f(x)g(y)$  for any continuous functions  $f$  and  $g$  on  $I$ . The specificity of such class is that we have an explicit



formulation of  $r_\infty$ , that is  $r_\infty = \int_I f(x)g(x)dx$  when  $\int_I g = 1$ . In the linear case, we obtain an explicit formula for  $\lambda_t$  in (Agathe-Nerine, 2022, Example 4.3), that is Example 3.4.3 in the present thesis.

#### 4.2.3.4 Extensions (weaker hypothesis on $W$ )

It is apparent from the proofs below that one can weaken the hypothesis of continuity of  $W$ . Under the hypothesis that  $W$  is bounded, Proposition 4.2.3 remains true when  $\mathcal{C}_b([0, T] \times I)$  is replaced by  $\mathcal{C}([0, T], L^\infty(I))$  (continuity of  $\lambda_t$  and  $X_t$  in  $x$  may not be satisfied). Supposing further that there exists a partition of  $I$  into  $p$  intervals  $I = \sqcup_{k=1, \dots, p} C_k$  such that for all  $\epsilon > 0$ , there exists  $\eta > 0$  such that  $\int_I |W(x, y) - W(x', y)| dy < \epsilon$  when  $|x - x'| < \eta$  and  $x, x' \in C_k$ , then for every  $k$ ,  $\lambda_{|[0, T] \times \overset{\circ}{C}_k}$  and  $X_{|[0, T] \times \overset{\circ}{C}_k}$  are both continuous. When  $p = 1$ , both  $\lambda$  and  $X$  are continuous on  $[0, T] \times I$ .

Concerning Theorem 4.2.7, defining for  $k \in \{1, 2\}$ :

$$R_{N,k}^W := \frac{1}{N} \sum_{i,j=1}^N \int_{B_{N,j}} |W(x_i, x_j) - W(x_i, y)|^k dy, \quad (4.2.24)$$

and

$$S_N^W := \sum_{i=1}^N \int_{B_{N,i}} \left( \int_I |W(x_i, y) - W(x, y)|^2 dy \right) dx, \quad (4.2.25)$$

Theorem 4.2.12 remains true when  $R_{N,1}^W, R_{N,2}^W, S_N^W \xrightarrow{N \rightarrow \infty} 0$ , see Lemmas 4.6.4, 4.6.5, 4.6.6 and 4.6.7. These particular conditions are met in the following cases (details of the computation are left to the reader):

- P-nearest neighbor model (Omelchenko *et al.*, 2012):  $W(x, y) = \mathbf{1}_{d_{\mathcal{S}_1}(x,y) < r}$  for any  $(x, y) \in I^2$  for some fixed  $r \in (0, \frac{1}{2})$ , with  $d_{\mathcal{S}_1}(x, y) = \min(|x - y|, 1 - |x - y|)$ ,
- Stochastic block model (Holland *et al.*, 1983; Ditlevsen & Löcherbach, 2017): it corresponds to considering  $p$  communities  $(C_k)_{1 \leq k \leq p}$ . An element of the community  $C_l$  communicates with an element of the community  $C_k$  with probability  $p_{kl}$ . This corresponds to the choice of interaction kernel  $W(x, y) = \sum_{k,l} p_{kl} \mathbf{1}_{x \in C_k, y \in C_l}$ .

#### 4.2.3.5 Extensions (subcritical case)

The point of this paragraph is to discuss the possibility of relaxing the subcriticality condition given in (4.2.3). This condition is used at several times in the paper:

- (a) as a sufficient condition to ensure the existence of a (unique) fixed-point  $X_\infty$  to (4.2.8) (see Theorem 4.2.7 (i)),
- (b) to prove the convergence of the deterministic  $X_t$  to  $X_\infty$  (see Theorem 4.2.7 (ii)), for general  $h$ , not necessarily exponential,

(c) to prove the long-term stability of  $X_N$  around  $X_\infty$ , in the exponential case (see Theorem 4.2.12).

The trickiest point is actually the first one (a), i.e. the existence of a fixed-point to (4.2.8). To fix ideas, let us think of the pure mean-field case seen in Section 4.2.3.1, for the generic example (4.1.1) (that is,  $F(X, \eta) = \mu + f(X)$ ). The fixed-point relation (4.2.8) is then finite-dimensional and it reduces to find  $X_\infty \in \mathbb{R}$  solution to

$$X_\infty = \|h\|_1 (\mu + f(X_\infty)). \quad (4.2.26)$$

The condition (4.2.3) is then the same as

$$\|f'\|_\infty \|h\|_1 < 1, \quad (4.2.27)$$

which is essentially the generic subcriticality condition that one finds in the literature for mean-field nonlinear Hawkes processes (Brémaud & Massoulié, 1996). In the linear case (corresponding to  $f(x) = x$ ), condition (4.2.27) reduces to  $\|h\|_1 < 1$  which is optimal. However, it is quite obvious that (4.2.3)/(4.2.27) is no longer optimal with respect to the existence of a solution to (4.2.26) for general  $f$ : there may very-well be a unique fixed-point to (4.2.26) whereas (4.2.27) is violated, for example in the case a sigmoid  $f$  (sufficiently close to the Heaviside function  $\mathbf{1}_{[\kappa, +\infty)}$  for some  $\kappa > 0$ ): as long as  $\kappa \notin [\|h\|_1 \mu, \|h\|_1 (\mu + 1)]$ , there is a unique solution to (4.2.26) whereas (4.2.27) is not true, as  $\|f'\|_\infty$  is very large. Not to mention the possibility of having several (three) fixed-points in this sigmoid case, while (4.2.27) still does not hold. In this homogenous mean-field case, one can compute the solution to (4.2.26) *by hand*, as it reduces to a simple equation in dimension 1. The situation gets really more complicated in the spatially-extended setting as (4.2.8) is intrinsically infinite dimensional. It is unclear if there exists a condition (that would be weaker than (4.2.3)) ensuring the existence of a (possibly non unique) solution to (4.2.8) for general  $W$ .

However, an important point is the following: *provided we have obtained the existence of such  $X_\infty$ , unique or not, solution to (4.2.8)* (which is again a straightforward task for (4.2.26) but may be complicated for (4.2.8)), points (b) (the convergence  $X_t \rightarrow X_\infty$  in Theorem 4.2.7) at least for  $h$  exponential and (c) (the long-term stability result in Theorem 4.2.12) remain valid under the weaker condition

$$\sup_x |\partial_X F(X_\infty(x), \eta_\infty(x))| \|h\|_1 r_\infty < 1. \quad (4.2.28)$$

Condition (4.2.28) is weaker than (4.2.3) in the sense that it is only local, around  $X_\infty$ , whereas (4.2.3) is global (note that in the homogeneous case (4.2.28) translates into  $|f'(X_\infty)| \|h\|_1 < 1$ , to compare with (4.2.27)). The only modification one has to make in the statements of Theorems 4.2.7 and 4.2.12 is that they are now essentially local, i.e. valid provided the initial condition  $X_0$  is sufficiently close to  $X_\infty$ . More precisely, an alternative statement of item (ii) of Theorem 4.2.7 would be:

**Proposition 4.2.16.** *Suppose that Hypothesis 4.2.1 is true and that we are in the exponential case (4.2.4). Assume the existence of some  $X_\infty \in L^2(I)$  solution to*

(4.2.8) such that (4.2.28) is satisfied. Then, there exists some  $\varepsilon_0 > 0$  such that whenever  $\|X_0 - X_\infty\|_2 < \varepsilon_0$ , one has  $X_t \xrightarrow[t \rightarrow \infty]{} X_\infty$  in  $L^2(I)$ .

However, this extension of point (b) is only valid when  $h$  is exponential. A convergence result under (4.2.28) for general  $h$  (not necessarily exponential) remains open: as it is, the proof of Theorem 4.2.7 uses in an essential way the uniform condition (4.2.3). In a same way, the corresponding local stability result concerning  $X_N$  is then

**Theorem 4.2.17.** *Choose some  $t_f > 0$  and  $m \geq 1$ . Assume Hypotheses 4.2.1 and 4.2.6. Assume the existence of some  $X_\infty \in L^2(I)$  solution to (4.2.8) such that (4.2.28) is satisfied. Let  $\varepsilon_0 > 0$  given by Proposition 4.2.16 and assume that  $\|X_0 - X_\infty\|_2 < \varepsilon_0$ . Suppose that for all  $\varepsilon > 0$ ,  $\mathbf{P}(\|X_N(0) - X_0\|_2 > \varepsilon) \xrightarrow[N \rightarrow \infty]{} 0$ . Then  $\mathbb{P}$ -a.s., for any  $\varepsilon > 0$  (4.2.17) is true, for some  $t_\varepsilon$  independent of  $N$ .*

Remark in particular that the operator  $\mathcal{L}$  in Proposition 4.2.10 (whose spectral properties are the main key to the long term stability result) is only expressed in terms of  $G = \partial_X F(X_\infty, \eta_\infty)$ , that is the exact local quantity appearing in (4.2.28). In particular, under (4.2.28), the spectral gap  $\gamma$  in (4.2.15) becomes  $\gamma = \alpha - \sup_x |\partial_X F(X_\infty(x), \eta_\infty(x))| > 0$  and the rest of the proof follows from the same arguments. We stress also that this result never requires the fixed-point  $X_\infty$  to be unique (it is indeed the case under the present condition (4.2.3) but it is never used in the proof of the long-term stability result, that is essentially a result of local nature, around  $X_\infty$ ).

#### 4.2.4 Link with the literature

Several previous works have complemented the propagation of chaos result mentioned in (4.1.4) in various situations: Central Limit Theorems (CLT) have been obtained in Delattre *et al.* (2016a); Ditlevsen & Löcherbach (2017) for homogeneous mean-field Hawkes processes (when both time and  $N$  go to infinity) or with age-dependence in Chevallier (2017). One should also mention the functional fluctuation result recently obtained in Heesen & Stannat (2021), also in a pure mean-field setting. A result closer to our case with spatial extension is Chevallier & Ost (2020), where a functional CLT is obtained for the spatial profile  $X_N$  around its limit. Some insights of the necessity of considering stochastic versions of the NFE (4.1.3) as second order approximations of the spatial profile are in particular given in Chevallier & Ost (2020). Note here that all of these works provide approximation results of quantities such that  $\lambda_N$  or  $X_N$  that are either valid on a bounded time interval  $[0, T]$  or under strict growth condition on  $T$  (see in particular the condition  $\frac{T}{N} \rightarrow 0$  for the CLT in Ditlevsen & Löcherbach (2017)), whereas we are here concerned with time-scales that grow polynomially with  $N$ .

The analysis of mean-field interacting processes on long time scales has a significant history in the case of interacting diffusions. The important issue of uniform propagation of chaos has been especially studied mostly in reversible situations (see e.g. the case of granular media equation in Bolley *et al.* (2013)) but also more recently in some irreversible situations (see Colombani & Bris (2022)).

There has been in particular a growing interest in the long-time analysis of phase oscillators (for a comprehensive review on the subject, see [Giacomin & Poquet \(2015\)](#) and references therein). We do not aim here to be exhaustive, but as the techniques used in this work present some formal similarities, let us nonetheless comment on the analysis of the simplest situation, i.e. the Kuramoto model. One is here interested in the longtime behavior of the empirical measure  $\mu_{N,t} := \frac{1}{N} \sum_{i=1}^N \delta_{\theta_{i,t}}$  of the system of interacting diffusions  $(\theta_1, \dots, \theta_N)$  solving the system of coupled SDEs  $d\theta_{i,t} = -\frac{K}{N} \sum_{j=1}^N \sin(\theta_{i,t} - \theta_{j,t}) dt + dB_{i,t}$ . Standard propagation of chaos techniques show that  $\mu_N$  converges weakly on a bounded time interval  $[0, T]$  to the solution  $\mu_t$  to the nonlinear Fokker-Planck (NFP) equation  $\partial_t \mu_t = \frac{1}{2} \partial_\theta^2 \mu_t + K \partial_\theta (\mu_t (\sin * \mu_t))$ . The simplicity of the Kuramoto model lies in the fact that one can easily prove the existence of a phase transition for this model: when  $K \leq 1$ ,  $\mu \equiv \frac{1}{2\pi}$  is the only (stable) stationary point of the previous NFP (subcritical case), whereas it coexists with a stable circle of synchronised profiles when  $K > 1$  (supercritical case). A series of papers have analysed the longtime behavior of the empirical measure  $\mu_N$  of the Kuramoto model (and extensions) in both the subcritical and supercritical cases (see in particular [Bertini \*et al.\* \(2014\)](#); [Luçon & Poquet \(2017\)](#); [Giacomin \*et al.\* \(2012\)](#); [Coppini \(2022\)](#) and references therein). The main arguments of the mentioned papers lie in a careful analysis of two contradictory phenomena that arise on a long-time scale: the stability of the deterministic dynamics around stationary points (that forces  $\mu_N$  to remain in a small neighborhood of these points) and the presence of noise in the microscopic system (which makes  $\mu_N$  diffuse around these points). In particular, the work that is somehow formally closest to the present article is [Coppini \(2022\)](#), where the long-time stability of  $\mu_N$  is analysed in both sub and supercritical cases for Kuramoto oscillators interacting on an Erdős-Rényi graph. We are here (at least formally) in a similar situation to the subcritical case of [Coppini \(2022\)](#): the deterministic dynamics of the spatial profile  $X_N$  (given by (4.1.11)) has a unique stationary point which possesses sufficient stability properties. The point of the analysis relies then on a time discretization and some careful control on the diffusive influence of noise that competes with the deterministic dynamics. The main difference (and present difficulty in the analysis) with the diffusion case in [Coppini \(2022\)](#) is that our noise (Poissonian rather than Brownian) is multiplicative (whereas it is essentially additive in [Coppini \(2022\)](#)). This explains in particular the stronger dilution conditions that we require in Hypothesis 4.2.6 (compared to the optimal  $N\rho_N \rightarrow \infty$  in [Coppini \(2022\)](#)) and also the fact that we only reach polynomial time scales (compared to the sub-exponential scale in [Coppini \(2022\)](#)). There is however every reason to believe that the stability result of Theorem 4.2.12 would remain valid up to this sub-exponential time scale.

Note here that we deal directly with the control of the Poisson noise. Another possibility would have been to use some Brownian approximation of the dynamics of  $X_N$ . Some results in this direction have been initiated in [Ditlevsen & Löcherbach \(2017\)](#) for spatially-extended Hawkes processes exhibiting oscillatory behaviors: some diffusive approximation of the dynamics of the (equivalent of) the spatial

profile is provided (see (Ditlevsen & Löcherbach, 2017, Section 5)). Note however that this approximation is based on the comparison of the corresponding semigroups and is not uniform in time. Hence, it is unclear how one could exploit these techniques for our case. Some stronger (pathwise) approximations between Hawkes and Brownian dynamics have been further proposed in Chevallier *et al.* (2021), based on Komlós, Major and Tusnády (KMT) coupling techniques (Ethier & Kurtz, 1986). Recently, Prodhomme (2023) used similar KMT coupling techniques applied to finite dimensional Markov chains and found the Gaussian approximation to remain precise for very large periods of time. However these results are valid for  $\mathbb{Z}^d$ -valued continuous-time Markov chains, and it is unclear how they can be applied in our situation (with infinite dimension and space extension). The proof we propose is direct and does not rely on such Brownian coupling. Another recent work Erny *et al.* (2022) about Hawkes processes with mean-field interactions in a diffusive regime extended also the propagation of chaos to longer time periods, but the scaling used there is different from ours. This diffusion scaling can also be found in Pfaffelhuber *et al.* (2022). To the author’s knowledge, this is the first result on large time stability of Hawkes process (not mentioning the issue of the random graph of interaction, we believe that our results remain also relevant in the pure mean-field case, see Theorem 4.2.14).

#### 4.2.5 Strategy of proof and organization of the paper

Section 4.3 is devoted to prove the convergence result as  $t \rightarrow \infty$  of Theorem 4.2.7. This in particular requires some spectral estimates on the operator  $\mathcal{L}$  defined in Proposition 4.2.10 that are gathered in Section 4.3.1.

The main lines of proof for Theorem 4.2.12 are given in Section 4.4. The strategy of proof is sketched here:

1. The starting point of the analysis is a semimartingale decomposition of  $Y_N := X_N - X$ , detailed in Section 4.4.1. The point is to decompose the dynamics of  $Y_N$  in terms of, at first order, the linear dynamics (4.2.11) governing the behavior of the deterministic profile  $X$ , modulo some drift terms coming from the graph and its mean-field approximation, some noise term and finally some quadratic remaining error coming from the nonlinearity of  $F$ .
2. A careful control on each of these terms in the semimartingale expansion on a bounded time interval is given in the remaining of Section 4.4.1. The proofs of these estimates are respectively given in Section 4.5 (for the noise term) and Section 4.6 (for the drift term).
3. The rest of Section 4.4 is devoted to the proof of Theorem 4.2.12, see Section 4.4.2. The first point is that for any given  $\varepsilon > 0$ , one has to wait a deterministic time  $t_\varepsilon > 0$ , so that the deterministic profile  $X_t$  reaches an  $\varepsilon$ -neighborhood of  $X_\infty$ . It is easy to see from the spectral gap estimate (4.2.14) that this  $t_\varepsilon$  is actually of order  $\frac{-\log(\varepsilon)}{\gamma}$ . Then, using Proposition 4.2.11, the microscopic process  $X_N$  is itself  $\varepsilon$ -close to  $X_\infty$  with high probability.

4. The previous argument is the starting point of an iterative procedure that works as follows: the point is to see that provided  $X_N$  is initially close to  $X_\infty$ , it will remain close to  $X_\infty$  on some  $[0, T]$  for some sufficiently large deterministic  $T > 0$ . The key argument is that on a bounded time interval, the deterministic linear dynamics dominates upon the contribution of the noise, so that one has only to wait some sufficiently large  $T$  so that the deterministic dynamics prevails upon the other contributions.
5. The rest of the proof consists in an iterative procedure from the previous argument, taking advantage of the Markovian structure of the dynamics of  $X_N$ . The time horizon at which one can pursue this recursion is controlled by moment estimates on the noise, proven in Section 4.5.

The rest of the paper is organised as follows: Section 4.7 collects the proofs for the finite time behavior of Proposition 4.2.11 whereas some technical estimates are gathered in the Appendix 4.A.

## 4.3 Asymptotic behavior of $(X_t)$

This section is related to the proof of Theorem 4.2.7.

### 4.3.1 Estimates on the operator $\mathcal{L}$

*Proof of Proposition 4.2.4.* The continuity and compactness of  $T_W$  come from the boundedness of  $W$ . The structure of the spectrum of  $T_W$  is a consequence of the spectral theorem for compact operators. The equality between the spectral radii is postponed to Lemma 4.A.8 where a more general result is stated (see also Proposition 4.7 of Agathe-Nerine (2022) for a similar result), that is Proposition 3.4.7 in the present thesis.  $\square$

*Proof of Proposition 4.2.10.* Let us introduce the operator

$$\begin{aligned} \mathcal{U} : L^2(I) &\longrightarrow L^2(I) \\ g &\longmapsto \mathcal{U}(g) = T_W(Gg), \end{aligned} \tag{4.3.1}$$

we have then  $\mathcal{L} = -\alpha Id + \mathcal{U}$ . By Hypothesis 4.2.1,  $G$  is bounded. Then, for any  $g \in L^2(I)$  using Cauchy-Schwarz inequality,  $\|\mathcal{U}(g)\|_2^2 \leq \|W\|_2^2 \|G\|_\infty^2 \|g\|_2^2$ . The operator  $\mathcal{U}$  is then compact and thus has a discrete spectrum. Moreover,  $r_2(\mathcal{U}) = r_\infty(\mathcal{U})$ , see Lemma 4.A.8, and  $r_\infty(\mathcal{U}) \leq r_\infty(T_W) \|G\|_\infty$  as for any  $g \in L^\infty$  and  $x \in I$ ,  $|\mathcal{U}g(x)| \leq \|T_W\|_\infty \|Gg\|_\infty \leq \|T_W\|_\infty \|G\|_\infty \|g\|_\infty$ . Then  $\mathcal{L}$  also has a discrete spectrum, which is the same as  $\mathcal{U}$  but shifted by  $\alpha$ . Since  $r_2(\mathcal{U}) = r_\infty(\mathcal{U})$  (see Lemma 4.A.8), for any  $\mu \in \sigma(\mathcal{L}) \setminus \{0\}$ ,  $|\mu + \alpha| \leq r_\infty(\mathcal{U})$  thus  $Re(\mu) \leq -\alpha + r_\infty(\mathcal{U}) \leq -\alpha + r_\infty\|\partial_u F\|_\infty < 0$  by (4.2.3). The estimate (4.2.14) follows then from functional analysis (see e.g. (Pazy, 1974, Theorem 3.1)).  $\square$

### 4.3.2 About the large time behavior of $X_t$

*Proof of Theorem 4.2.7.* We prove that

- there exists a unique function  $\ell : I \mapsto \mathbb{R}^+$  solution of (4.2.9), continuous and bounded on  $I$ , and that
- $(\lambda_t)_{t \geq 0}$  converges uniformly when  $t \rightarrow \infty$  towards  $\ell$ .

It gives then that  $X_\infty := \|h\|_1 T_W \ell$  is the unique solution of (4.2.8). Then, as  $X_t(x) = \int_I W(x, y) \int_0^t h(t-s) \lambda_s(y) ds dy$ , as  $(\lambda_t)$  is uniformly bounded, and as  $h$  is integrable and  $\lambda_t \rightarrow \ell$  uniformly, we conclude by dominated convergence that uniformly on  $y$ ,  $\int_0^t h(t-s) \lambda_s(y) ds \xrightarrow[t \rightarrow \infty]{} \|h\|_1 \ell(y)$ . As  $T_W$  is continuous, the result follows:  $X_t$  converges uniformly towards  $X_\infty$ . We show first that  $(\lambda_t)$  is uniformly bounded. Let  $\bar{\lambda}_t(x) = \sup_{s \in [0, t]} \lambda_s(x)$ , we have then with (4.1.13), for  $s \in [0, t]$

$$\begin{aligned} \lambda_s(x) &\leq F(0, 0) + \|F\|_L |\eta_s(x)| + \|\partial_X F\|_\infty \int_I W(x, y) \int_0^s h(s-u) \lambda_u(y) du dy \\ &\leq F(0, 0) + \|F\|_L \sup_{s, x} |\eta_s(x)| + \|\partial_X F\|_\infty \|h\|_1 T_W \bar{\lambda}_t(x), \end{aligned}$$

hence  $\bar{\lambda}_t(x) \leq C_{F, \eta} + \|\partial_X F\|_\infty \|h\|_1 T_W \bar{\lambda}_t(x)$ . An immediate iteration gives then  $\bar{\lambda}_t(x) \leq C_{F, \eta, n_0, h} + \|\partial_X F\|_\infty^{n_0} \|h\|_1^{n_0} |T_W^{n_0} \bar{\lambda}_t(x)|$ , so that, by (4.2.3) and choosing  $n_0$  sufficiently large such that  $\|\partial_X F\|_\infty^{n_0} \|h\|_1^{n_0} \|T_W\|^{n_0} < 1$ , we obtain that  $\|\bar{\lambda}_t\|_\infty < C$ , where  $C$  is independent of  $t$ . Passing to the limit as  $t \rightarrow \infty$ , this implies that  $(\lambda_t)_{t > 0}$  is then uniformly bounded, i.e.  $\sup_{t > 0} \sup_{x \in I} |\lambda_t(x)| =: \|\lambda\|_\infty < \infty$ .

We show next that  $(\lambda_t)$  converges pointwise. We start by studying the supremum limit of  $\lambda_t$ , denoted by  $x \mapsto \bar{\ell}(x) := \limsup_{t \rightarrow \infty} \lambda_t(x) = \inf_{r > 0} \sup_{t > r} \lambda_t(x) =: \inf_{r > 0} \Lambda(r, x)$ . Then for any  $x \in I$ ,  $r > 0$  and  $t > r$ ,  $\lambda_t(x)$  can be decomposed into

$$F \left( \int_I W(x, y) \int_0^r h(t-s) \lambda_s(y) ds dy + \int_I W(x, y) \int_r^t h(t-s) \lambda_s(y) ds dy, \eta_t(x) \right),$$

which by monotonicity of  $F$  in the first variable and by positivity of  $W$  and  $h$  is upper bounded by the term

$$F \left( \int_I W(x, y) \int_0^r h(t-s) \lambda_s(y) ds dy + \int_I W(x, y) \Lambda(r, y) \int_r^t h(t-s) ds dy, \eta_t(x) \right).$$

As  $\int_r^t h(t-s) ds \leq \|h\|_1$ , it gives

$$\lambda_t(x) \leq F \left( \int_I W(x, y) \int_0^r h(t-s) \lambda_s(y) ds dy + \|h\|_1 \int_I W(x, y) \Lambda(r, y) dy, \eta_t(x) \right),$$

and as  $h(t) \rightarrow 0$ , by dominated convergence  $\int_I W(x, y) \int_0^r h(t-s) \lambda_s(y) ds dy \xrightarrow[t \rightarrow \infty]{} 0$  and by continuity and monotonicity of  $F$ , we obtain

$$\bar{\ell}(x) \leq F \left( \|h\|_1 (T_W \bar{\ell})(x), \eta_\infty(x) \right). \quad (4.3.2)$$

Note that  $\|\bar{\ell}\|_\infty \leq \|\lambda\|_\infty < \infty$ , by the first step of this proof. Denote in a same way  $\underline{\ell}(x) := \liminf_{t \rightarrow \infty} \lambda_t(x) = \sup_{r > 0} \inf_{t > r} \lambda_t(x) =: \sup_{r > 0} v(r, x)$ , for any  $t > 0$  we have by monotonicity of  $F$  in the first variable:

$$\begin{aligned} & \lambda_t(x) \\ = & F \left( \int_0^{\frac{t}{2}} \int_I W(x, y) h(t-s) \lambda_s(y) dy ds + \int_{\frac{t}{2}}^t \int_I W(x, y) h(t-s) \lambda_s(y) dy ds, \eta_t(x) \right) \\ \geq & F \left( \int_{\frac{t}{2}}^t \int_I W(x, y) h(t-s) v \left( \frac{t}{2}, y \right) dy ds, \eta_t(x) \right) \\ = & F \left( \int_0^{\frac{t}{2}} h(u) du \int_I W(x, y) v \left( \frac{t}{2}, y \right) dy, \eta_t(x) \right). \end{aligned}$$

Taking  $\liminf_{t \rightarrow \infty}$  on both sides, by monotone convergence, we obtain

$$\underline{\ell}(x) \geq F(\|h\|_1 (T_W \underline{\ell})(x), \eta_\infty(x)). \quad (4.3.3)$$

Combining (4.3.2) and (4.3.3), setting  $H : l \in L^\infty \mapsto F(\|h\|_1 T_W l, \eta_\infty) \in L^\infty$ , we have shown

$$H \underline{\ell} \leq \underline{\ell} \leq \bar{\ell} \leq H \bar{\ell}. \quad (4.3.4)$$

For any  $l$  and  $l'$  in  $L^\infty(I)$  and any  $x \in I$ , we have

$$\begin{aligned} |Hl(x) - Hl'(x)| & \leq |F(\|h\|_1 (T_W l)(x), \eta_\infty(x)) - F(\|h\|_1 (T_W l')(x), \eta_\infty(x))| \\ & \leq \|\partial_X F\|_\infty \|h\|_1 |(T_W(l - l'))(x)|. \end{aligned}$$

By iteration one can show that  $\|H^{n_0} l - H^{n_0} l'\|_\infty \leq \|\partial_u F\|_\infty^{n_0} \|h\|_1^{n_0} \|T_W^{n_0}\| \|l - l'\|_\infty$ , so that, choosing again  $n_0$  sufficiently large,  $H^{n_0}$  is a contraction mapping by (4.2.3). Hence, by (4.3.4), one has necessarily that for all  $x \in I$   $\underline{\ell}(x) = \bar{\ell}(x) < +\infty$  thus  $(\lambda_t)$  converges pointwise towards  $\ell = \underline{\ell} = \bar{\ell}$  the unique fixed point of  $H$  which satisfies (4.2.9).

We show now that the family  $(\lambda_t)_{t \geq 0}$  is equicontinuous so that the pointwise convergence will imply uniform convergence on the compact set  $I$ . For any  $(x, y) \in I$  and  $t \geq 0$ , we have

$$\begin{aligned} |\lambda_t(x) - \lambda_t(y)| & = |F(X_t(x), \eta_t(x)) - F(X_t(y), \eta_t(y))| \\ & \leq \|F\|_L (|X_t(x) - X_t(y)| + |\eta_t(x) - \eta_t(y)|). \end{aligned}$$

With (4.2.1), we have

$$\begin{aligned} |\eta_t(x) - \eta_t(y)| & \leq |\eta_t(x) - \eta_\infty(x)| + |\eta_\infty(x) - \eta_\infty(y)| + |\eta_\infty(y) - \eta_t(y)| \\ & \leq 2\delta_t + \|\eta_\infty\|_L |x - y|, \end{aligned}$$

and as  $\lambda$  is bounded, we have

$$\begin{aligned} |X_t(x) - X_t(y)| & = \left| \int_I (W(x, z) - W(y, z)) \int_0^t h(t-s) \lambda_s(z) ds dz \right| \\ & \leq \|\lambda\|_\infty \|h\|_1 \int_I |W(x, z) - W(y, z)| dz. \end{aligned} \quad (4.3.5)$$



Then  $|\lambda_t(x) - \lambda_t(y)| \leq C_{F,\lambda,h,W} (\delta_t + |x - y| + \int_I |W(x, z) - W(y, z)| dz)$ . Fix  $\varepsilon > 0$ , with (4.2.1), one can find  $T$  such that  $C_{F,\lambda,h,W} \delta_t \leq \frac{\varepsilon}{2}$  for any  $t \geq T$ , and as  $W$  is uniformly continuous on  $I^2$ , one can find  $\eta > 0$  such that

$$C_{F,\lambda,h,W} \left( |x - y| + \int_I |W(x, z) - W(y, z)| dz \right) \leq \frac{\varepsilon}{2}$$

when  $|x - y| \leq \eta$ . We can divide  $[0, 1]$  in intervals  $[z_k, z_{k+1}]$  such that for any  $k$ ,  $z_{k+1} - z_k \leq \eta$ . Then, for any  $x \in [0, 1]$ , one can find  $z_k$  such that  $|z_k - x| \leq \eta$ , and  $|\lambda_t(x) - \ell(x)| \leq |\lambda_t(x) - \lambda_t(z_k)| + |\lambda_t(z_k) - \ell(z_k)| + |\ell(z_k) - \ell(x)|$ . By pointwise convergence,  $|\lambda_t(z_k) - \ell(z_k)| \leq \varepsilon$  for  $t$  large enough (but independent of the choice of  $x$ ), and  $|\ell(z_k) - \ell(x)| \leq \varepsilon$  by taking the limit when  $t \rightarrow \infty$  in  $|\lambda_t(z_k) - \lambda_t(x)| \leq \varepsilon$ . It gives then  $|\lambda_t(x) - \ell(x)| \leq 3\varepsilon$  hence  $\sup_{x \in I} |\lambda_t(x) - \ell(x)| \xrightarrow{t \rightarrow \infty} 0$ , i.e.  $(\lambda_t)$  converges uniformly towards  $\ell$ . Similarly to (4.3.5), for any  $x, x' \in I$ ,

$$|X_\infty(x) - X_\infty(x')| \leq \|h\|_1 \|\ell\|_\infty \int_I |W(x, y) - W(x', y)| dy$$

which gives, as  $W$  is uniformly continuous, the continuity of  $X_\infty$ .  $\square$

## 4.4 Large time behavior of $X_N(t)$

The aim of the present section is to prove Theorem 4.2.12. To study the behavior of  $\|X_N(t) - X_\infty\|_2$ , let

$$Y_N := X_N - X_\infty. \quad (4.4.1)$$

The first step is to write the semimartingale decomposition of  $Y_N$ , written in a mild form (see Section 4.4.1). The proper control on the drift and noise terms are given in Propositions 4.4.2 and 4.4.3. In Section 4.4.2, we give the proof of Theorem 4.2.12, based in particular on the convergence on a bounded time interval in Proposition 4.2.11.

### 4.4.1 Mild formulation

**Proposition 4.4.1.** *The process  $(Y_N(t))_{t \geq 0}$  satisfies the following semimartingale decomposition in  $D([0, T], L^2(I))$ , written in a mild form: for any  $0 \leq t_0 \leq t$*

$$Y_N(t) = e^{(t-t_0)\mathcal{L}} Y_N(t_0) + \phi_N(t_0, t) + \zeta_N(t_0, t) \quad (4.4.2)$$

where (recall (4.1.7) the partition of  $I$  in  $N$  segments of equal length):

$$\phi_N(t_0, t) = \int_{t_0}^t e^{(t-s)\mathcal{L}} r_N(s) ds \quad (4.4.3)$$

$$\zeta_N(t_0, t) = \int_{t_0}^t e^{(t-s)\mathcal{L}} dM_N(s), \quad (4.4.4)$$

with

$$M_N(t) = \sum_{i=1}^N \sum_{j=1}^N \frac{w_{ij}}{N} \left( Z_{N,j}(t) - \int_0^t \lambda_{N,j}(s) ds \right) \mathbf{1}_{B_{N,i}} \quad (4.4.5)$$

and

$$r_N(t)(x) = T_W(g_N(t))(x) + \sum_{i=1}^N \left( \frac{1}{N\rho_N} \sum_{j=1}^N \xi_{ij}^{(N)} F(X_{N,j}(t-), \eta_t(x_j)) - \int_I W(x, y) F(X_N(t, y), \eta_t(y)) dy \right) \mathbf{1}_{B_{N,i}}(x) \quad (4.4.6)$$

with

$$\begin{aligned} g_N(t)(y) &:= \partial_\eta F(X_\infty(y), \eta_\infty(y)) \cdot (\eta_t(y) - \eta_\infty(y)) + \int_0^1 (1-r) (\eta_t(y) - \eta_\infty(y)) \cdot \\ &\quad \partial_\eta^2 F(X_\infty(y) + rY_N(t)(y), (1-r)\eta_\infty(y) + r\eta_t(y)) (\eta_t(y) - \eta_\infty(y)) dr \\ &+ \int_0^1 (1-r) \partial_x^2 F(X_\infty(y) + rY_N(t)(y), (1-r)\eta_\infty(y) + r\eta_t(y)) Y_N(t)(y)^2 dr \\ &\quad + \int_0^1 2(1-r) \partial_{x,\eta}^2 Y_N(t)(y) \cdot \\ &\quad F(X_\infty(y) + rY_N(t)(y), (1-r)\eta_\infty(y) + r\eta_t(y)) \cdot (\eta_t(y) - \eta_\infty(y)) dr. \end{aligned} \quad (4.4.7)$$

$\phi_N$  is the drift term and  $\zeta_N$  is the noise term coming from the jumps of the process  $X_N$ .

*Proof of Proposition 4.4.1.* From (4.1.10) and (4.1.11), as we are in the exponential case (4.2.4), we obtain that  $X_N$  verifies

$$dX_N(t) = -\alpha X_N(t) dt + \sum_{i=1}^N \sum_{j=1}^N \frac{w_{ij}}{N} dZ_{N,j}(t) \mathbf{1}_{B_{N,i}}. \quad (4.4.8)$$

The centered noise  $M_N$  defined in (4.4.5) verifies

$$dM_N(t) := \sum_{i=1}^N \sum_{j=1}^N \frac{w_{ij}}{N} (dZ_{N,j}(t) - F(X_{N,j}(t-), \eta_t(x_j)) dt) \mathbf{1}_{B_{N,i}},$$

and is a martingale in  $L^2(I)$ . Thus recalling the definition of  $X_\infty$  in (4.2.8) and by inserting the term  $\sum_{i=1}^N \sum_{j=1}^N \frac{w_{ij}}{N} F(X_{N,j}(t-), \eta_t(x_j)) dt \mathbf{1}_{B_{N,i}}$  in (4.4.8), we obtain

$$\begin{aligned} dY_N(t) &= -\alpha Y_N(t) + dM_N(t) + \sum_{i=1}^N \left( \sum_{j=1}^N \frac{w_{ij}}{N} F(X_{N,j}(t-), \eta_t(x_j)) \right) \mathbf{1}_{B_{N,i}} dt \\ &\quad - T_W F(X_\infty, \eta_\infty) dt. \end{aligned} \quad (4.4.9)$$

A Taylor's expansion gives

$$F(X_N(t, y), \eta_t(y)) - F(X_\infty(y), \eta_\infty(y)) = \partial_X F(X_\infty(y), \eta_\infty(y)) (X_N(t, y) - X_\infty(y)) + g_N(t)(y),$$

with  $g_N$  given in (4.4.7). Hence, we have with  $G$  defined in (4.2.12)

$$T_W F(X_\infty, \eta_\infty)(x) = \int_I W(x, y) F(X_N(t, y), \eta_t(y)) dy - T_W(GY_N(t)) - T_W g_N(t)(x),$$

hence coming back to (4.4.9) and recognizing the operator  $\mathcal{L}$  (4.2.13)

$$dY_N(t) = \mathcal{L}Y_N(t) + dM_N(t) + \sum_{i=1}^N \left( \sum_{j=1}^N \frac{w_{ij}}{N} F(X_{N,j}(t-), \eta_t(x_j)) \right) \mathbf{1}_{B_{N,i}} dt - T_W F(X_N(t, \cdot), \eta_t(\cdot)) dt + T_W g_N(t).$$

We recognize  $r_N$  defined in (4.4.6), and obtain exactly

$$dY_N(t) = \mathcal{L}Y_N(t) dt + r_N(t) dt + dM_N(t). \quad (4.4.10)$$

Then the mild formulation (4.4.2) is a direct consequence of (Zhu *et al.*, 2017, Lemma 3.2): the unique strong solution to (4.4.10) is indeed given by (4.4.2).  $\square$

**Proposition 4.4.2** (Noise perturbation). *Let  $m \geq 1$  and  $T > t_0 \geq 0$ . Under Hypotheses 4.2.1 and 4.2.6, there exists a constant  $C = C(T, m, F, \eta_0) > 0$  such that  $\mathbb{P}$ -almost surely for  $N$  large enough:*

$$\mathbf{E} \left[ \sup_{s \leq T} \|\zeta_N(t_0, s)\|_2^{2m} \right] \leq \frac{C}{(N\rho_N)^m}.$$

The proof is postponed to Section 4.5.

**Proposition 4.4.3** (Drift term). *Under Hypothesis 4.2.1, for any  $t \geq t_0 > 0$ ,  $\mathbb{P}$ -almost surely if  $N$  is large enough,*

$$\|\phi_N(t_0, t)\|_2 \leq C_{drift} \left( \int_{t_0}^t e^{-(t-s)\gamma} \|Y_N(s)\|_2^2 ds + G_N + \int_{t_0}^t e^{-\gamma(t-s)} (\delta_s^2 + \delta_s) ds \right), \quad (4.4.11)$$

where  $C_{drift} = C_{W,F,\alpha}$ ,  $\gamma$  is defined in (4.2.15),  $\delta_s$  is defined in (4.2.1) and  $G_N = G_N(\xi)$  is an explicit quantity to be found in the proof that tends to 0 as  $N \rightarrow \infty$ .

The proof is postponed to Section 4.6.

## 4.4.2 Proof of the large time behaviour

We prove here Theorem 4.2.12, based on the results of Section 4.4.1. The approach followed is somehow formally similar to the strategy of proof developed in Coppini (2022) for the diffusion case.

*Proof of Theorem 4.2.12.* Choose  $m \geq 1$  and  $t_f > 0$ . Let

$$\varepsilon_0 = \frac{\gamma}{6C_{drift}}, \quad (4.4.12)$$

where  $\gamma$  is defined in (4.2.15) and the constant  $C_{drift}$  comes from Proposition 4.4.3 above. Note that it suffices to consider  $\varepsilon$  small enough, such that  $\varepsilon < \varepsilon_0$ :  $t_\varepsilon$  defined below increases as  $\varepsilon \searrow 0$ , so that it suffices to take  $t_\varepsilon = t_{\varepsilon_0}$  whenever  $\varepsilon \geq \varepsilon_0$ . As  $(X_t)$  converges uniformly towards  $X_\infty$  (Theorem 4.2.7), there exists  $t_\varepsilon^1 < \infty$  such that

$$\|X_t - X_\infty\|_2 \leq \frac{\varepsilon}{4}, \quad t \geq t_\varepsilon^1. \quad (4.4.13)$$

Moreover, with (4.2.1), we also have that  $\int_0^t e^{-\gamma(t-s)} (\delta_s^2 + \delta_s) ds \xrightarrow[t \rightarrow \infty]{} 0$ , hence there exists  $t_\varepsilon^2 < \infty$  such that

$$C_{\text{drift}} \int_0^t e^{-\gamma(t-s)} (\delta_s^2 + \delta_s) ds \leq \frac{\varepsilon}{18}, \quad t \geq t_\varepsilon^2. \quad (4.4.14)$$

We set now  $t_\varepsilon = \max(t_\varepsilon^1, t_\varepsilon^2)$ . Let  $T$  such that

$$e^{-\gamma T} < \frac{1}{3}, \quad T > t_f. \quad (4.4.15)$$

The strategy of proof relies on the following time discretisation. The point is to control  $\|X_N(t) - X_\infty\|_2$  on  $[t_\varepsilon, T_N]$  with

$$T_N := a_N T + t_\varepsilon, \quad \text{with } a_N := \lceil (N\rho_N)^m \rceil, \quad (4.4.16)$$

which will imply the result (4.2.17) as  $[t_\varepsilon, (N\rho_N)^m t_f] \subset [t_\varepsilon, T_N]$  since  $T > t_f$ . We decompose below the interval  $[t_\varepsilon, T_N]$  into  $a_N$  intervals of length  $T$ . We define the following events, with  $0 \leq t_a \leq t_b$  (recall that  $Y_N(t) = X_N(t) - X_\infty$ )

$$A_1^N(t, \varepsilon) := \left\{ \|Y_N(t)\|_2 \leq \frac{\varepsilon}{2} \right\} \quad \text{for } t \geq 0, \quad (4.4.17)$$

$$A_2^N(\varepsilon) := \left\{ \sup_{t \in [t_\varepsilon, t_\varepsilon + T]} \|\zeta_N(t_\varepsilon, t)\|_2 \leq \frac{\varepsilon}{18} \right\}, \quad (4.4.18)$$

$$E(t_a, t_b) := \left\{ \max \left( 2\|Y_N(t_a)\|_2, \sup_{t \in [t_a, t_b]} \|Y_N(t)\|_2, 2\|Y_N(t_b)\|_2 \right) \leq \varepsilon \right\}. \quad (4.4.19)$$

As Proposition 4.2.11 gives that  $\mathbf{P} \left( \sup_{t \in [0, t_\varepsilon]} \|Y_N(t)\|_2 > \frac{\varepsilon}{4} \right) \xrightarrow[N \rightarrow \infty]{} 0$  and by (4.4.13), we have by triangle inequality

$$\mathbf{P} (A_1^N(t_\varepsilon, \varepsilon)) \xrightarrow[N \rightarrow \infty]{} 1. \quad (4.4.20)$$

**Step 1** We have from the definition (4.4.19) of  $E(t_a, t_b)$  that

$$\mathbf{P} \left( \sup_{t \in [t_\varepsilon, T_N]} \|Y_N(t)\|_2 \leq \varepsilon \right) \geq \mathbf{P} (E(t_\varepsilon, T_N)) = \mathbf{P} (E(t_\varepsilon, T_N) | A_1^N(t_\varepsilon, \varepsilon)) \mathbf{P} (A_1^N(t_\varepsilon, \varepsilon)). \quad (4.4.21)$$

Moreover,

$$\begin{aligned} \mathbf{P} (E(t_\varepsilon, T_N) | A_1^N(t_\varepsilon, \varepsilon)) &= \mathbf{P} (E(t_\varepsilon, t_\varepsilon + a_N T) | A_1^N(t_\varepsilon, \varepsilon)) \\ &\geq \mathbf{P} (E(t_\varepsilon, t_\varepsilon + a_N T) \cap E(t_\varepsilon, t_\varepsilon + (a_N - 1)T) | A_1^N(t_\varepsilon, \varepsilon)) \\ &= \mathbf{P} (E(t_\varepsilon, t_\varepsilon + a_N T) | E(t_\varepsilon, t_\varepsilon + (a_N - 1)T) \cap A_1^N(t_\varepsilon, \varepsilon)) \\ &\quad \cdot \mathbf{P} (E(t_\varepsilon, t_\varepsilon + (a_N - 1)T) | A_1^N(t_\varepsilon, \varepsilon)). \end{aligned}$$

Recall that we are in the exponential case (4.2.4), so that  $(X_N(t))_t$  is a Markov process. Thus by Markov property

$$\begin{aligned} & \mathbf{P} \left( E(t_\varepsilon, t_\varepsilon + a_N T) | E(t_\varepsilon, t_\varepsilon + (a_N - 1)T) \cap A_1^N(t_\varepsilon, \varepsilon) \right) \\ &= \mathbf{P} \left( E(t_\varepsilon + (a_N - 1)T, t_\varepsilon + a_N T) | E(t_\varepsilon, t_\varepsilon + (a_N - 1)T) \right) \\ &= \mathbf{P} \left( E(t_\varepsilon + (a_N - 1)T, t_\varepsilon + a_N T) \left| \left\{ \|Y_N(t_\varepsilon + (a_N - 1)T)\|_2 \leq \frac{\varepsilon}{2} \right\} \right. \right) \\ &= \mathbf{P} \left( E(t_\varepsilon + (a_N - 1)T, t_\varepsilon + a_N T) | A_1^N(t_\varepsilon + (a_N - 1)T, \varepsilon) \right). \end{aligned}$$

$\mathbf{P} \left( E(t_\varepsilon + (a_N - 1)T, t_\varepsilon + a_N T) | A_1^N(t_\varepsilon + (a_N - 1)T, \varepsilon) \right)$  means that, under an initial condition at  $t_\varepsilon + (a_N - 1)T$ , we look at the probability that  $Y_N$  stays below  $\varepsilon$  on the interval  $[t_\varepsilon + (a_N - 1)T, t_\varepsilon + a_N T]$  of size  $T$  and comes back below  $\frac{\varepsilon}{2}$  at the final time  $t_\varepsilon + a_N T$ . By Markov's property, it is exactly  $\mathbf{P} \left( E(t_\varepsilon, t_\varepsilon + T) | A_1^N(t_\varepsilon, \varepsilon) \right)$ . An immediate iteration gives then

$$\mathbf{P} \left( E(t_\varepsilon, T_N) | A_1^N(t_\varepsilon, \varepsilon) \right) \geq \mathbf{P} \left( E(t_\varepsilon, t_\varepsilon + T) | A_1^N(t_\varepsilon, \varepsilon) \right)^{a_N}. \quad (4.4.22)$$

By (4.4.20), from now on we consider that we are on this event  $A_1^N(t_\varepsilon, \varepsilon)$  and omit this notation for simplicity.

**Step 2** We show that

$$A_2^N(\varepsilon) \subset E(t_\varepsilon, t_\varepsilon + T). \quad (4.4.23)$$

Let us place ourselves in  $A_2^N(\varepsilon)$ . As we are also under  $A_1^N(t_\varepsilon, \varepsilon)$ , we have indeed  $\|Y_N(t_\varepsilon)\|_2 \leq \frac{\varepsilon}{2}$  for the first condition of  $E(t_\varepsilon, t_\varepsilon + T)$ . As  $Y_N$  verifies (4.4.1), it can be written for  $t \geq t_\varepsilon$

$$Y_N(t) = e^{\mathcal{L}(t-t_\varepsilon)} Y_N(t_\varepsilon) + \phi_N(t_\varepsilon, t) + \zeta_N(t_\varepsilon, t). \quad (4.4.24)$$

For any  $t \in [t_\varepsilon, t_\varepsilon + T]$ ,

$$\begin{aligned} \|\phi_N(t_\varepsilon, t)\|_2 &\leq C_{\text{drift}} \left( \int_{t_\varepsilon}^t e^{-(t-s)\gamma} \|Y_N(s)\|_2^2 ds + G_N + \int_{t_0}^t e^{-\gamma(t-s)} (\delta_s^2 + \delta_s) ds \right) \\ &\leq C_{\text{drift}} \left( \int_{t_\varepsilon}^t e^{-(t-s)\gamma} \|Y_N(s)\|_2^2 ds \right) + \frac{\varepsilon}{9} \end{aligned} \quad (4.4.25)$$

where the first inequality comes from Proposition 4.4.3, and the second is true for  $N$  large enough using  $G_N \rightarrow 0$  and (4.4.14). Coming back to (4.4.24), using that by Proposition 4.2.10

$$\|e^{\mathcal{L}(t-t_\varepsilon)} Y_N(t_\varepsilon)\|_2 \leq e^{-\gamma(t-t_\varepsilon)} \|Y_N(t_\varepsilon)\|_2, \quad (4.4.26)$$

and using (4.4.25), we have on  $A_1^N(t_\varepsilon, \varepsilon) \cap A_2^N(\varepsilon)$

$$\|Y_N(t)\|_2 \leq \frac{\varepsilon}{2} + C_{\text{drift}} \left( \int_{t_\varepsilon}^t e^{-(t-s)\gamma} \|Y_N(s)\|_2^2 ds \right) + \frac{\varepsilon}{9} + \frac{\varepsilon}{18}.$$

Let  $\delta > 0$  such that  $\delta \leq \min \left( \frac{\varepsilon}{6}, \frac{\gamma}{9C_{\text{drift}}} \right)$ . Recall that  $\|Y_N(\cdot)\|_2$  is not a continuous function, it jumps whenever a spike of the process  $(Z_{N,1}, \dots, Z_{N,N})$  occurs, but the

size jump never exceeds  $\frac{1}{N}$ , and for  $N$  large enough  $\frac{1}{N} \leq \frac{\delta}{2}$ . Then, one can apply Lemma 4.A.9 and obtain that for all  $N$  large enough,

$$\sup_{t \in [t_\varepsilon, t_\varepsilon + T]} \|Y_N(t)\|_2 \leq \frac{\varepsilon}{2} + 3\delta \leq \varepsilon. \quad (4.4.27)$$

It remains to prove that  $\|Y_N(t_\varepsilon + T)\|_2 \leq \frac{\varepsilon}{2}$ . We obtain from (4.4.24), (4.4.25) and (4.4.26) for  $t = t_\varepsilon + T$  on  $A_1^N(t_\varepsilon, \varepsilon) \cap A_2^N(\varepsilon)$

$$\|Y_N(t_\varepsilon + T)\|_2 \leq e^{-\gamma T} \frac{\varepsilon}{2} + \frac{\varepsilon}{6} + C_{\text{drift}} \int_{t_\varepsilon}^{t_\varepsilon + T} e^{-(t_\varepsilon + T - s)\gamma} \|Y_N(s)\|_2^2 ds.$$

Using the a priori bound (4.4.27)

$$\|Y_N(t_\varepsilon + T)\|_2 \leq e^{-\gamma T} \frac{\varepsilon}{2} + \frac{\varepsilon}{12} + \varepsilon^2 \frac{C_{\text{drift}}}{\gamma} \leq e^{-\gamma T} \frac{\varepsilon}{2} + \frac{\varepsilon}{6} + \frac{\varepsilon}{6} \leq \frac{\varepsilon}{2},$$

where we recall the particular choices of  $T$  in (4.4.15) and  $\varepsilon < \varepsilon_0$  in (4.4.12). This concludes the proof of (4.4.23).

**Step 3** We obtain with (4.4.22) and Markov's inequality,

$$\begin{aligned} \mathbf{P}(E(t_\varepsilon, T_N)) &\geq \mathbf{P}(E(t_\varepsilon, t_\varepsilon + T))^{a_N} \geq \mathbf{P}(A_2^N(\varepsilon))^{a_N} \\ &= \left(1 - \mathbf{P}\left(\sup_{t \in [t_\varepsilon, t_\varepsilon + T]} \|\zeta_N(t_\varepsilon, t)\|_2 > \frac{\varepsilon}{18}\right)\right)^{a_N} \\ &\geq \left(1 - 18^{2m'} \frac{\mathbb{E}\left[\sup_{t \in [t_\varepsilon, t_\varepsilon + T]} \|\zeta_N(t_\varepsilon, t)\|_2^{2m'}\right]}{\varepsilon^{2m'}}\right)^{a_N}, \end{aligned}$$

where we have taken  $m' > m$ . With Proposition 4.4.2, it gives

$$\mathbf{P}(E(t_\varepsilon, T_N)) \geq \left(1 - \frac{C}{(\varepsilon^2 N \rho_N)^{m'}}\right)^{a_N} = \exp\left(a_N \ln\left(1 - \frac{C}{(\varepsilon^2 N \rho_N)^{m'}}\right)\right).$$

By definition (4.4.16),  $a_N = o(N \rho_N)^{m'}$ , the right term tends to 1 as  $N$  goes to  $\infty$  under Hypothesis 4.2.6. By (4.4.21), we conclude that

$$\mathbf{P}\left(\sup_{t \in [t_\varepsilon, T_N]} \|X_N(t) - X_\infty\|_2 \leq \varepsilon\right) \xrightarrow{N \rightarrow \infty} 1.$$

This concludes the proof of Theorem 4.2.12.  $\square$

## 4.5 Noise perturbation (proofs)

In this section, we prove Proposition 4.4.2 concerning the control of the noise perturbation  $\zeta_N(t_0, t)$  defined in (4.4.4). For simplicity of notation, we assume that

$t_0 = 0$ . Recall the expression of  $(Z_{N,j})_{1 \leq j \leq N}$  in (4.1.8). Introduce the compensated measure  $\tilde{\pi}_j(ds, dz) := \pi_j(ds, dz) - dsdz$ , so that with the linearity of  $(e^{t\mathcal{L}})_{t \geq 0}$ , we obtain that  $\zeta_N$  can be written as

$$\zeta_N(0, t) = \sum_{j=1}^N \int_0^t \int_0^\infty e^{(t-s)\mathcal{L}} \chi_j(s, z) \tilde{\pi}_j(ds, dz), \quad (4.5.1)$$

with  $\chi_j(s, z) := \left( \sum_{i=1}^N \mathbf{1}_{B_{N,i}} \frac{w_{ij}}{N} \right) \mathbf{1}_{z \leq \lambda_{N,j}(s)} \in L^2(I)$ . The proof of Proposition 4.4.3 relies on a adaptation of an argument given in (Zhu *et al.*, 2017, Theorem 4.3), where a similar quantity to (4.5.1) is considered for  $N = 1$ .

#### 4.5.1 Control of the moments of the process $Z_{N,i}$

**Proposition 4.5.1.** *Let  $m \geq 1$  and  $T > 0$ . Under Hypotheses 4.2.1 and 4.2.6,  $\mathbb{P}$ -almost surely*

$$\sup_{N \geq 1} \mathbf{E} \left[ \frac{1}{N} \sum_{j=1}^N Z_{N,j}(T)^m \right] < \infty.$$

*Proof.* Let  $N \geq 1$ . We have for any  $i \in \llbracket 1, N \rrbracket$

$$\begin{aligned} \mathbf{E}[Z_{N,i}(T)^m] &\leq \mathbf{E} \left[ \left( \left( Z_{N,i}(T) - \int_0^T \lambda_{N,i}(t) dt \right) + \int_0^T \lambda_{N,i}(t) dt \right)^m \right] \\ &\leq 2^{m-1} \mathbf{E} \left[ \left( Z_{N,i}(T) - \int_0^T \lambda_{N,i}(t) dt \right)^m \right] + 2^{m-1} \mathbf{E} \left[ \left( \int_0^T \lambda_{N,i}(t) dt \right)^m \right] \\ &\leq 2^{m-1} C \mathbf{E} \left[ \left( \int_0^T \lambda_{N,i}(t) dt \right)^{\frac{m}{2}} \right] + (2T)^{m-1} \mathbf{E} \left[ \int_0^T \lambda_{N,i}(t)^m dt \right], \end{aligned} \quad (4.5.2)$$

where we used Jensen's inequality and Burkholder-Davis-Gundy Inequality on the martingale  $\left( Z_{N,i}(T) - \int_0^T \lambda_{N,i}(t) dt \right)$ . Similarly, we obtain

$$\mathbf{E} \left[ \left( \int_0^T \lambda_{N,i}(t) dt \right)^{\frac{m}{2}} \right] \leq T^{\frac{m}{2}-1} \mathbf{E} \left[ \int_0^T (\lambda_{N,i}(t))^{\frac{m}{2}} dt \right].$$

We focus now on the term  $\mathbf{E} \left[ \int_0^T \lambda_{N,i}(t)^k dt \right]$  for  $k \geq 1$ . From the definition of  $\lambda_{N,i}$  in (4.1.9), by Lipschitz continuity of  $F$  and with Jensen's inequality

$$\begin{aligned} \mathbf{E} \left[ \int_0^T \lambda_{N,i}(t)^k dt \right] &\leq 2^{k-1} T F(0, \eta_t(x_i))^k \\ &\quad + 2^{k-1} \|F\|_L^k \mathbf{E} \left[ \int_0^T \left( \frac{1}{N} \sum_{j=1}^N \int_0^{t-} w_{ij} e^{-\alpha(t-s)} dZ_{N,j}(s) \right)^k dt \right]. \end{aligned}$$

Let  $S_i := \sum_{j=1}^N \frac{w_{ij}}{N}$ . By Lemma 4.A.4, we have that  $\mathbb{P}$ -a.s.,  $\limsup_{N \rightarrow \infty} \sup_{1 \leq i \leq N} S_i \leq 2$ . We obtain with discrete Jensen's inequality that for any  $t \geq 0$

$$\begin{aligned} \left( \frac{1}{N} \sum_{j=1}^N \int_0^{t^-} w_{ij} e^{-\alpha(t-s)} dZ_{N,j}(s) \right)^k &\leq S_i^k \left( \sum_{j=1}^N \frac{w_{ij}}{N S_i} Z_{N,j}(t) \right)^k \\ &\leq S_i^{k-1} \sum_{j=1}^N \frac{w_{ij}}{N} Z_{N,j}(t)^k. \end{aligned}$$

We obtain then

$$\mathbf{E} \left[ \int_0^T \lambda_{N,i}(t)^k dt \right] \leq C_{T,F,\eta_0,k} + C_{k,F} \sum_{j=1}^N \frac{w_{ij}}{N} \mathbf{E} \left[ \int_0^T Z_{N,j}(t)^k dt \right],$$

thus going back to (4.5.2), with  $C = C_{T,F,\eta_0,m}$

$$\begin{aligned} \mathbf{E} \left[ \frac{1}{N} \sum_{j=1}^N Z_{N,j}(T)^m \right] &\leq \frac{C}{N} \sum_{i=1}^N \left( \mathbf{E} \left[ \int_0^T \lambda_{N,i}(t)^{\frac{m}{2}} dt \right] + C \mathbf{E} \left[ \int_0^T \lambda_{N,i}(t)^m dt \right] \right) \\ &\leq C \left( 1 + \sum_{i,j=1}^N \frac{w_{ij}}{N^2} \mathbf{E} \left[ \int_0^T Z_{N,j}(t)^{\frac{m}{2}} dt \right] + \sum_{i,j=1}^N \frac{w_{ij}}{N^2} \mathbf{E} \left[ \int_0^T Z_{N,j}(t)^m dt \right] \right) \\ &\leq C \left( 1 + \int_0^T \mathbf{E} \left[ \frac{1}{N} \sum_{j=1}^N Z_{N,j}(t)^{\frac{m}{2}} \right] dt + \int_0^T \mathbf{E} \left[ \frac{1}{N} \sum_{j=1}^N Z_{N,j}(t)^m \right] dt \right) \end{aligned}$$

$\mathbb{P}$ -a.s. for  $N$  large enough, where we used (4.A.2). As for any  $t \geq 0$

$$\begin{aligned} \mathbf{E} \left[ \frac{1}{N} \sum_{i=1}^N Z_{N,i}(t) \right] &= \frac{1}{N} \sum_{i=1}^N \mathbf{E} \left[ \int_0^t \lambda_{N,i}(s) ds \right] \\ &\leq C_{T,\eta_0,F} + C_{T,\eta_0,F} \int_0^t \mathbf{E} \left[ \frac{1}{N} \sum_{j=1}^N Z_{N,j}(s) \right] ds, \end{aligned}$$

Grönwall's lemma gives that  $\sup_{t \leq T} \mathbf{E} \left[ \frac{1}{N} \sum_{i=1}^N Z_{N,i}(t) \right] < \infty$  (independently of  $N$ ) and

similarly an immediate iteration gives that  $\sup_{N \geq 1} \mathbf{E} \left[ \frac{1}{N} \sum_{j=1}^N Z_{N,j}(T)^{2^k} \right] < \infty$  for any  $k \geq 0$ , which concludes the proof.  $\square$

## 4.5.2 Proof of Proposition 4.4.2

*Proof.* We divide the proof in different steps. Fix  $m \geq 1$ . We prove Proposition 4.4.2 for the choice  $t_0 = 0$ , but it remains the same for a general initial time  $t_0 \geq 0$ .



**Step 1** The functional  $\phi : L^2(I) \rightarrow \mathbb{R}$  given by  $\phi(v) = \|v\|_2^{2m}$  is of class  $\mathcal{C}^2$  (recall that  $\zeta_N \in L^2(I)$ ) so that by Itô formula on the expression (4.5.1) we obtain

$$\begin{aligned} \phi(\zeta_N(t)) &= \int_0^t \phi'(\zeta_N(s)) \mathcal{L}(\zeta_N(s)) ds + \sum_{j=1}^N \int_0^t \int_0^\infty \phi'(\zeta_N(s-)) \chi_j(s, z) \tilde{\pi}_j(ds, dz) \\ &+ \sum_{j=1}^N \int_0^t \int_0^\infty [\phi(\zeta_N(s-) + \chi_j(s, z)) - \phi(\zeta_N(s-)) - \phi'(\zeta_N(s-)) \chi_j(s, z)] \pi_j(ds, dz) \\ &:= I_0(t) + I_1(t) + I_2(t). \end{aligned} \quad (4.5.3)$$

We have then that for any  $v, h, k \in L^2(I)$ ,  $\phi'(v)h = 2m\|v\|_2^{2m-2}\langle v, h \rangle \in \mathbb{R}$  and  $\phi''(v)(h, k) = 2m(2m-1)\|v\|_2^{2m-4}\langle v, k \rangle \langle v, h \rangle + 2m\|v\|_2^{2m-2}\langle h, k \rangle$ .

**Step 2** We have  $I_0(t) = \int_0^t 2m\|\zeta_N(s)\|_2^{2m-2}\langle \zeta_N(s), \mathcal{L}(\zeta_N(s)) \rangle ds$ . From Proposition 4.2.10,  $\mathcal{L}$  generates a contraction semi-group hence by Lumer-Philipp's Theorem (see (Pazy, 1974, Section 1.4)), for any  $s \geq 0$ ,  $\langle \zeta_N(s), \mathcal{L}(\zeta_N(s)) \rangle \leq 0$ . Then for any  $t \geq 0$  we have  $I_0(t) \leq 0$ .

**Step 3** Let us focus now on  $I_1$  in (4.5.3), denoting by  $\alpha_j(s, z) := 2m\|\zeta_N(s-)\|_2^{2m-2}\langle \zeta_N(s-), \chi_j(s, z) \rangle \in \mathbb{R}$ , we have

$$I_1(t) = \sum_{j=1}^N \int_0^t \int_0^\infty \alpha_j(s, z) \tilde{\pi}_j(ds, dz).$$

$I_1$  is then a real martingale. Using that the  $(\pi_j)_{1 \leq j \leq N}$  are independent so that there are almost surely no simultaneous jumps and hence  $[\tilde{\pi}_j, \tilde{\pi}_{j'}] = 0$  if  $j \neq j'$ ,

$$\begin{aligned} [I_1]_t &= \sum_{j=1}^N \int_0^t \int_0^\infty \alpha_j(s, z)^2 \pi_j(ds, dz) \\ &= \sum_{j=1}^N \int_0^t \int_0^\infty (2m\|\zeta_N(s-)\|_2^{2m-2}\langle \zeta_N(s-), \chi_j(s, z) \rangle)^2 \pi_j(ds, dz) \\ &\leq 4m^2 \sup_{0 \leq s \leq t} (\|\zeta_N(s)\|_2^{4m-2}) \sum_{j=1}^N \int_0^t \int_0^\infty \|\chi_j(s, z)\|_2^2 \pi_j(ds, dz). \end{aligned}$$

Then, by Burkholder-Davis-Gundy inequality, for some  $C > 0$ ,

$$\begin{aligned} &\mathbf{E} \left[ \sup_{s \leq t} |I_1(s)| \right] \\ &\leq C2m \mathbf{E} \left[ \sup_{0 \leq s \leq t} (\|\zeta_N(s)\|_2^{2m-1}) \left( \sum_{j=1}^N \int_0^t \int_0^\infty \|\chi_j(s, z)\|_2^2 \pi_j(ds, dz) \right)^{\frac{1}{2}} \right]. \end{aligned}$$

Applying Hölder inequality with parameter  $\frac{2m-1}{2m} + \frac{1}{2m} = 1$  for the random variables  $\sup_{0 \leq s \leq t} (\|\zeta_N(s)\|_2^{2m-1})$  and  $\left( \sum_{j=1}^N \int_0^t \int_0^\infty \|\chi_j(s, z)\|_2^2 \pi_j(ds, dz) \right)^{\frac{1}{2}}$ , we obtain that  $\mathbf{E} \left[ \sqrt{[I_1]_t} \right]$  is upper bounded by

$$2m \left( \mathbf{E} \left[ \sup_{0 \leq s \leq t} (\|\zeta_N(s)\|_2^{2m}) \right] \right)^{\frac{2m-1}{2m}} \left( \mathbf{E} \left[ \left( \sum_{j=1}^N \int_0^t \int_0^\infty \|\chi_j(s, z)\|_2^2 \pi_j(ds, dz) \right)^m \right] \right)^{\frac{1}{2m}}.$$

Let  $\varepsilon > 0$  to be chosen later. From Young's inequality, for any  $a, b \geq 0$ , we can write  $ab = \left( \varepsilon^{\frac{2m-1}{2m}} a \right) \left( \varepsilon^{-\frac{(2m-1)}{2m}} b \right) \leq \frac{2m-1}{2m} \left( \varepsilon^{\frac{2m-1}{2m}} a \right)^{\frac{2m}{2m-1}} + \frac{1}{2m} \left( \varepsilon^{-\frac{(2m-1)}{2m}} b \right)^{2m} = \frac{2m-1}{2m} \varepsilon a^{\frac{2m}{2m-1}} + \frac{1}{2m} \varepsilon^{-(2m-1)} b^{2m}$ . Then this gives for the choice

$$a = \left( \mathbf{E} \left[ \sup_{0 \leq s \leq t} (\|\zeta_N(s)\|_2^{2m}) \right] \right)^{\frac{2m-1}{2m}} \quad \text{and} \\ b = \left( \mathbf{E} \left[ \left( \sum_{j=1}^N \int_0^t \int_0^\infty \|\chi_j(s, z)\|_2^2 \pi_j(ds, dz) \right)^m \right] \right)^{\frac{1}{2m}},$$

that

$$\mathbf{E} \left[ \sqrt{[I_1]_t} \right] \leq (2m-1)\varepsilon \mathbf{E} \left[ \sup_{0 \leq s \leq t} (\|\zeta_N(s)\|_2^{2m}) \right] \\ + \varepsilon^{-(2m-1)} \mathbf{E} \left[ \left( \sum_{j=1}^N \int_0^t \int_0^\infty \|\chi_j(s, z)\|_2^2 \pi_j(ds, dz) \right)^m \right].$$

We have then shown that, for the constant  $C$  given by Burkholder-Davis-Gundy Inequality,

$$\mathbf{E} \left[ \sup_{s \leq T} |I_1(s)| \right] \leq C(2m-1)\varepsilon \mathbf{E} \left[ \sup_{0 \leq s \leq T} (\|\zeta_N(s)\|_2^{2m}) \right] \\ + C\varepsilon^{-(2m-1)} \mathbf{E} \left[ \left( \sum_{j=1}^N \int_0^T \int_0^\infty \|\chi_j(s, z)\|_2^2 \pi_j(ds, dz) \right)^m \right]. \quad (4.5.4)$$

**Step 4** Let us focus now on  $I_2$  in (4.5.3), recall that it is

$$\sum_{j=1}^N \int_0^t \int_0^\infty [\phi(\zeta_N(s-) + \chi_j(s, z)) - \phi(\zeta_N(s-)) - \phi'(\zeta_N(s-)) \chi_j(s, z)] \pi_j(ds, dz).$$

For any jump  $(s, z)$  of the Poisson measure  $\pi_j$ , from Taylor's Lagrange formula there exists  $\tau_s \in (0, 1)$  such that

$$\phi(\zeta_N(s-) + \chi_j(s, z)) - \phi(\zeta_N(s-)) - \phi'(\zeta_N(s-)) \chi_j(s, z) \\ = \frac{1}{2} \phi''(\zeta_N(s-) + \tau_s \chi_j(s, z)) (\chi_j(s, z), \chi_j(s, z)).$$

As  $\phi''(v)(h, k) = 2m(2m-1)\|v\|_2^{2m-4}\langle v, k\rangle\langle v, h\rangle + 2m\|v\|_2^{2m-2}\langle h, k\rangle$  for any  $v, h, k \in L^2(I)$ , one has with Cauchy–Schwarz inequality that

$$\phi''(\zeta_N(s-) + \tau_s \chi_j(s, z))(\chi_j(s, z))^2 \leq 4m^2 \|\zeta_N(s-) + \tau_s \chi_j(s, z)\|_2^{2m-2} \|\chi_j(s, z)\|_2^2.$$

But as  $\|x + \tau y\|_2^2 \leq \max(\|x\|_2^2, \|x + y\|_2^2)$  for any  $x, y \in L^2(I)$  and  $\tau \in (0, 1)$ , we have here

$$\|\zeta_N(s-) + \tau_s \chi_j(s, z)\|_2^{2m-2} \leq \max(\|\zeta_N(s-)\|_2^{2m-2}, \|\zeta_N(s-) + \chi_j(s, z)\|_2^{2m-2}).$$

As  $\|\zeta_N(s-) + \chi_j(s, z)\|_2^{2m-2} = \|\zeta_N(s)\|_2^{2m-2} \leq \sup_{s \leq t} \|\zeta_N(s)\|_2^{2m-2}$  and  $\|\zeta_N(s-)\|_2^{2m-2} \leq \sup_{s \leq t} \|\zeta_N(s)\|_2^{2m-2}$ , we obtain

$$\mathbf{E} \left[ \sup_{s \leq t} |I_2(s)| \right] \leq 2m^2 \mathbf{E} \left[ \sup_{s \leq t} \|\zeta_N(s)\|_2^{2m-2} \sum_{j=1}^N \int_0^t \int_0^\infty \|\chi_j(s, z)\|_2^2 \pi_j(ds, dz) \right].$$

We proceed now similarly as for  $I_1$ . From Hölder inequality, as  $\frac{2m-2}{2m} + \frac{1}{m} = 1$  we know that for any  $A, B$  random non-negative variables,

$$\mathbf{E}[AB] \leq \left( \mathbf{E} \left[ A^{\frac{2m}{2m-2}} \right] \right)^{\frac{2m-2}{2m}} \left( \mathbf{E}[B^m] \right)^{\frac{1}{m}}.$$

For the choice  $A = \sup_{0 \leq s \leq t} (\|\zeta_N(s)\|_2^{2m-2})$  and  $B = \sum_{j=1}^N \int_0^t \int_0^\infty \|\chi_j(s, z)\|_2^2 \pi_j(ds, dz)$ , it leads to  $\mathbf{E}[\sup_{s \leq t} |I_2(s)|]$  upper bounded by

$$2m^2 \left( \mathbf{E} \left[ \sup_{0 \leq s \leq t} (\|\zeta_N(s)\|_2^{2m}) \right] \right)^{\frac{2m-2}{2m}} \left( \mathbf{E} \left[ \left( \sum_{j=1}^N \int_0^t \int_0^\infty \|\chi_j(s, z)\|_2^2 \pi_j(ds, dz) \right)^m \right] \right)^{\frac{1}{m}}.$$

With the same  $\varepsilon$  introduced for  $I_1$ , from Young's inequality, for any  $a, b \geq 0$ , we can write  $ab = \left( \varepsilon^{\frac{2m-2}{2m}} a \right) \left( \varepsilon^{-\frac{(2m-2)}{2m}} b \right) \leq \frac{2m-2}{2m} \left( \varepsilon^{\frac{2m-2}{2m}} a \right)^{\frac{2m}{2m-2}} + \frac{1}{m} \left( \varepsilon^{-\frac{(2m-2)}{2m}} b \right)^m = \frac{2m-2}{2m} \varepsilon a^{\frac{2m}{2m-2}} + \frac{1}{m} \varepsilon^{-(2m-2)} b^m$ . For the choice

$$a = \left( \mathbf{E} \left[ \sup_{0 \leq s \leq t} (\|\zeta_N(s)\|_2^{2m}) \right] \right)^{\frac{2m-2}{2m}} \quad \text{and} \\ b = \left( \mathbf{E} \left[ \left( \sum_{j=1}^N \int_0^t \int_0^\infty \|\chi_j(s, z)\|_2^2 \pi_j(ds, dz) \right)^m \right] \right)^{\frac{1}{m}},$$

it gives that

$$\mathbf{E} \left[ \sup_{s \leq t} |I_2(s)| \right] \leq m(2m-2)\varepsilon \mathbf{E} \left[ \sup_{0 \leq s \leq t} (\|\zeta_N(s)\|_2^{2m}) \right] \\ + 2m\varepsilon^{-(2m-2)} \mathbf{E} \left[ \left( \sum_{j=1}^N \int_0^t \int_0^\infty \|\chi_j(s, z)\|_2^2 \pi_j(ds, dz) \right)^m \right]. \quad (4.5.5)$$

Taking the expectation in (4.5.3) and combining (4.5.4) and (4.5.5), we obtain that

$$\begin{aligned} \mathbf{E} \left[ \sup_{s \leq T} \|\zeta_N(s)\|_2^{2m} \right] &\leq \varepsilon (C(2m-1) + m(2m-2)) \mathbf{E} \left[ \sup_{0 \leq s \leq T} (\|\zeta_N(s)\|_2^{2m}) \right] \\ &\quad + (C\varepsilon^{-(2m-1)} + 2m\varepsilon^{-(2m-2)}) \mathbf{E} \left[ \left( \sum_{j=1}^N \int_0^T \int_0^\infty \|\chi_j(s, z)\|_2^2 \pi_j(ds, dz) \right)^m \right]. \end{aligned} \quad (4.5.6)$$

**Step 5** We can now fix  $\varepsilon$  such that  $\varepsilon (C(2m-1) + m(2m-2)) \leq \frac{1}{2}$  so that (4.5.6) leads to

$$\mathbf{E} \left[ \sup_{s \leq T} \|\zeta_N(s)\|_2^{2m} \right] \leq 2C \mathbf{E} \left[ \left( \sum_{j=1}^N \int_0^T \int_0^\infty \|\chi_j(s, z)\|_2^2 \pi_j(ds, dz) \right)^m \right], \quad (4.5.7)$$

where  $C > 0$  depends only on  $m$ .

**Step 6** Let  $A_N := \mathbf{E} \left[ \left( \sum_{j=1}^N \int_0^T \int_0^\infty \|\chi_j(s, z)\|_2^2 \pi_j(ds, dz) \right)^m \right]$ . We have

$$\|\chi_j(s, z)\|_2^2 = \int_I \left( \sum_{i=1}^N \mathbf{1}_{B_{N,i}}(x) \frac{w_{ij}}{N} \mathbf{1}_{z \leq \lambda_{N,j}(s)} \right)^2 dx = \mathbf{1}_{z \leq \lambda_{N,j}(s)} \sum_{i=1}^N \frac{\xi_{ij}}{N^3 \rho_N^2},$$

which leads to, with the definition of  $Z_{N,j}$  in (4.1.8)

$$\begin{aligned} A_N &= \mathbf{E} \left[ \left( \sum_{i,j=1}^N \int_0^T \int_0^\infty \mathbf{1}_{z \leq \lambda_{N,j}(s)} \frac{\xi_{ij}}{N^3 \rho_N^2} \pi_j(ds, dz) \right)^m \right] \\ &\leq \left( \frac{1}{N \rho_N} \right)^m \mathbf{E} \left[ \left( \sum_{i,j=1}^N \frac{\xi_{ij}}{N^2 \rho_N} Z_{N,j}(T) \right)^m \right]. \end{aligned}$$

With (4.A.2), Jensen's discrete inequality and (4.5.7), it leads to

$$\begin{aligned} A_N &\leq \left( \frac{1}{N \rho_N} \right)^m \mathbf{E} \left[ \left( \sum_{j=1}^N \frac{1}{N} \left( \sup_j \sum_{i=1}^N \frac{\xi_{ij}}{N \rho_N} \right) Z_{N,j}(T) \right)^m \right] \\ &\leq \frac{C}{(N \rho_N)^m} \mathbf{E} \left[ \frac{1}{N} \sum_{j=1}^N Z_{N,j}(T)^m \right], \end{aligned}$$

hence the result with Proposition 4.5.1.  $\square$

## 4.6 Drift term (proofs)

In this section, we prove Proposition 4.4.3 concerning the control of the drift term perturbation  $\phi_N(t_0, t)$  defined in (4.4.3).

### 4.6.1 Notation

We introduce the following constants

$$\Theta_{t,i,1} := \frac{1}{N\rho_N} \sum_{j=1}^N \left( \xi_{ij}^{(N)} - \rho_N W(x_i, x_j) \right) F(X_{N,j}(t-), \eta_t(x_j)), \quad (4.6.1)$$

$$\Theta_{t,i,2} := \frac{1}{N} \sum_{j=1}^N W(x_i, x_j) (F(X_{N,j}(t-), \eta_t(x_j)) - F(X_{N,j}(t), \eta_t(x_j))), \quad (4.6.2)$$

$$\Theta_{t,i,3} := \frac{1}{N} \sum_{j=1}^N W(x_i, x_j) F(X_{N,j}(t), \eta_t(x_j)) - \int_I W(x_i, y) F(X_N(t, y), \eta_t(y)) dy, \quad (4.6.3)$$

and the auxiliary function in  $L^2(I)$

$$\Theta_{t,i,4}(x) := \int_I (W(x_i, y) - W(x, y)) F(X_N(t, y), \eta_t(y)) dy. \quad (4.6.4)$$

From the expression of  $r_N$  in (4.4.6), we have then

$$r_N(t) = \sum_{i=1}^N \left( \sum_{k=1}^4 \Theta_{t,i,k} \right) \mathbf{1}_{B_{N,i}} + T_W(g_N(t)), \quad (4.6.5)$$

and we can divide  $\phi_N$  defined in (4.4.3) in several terms  $\phi_N(t) = \sum_{k=0}^4 \phi_{N,k}(t)$  with

$$\phi_{N,0}(t) := \int_{t_0}^t e^{(t-s)\mathcal{L}} T_W(g_N(s)) ds, \quad (4.6.6)$$

$$\phi_{N,k}(t) := \int_{t_0}^t e^{(t-s)\mathcal{L}} \sum_{i=1}^N \frac{1}{N} \Theta_{s,i,k} \mathbf{1}_{B_{N,i}} ds \quad \text{for } k \in \llbracket 1, 4 \rrbracket, \quad (4.6.7)$$

### 4.6.2 Preliminary results

**Lemma 4.6.1.** Denoting by  $\tilde{Y}_N(s)(v) := Y_N(s) \left( \frac{\lceil Nv \rceil}{N} \right)$ , we have

$$\sup_{s \geq 0} \|\tilde{Y}_N(s) - Y_N(s)\|_2 \xrightarrow{N \rightarrow \infty} 0. \quad (4.6.8)$$

*Proof.* A direct computation gives, for any  $s \geq 0$ ,

$$\|\tilde{Y}_N(s) - Y_N(s)\|_2^2 = \sum_j \int_{B_{N,j}} (X_{N,j}(s) - X_\infty(x_j) - X_N(s)(y) + X_\infty(y))^2 dy.$$

By definition of  $X_N(s)$  in (4.1.11),  $X_N = X_{N,j}$  on  $B_{N,j}$  hence

$$\|\tilde{Y}_N(s) - Y_N(s)\|_2^2 = \sum_j \int_{B_{N,j}} (X_\infty(y) - X_\infty(x_j))^2 dy.$$

Then (4.6.8) is a straightforward consequence of the uniform continuity of  $X_\infty$  on the compact  $I$  (see Theorem 4.2.7). It still holds under the hypotheses of Section 4.2.3.4 by decomposing the sum on each interval  $C_k$ .  $\square$

We will often use

$$\frac{1}{N} \sum_{j=1}^N |Y_N(s)(x_j)|^2 = \|\tilde{Y}_N(s)\|_2^2 \leq \frac{1}{2} \left(1 + \|\tilde{Y}_N(s)\|_2^4\right) \leq \frac{1}{2} \left(2 + \|Y_N(s)\|_2^4\right), \quad (4.6.9)$$

the last inequality being true for  $N$  large enough (independently of  $s$ ) using Lemma 4.6.1.

**Lemma 4.6.2.** *Under Hypothesis 4.2.1,*

$$R_{N,k}^W \xrightarrow[N \rightarrow \infty]{} 0, \quad k \in \{1, 2\}, \quad S_N^W \xrightarrow[N \rightarrow \infty]{} 0, \quad (4.6.10)$$

where  $R_{N,k}^W$  and  $S_N^W$  are respectively defined in (4.2.24) and (4.2.25).

*Proof.* Fix  $\varepsilon > 0$ . As  $W$  is uniformly continuous on  $I$ , there exists  $\eta > 0$  such that  $|W(x, y) - W(x, z)| \leq \varepsilon$  for any  $(x, y, z) \in I^3$  with  $|y - z| \leq \eta$ . Then, for  $N$  large enough (such that  $\frac{1}{N} \leq \eta$ , we have directly that  $R_{N,1}^W \leq \varepsilon$  and  $R_{N,2}^W \leq \varepsilon$  hence the result. We can do the same for  $S_N^W$ .  $\square$

**Lemma 4.6.3.** *Under Hypothesis 4.2.1, for any  $t > t_0 \geq 0$ ,*

$$\|\phi_{N,0}(t)\|_2 \leq C_{F,W} \int_{t_0}^t e^{-\gamma(t-s)} (\|Y_N(s)\|_2^2 + \delta_s + \delta_s^2) ds. \quad (4.6.11)$$

*Proof.* By Proposition 4.2.10 we have  $\|\phi_{N,0}(t)\|_2 \leq \int_{t_0}^t e^{-\gamma(t-s)} \|T_W g_N(s)\|_2 ds$ . As for any  $x \in I$ ,  $|T_W g_N(s)(x)| \leq \int_I W(x, y) |g_N(s)(y)| dy$ , and as

$$\begin{aligned} |g_N(s)(y)| &\leq \|\partial_X^2 F\|_\infty Y_N(s)(y)^2 + \|\partial_\eta^2 F\|_\infty |\eta_t(y) - \eta_\infty(y)|^2 \\ &\quad + 2 \|\partial_{X,\eta}^2 F\|_\infty |Y_N(s)(y)| |\eta_t(y) - \eta_\infty(y)| + \|\partial_\eta F\|_\infty |\eta_t(y) - \eta_\infty(y)|, \end{aligned}$$

with Hypothesis 4.2.1 it gives

$$\begin{aligned} \|T_W g_N(s)\|_2^2 &= \int_I \left( \int_I W(x, y) g_N(s)(y) dy \right)^2 dx \\ &\leq C_F \int_I \left( \int_I W(x, y) (Y_N(s)(y)^2 + \delta_s^2 + Y_N(s)(y)\delta_s + \delta_s) dy \right)^2 dx \\ &\leq C_{F,W} (\|Y_N(s)\|_2^4 + \|Y_N(s)\|_2^2 \delta_s^2 + \delta_s^2 + \delta_s^4) \\ &\leq C_{F,W} \left( \frac{3}{2} \|Y_N(s)\|_2^4 + \frac{3}{2} \delta_s^2 + \delta_s^4 \right) \end{aligned}$$

as  $W$  is bounded. We obtain then, as  $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$ ,

$$\|\phi_{N,0}(t)\|_2 \leq C_{F,W} \int_{t_0}^t e^{-\gamma(t-s)} (\|Y_N(s)\|_2^2 + \|Y_N(s)\|_2 \delta_s + \delta_s + \delta_s^2) ds.$$

Then (4.6.11) follows as  $\|Y_N(s)\|_2 \leq \frac{1}{2} (1 + \|Y_N(s)\|_2^2)$  and  $\sup_s \delta_s < \infty$ .  $\square$

**Lemma 4.6.4.** *Under Hypotheses 4.2.1 and 4.2.6,  $\mathbb{P}$ -almost surely for  $N$  large enough and for any  $t > t_0 \geq 0$ ,*

$$\|\phi_{N,1}(t)\|_2 \leq C_F \int_{t_0}^t e^{-(t-s)\gamma} \|Y_N(s)\|_2^2 ds + G_{N,1}, \quad (4.6.12)$$

where  $G_{N,1} = G_{N,1}(\xi)$  is explicit in  $N$  and tends to 0 as  $N \rightarrow \infty$ . Moreover, if we suppose  $F$  bounded, we have a better bound

$$\sup_{t>0} \|\phi_{N,1}(t)\|_2 \leq \frac{C_F}{\sqrt{N\rho_N^2}}. \quad (4.6.13)$$

*Proof of Lemma 4.6.4.* Proposition 4.2.10 gives that

$$\|\phi_{N,1}(t)\|_2 \leq K \int_{t_0}^t e^{-(t-s)\gamma} \|\gamma_N(s)\|_2 ds \quad (4.6.14)$$

with

$$\gamma_N(s) := \sum_{i=1}^N \Theta_{i,s,1} \mathbf{1}_{B_{N,i}} = \sum_{i,j=1}^N \frac{1}{N\rho_N} \overline{\xi_{ij}} F(X_{N,j}(s-), \eta_s(x_j)) \mathbf{1}_{B_{N,i}}. \quad (4.6.15)$$

where we have used the notation

$$\overline{\xi_{ij}} = \xi_{ij}^{(N)} - W_N(x_i, x_j), \quad (4.6.16)$$

Forgetting about the term  $F(X_{N,j}(s-), \eta_s(x_j))$  in (4.6.15),  $\gamma_N$  is essentially an empirical mean of the independent centered variables  $\overline{\xi_{ij}}$  and thus should be small as  $N \rightarrow \infty$ . One difficulty here is that concentration bounds (e.g. Bernstein inequality) for weighted sums such as  $\sum_j \overline{\xi_{ij}} u_{i,j}$  (for some deterministic fixed weight  $u_{i,j}$ ) are not directly applicable, as  $u_{i,j} = F(X_{N,j}(s-), \eta_s(x_j)) \mathbf{1}_{B_{N,i}}$  depends in a highly nontrivial way on the variables  $\xi_{i,j}^{(N)}$  themselves. A strategy would be to use Grothendieck inequality (see Theorem 4.A.1). We refer here to Coppini (2022); Coppini *et al.* (2022) where the use of such Grothendieck inequality (and extensions) has been implemented in a similar context of interacting diffusions on random graphs. However here, a supplementary difficulty lies in the fact that  $F$  need not be bounded (recall that a particular example considered here concerns the linear case where  $F(x, \eta) = x + \mu$ ). Hence the application of Grothendieck inequality is not straightforward when  $F$  is unbounded. For this reason, we give below two different controls on  $\gamma_N$ : a general one, without assuming that  $F$  is bounded and a second (sharper) one, when  $F$  is bounded (using Grothendieck inequality). In the first case, we get around the difficulty of unboundedness of  $F$  by introducing  $F(X_\infty(x_j), \eta_\infty(x_j))$  which is bounded, since  $X_\infty$  is.

First begin with the general control on  $\gamma_N$ : we can write

$$\begin{aligned} \gamma_N(s) &= \sum_{i,j=1}^N \frac{1}{N\rho_N} \overline{\xi_{ij}} (F(X_{N,j}(s-), \eta_s(x_j)) - F(X_\infty(x_j), \eta_\infty(x_j))) \mathbf{1}_{B_{N,i}} \\ &\quad + \sum_{i,j=1}^N \frac{1}{N\rho_N} \overline{\xi_{ij}} F(X_\infty(x_j), \eta_\infty(x_j)) \mathbf{1}_{B_{N,i}} =: \gamma_{N,1}(s) + \gamma_{N,2}. \end{aligned} \quad (4.6.17)$$

Denoting by

$$\Delta F_j := F(X_{N,j}(s-), \eta_s(x_j)) - F(X_\infty(x_j), \eta_\infty(x_j)),$$

we have, as  $\langle \mathbf{1}_{B_{N,i}}, \mathbf{1}_{B_{N,i'}} \rangle = \frac{\mathbf{1}_{i=i'}}{N}$  and with  $S_{jj'} := \frac{1}{N} \sum_{i=1}^N \bar{\xi}_{ij} \bar{\xi}_{ij'}$ ,  $\|\gamma_{N,1}(s)\|_2^2 = \frac{1}{N^2 \rho_N^2} \sum_{j,j'=1}^N \Delta F_j \Delta F_{j'} \frac{1}{N} S_{jj'}$ . Define the following quantity  $S_N^{\max} := \sup_{1 \leq j \neq j' \leq N} |S_{jj'}|$ .

The purpose of Lemma 4.A.5 is exactly to control  $S_N^{\max}$ , see in particular (4.A.3). We have

$$\begin{aligned} \|\gamma_{N,1}(s)\|_2^2 &= \left( \frac{1}{N^2 \rho_N^2} \sum_{j \neq j'=1}^N \Delta F_j \Delta F_{j'} \frac{S_{jj'}}{S_N^{\max}} \right) S_N^{\max} + \frac{1}{N^3 \rho_N^2} \sum_{i,j=1}^N \Delta F_j^2 \bar{\xi}_{ij}^2 \\ &\leq S_N^{\max} \left( \frac{1}{N \rho_N^2} \sum_{j=1}^N |\Delta F_j|^2 \right) + \frac{1}{N^2 \rho_N^2} \sum_{j=1}^N \Delta F_j^2. \end{aligned}$$

As  $|\Delta F_j| \leq \|F\|_L (|Y_N(s-)(x_j)| + \delta_s)$ , we obtain as  $s \mapsto \delta_s$  is bounded

$$\|\gamma_{N,1}(s)\|_2^2 \leq C_F \left( \|\tilde{Y}_N(s-)\|_2^2 + 1 \right) \left( \frac{S_N^{\max}}{\rho_N^2} + \frac{1}{N \rho_N^2} \right),$$

hence using (4.6.9) and the fact that  $\|Y_N(s-)\|_2 \leq \|Y_N(s)\|_2 + \frac{C}{N \rho_N}$ ,

$$\|\gamma_{N,1}(s)\|_2^2 \leq C_F (\|Y_N(s)\|_2^4 + 1) \left( \frac{S_N^{\max}}{\rho_N^2} + \frac{1}{N \rho_N^2} \right). \quad (4.6.18)$$

For the second term of (4.6.17), we have

$$\begin{aligned} \|\gamma_{N,2}\|_2^2 &= \frac{1}{N} \sum_{i=1}^N \left( \frac{1}{N \rho_N} \sum_{j=1}^N \bar{\xi}_{ij} F(X_\infty(x_j), \eta_\infty(x_j)) \right)^2 \\ &= \frac{1}{N^3 \rho_N^2} \sum_{i=1}^N \sum_{j,j'=1}^N \bar{\xi}_{ij} \bar{\xi}_{ij'} F(X_\infty(x_j), \eta_\infty(x_j)) F(X_\infty(x_{j'}), \eta_\infty(x_{j'})). \end{aligned}$$

Define

$$\alpha_{i,j,j'} := \frac{F(X_\infty(x_j), \eta_\infty(x_j)) F(X_\infty(x_{j'}), \eta_\infty(x_{j'}))}{\|F(X_\infty, \eta_\infty)\|_\infty^2} \in [0, 1],$$

$$R_k := \sum_{\substack{i,j,j'=1 \\ j \neq j'}}^k \alpha_{i,j,j'} \bar{\xi}_{ij} \bar{\xi}_{ij'} \text{ and}$$

$$\mathcal{F}_k = \sigma(\xi_{ij}, 1 \leq i, j \leq k).$$

We have then  $\|\gamma_{N,2}\|_2^2 = \frac{C_{F,X_\infty}}{N^3 \rho_N^2} \sum_{i,j=1}^N \alpha_{i,j,j} \bar{\xi}_{ij}^2 + \frac{C_{F,X_\infty}}{N^3 \rho_N^2} R_N \leq \frac{C_{F,X_\infty}}{N \rho_N^2} + \frac{C_{F,X_\infty}}{N^3 \rho_N^2} R_N$ .

We show next that  $(R_k)_{1 \leq k \leq N}$  is a  $(\mathcal{F}_k)$ -martingale. Let  $\Delta R_k = R_{k+1} - R_k$ . For



any  $k \geq 1$  (note that  $R_1 = 0$ ), we have

$$\Delta R_k = \sum_{\substack{j,j'=1 \\ j \neq j'}}^k \alpha_{k+1,j,j'} \overline{\xi_{k+1,j}} \overline{\xi_{k+1,j'}} + \sum_{\substack{1 \leq i \leq k+1 \\ 1 \leq j \leq k}} (\alpha_{i,j,k+1} + \alpha_{i,k+1,j'}) \overline{\xi_{i,k+1}} \overline{\xi_{ij}},$$

and thus  $\mathbb{E}[\Delta R_k | \mathcal{F}_k] = 0$  as  $\mathbb{E}[\overline{\xi_{ij}} \overline{\xi_{ij'}} | \mathcal{F}_k] = 0$  if  $j \neq j'$  and at least one of the indexes  $i, j, j'$  is equal to  $k+1$  by independence of the family of random variables  $(\xi_{ij})_{i,j}$ . Moreover, as each  $|\overline{\xi_{i,j}}| \leq 1$  and  $|\alpha_{i,j,k}| \leq 1$ , it gives  $|\Delta R_k| \leq 3k^2 + k$ . Theorem 4.A.2 gives then that

$$\begin{aligned} \mathbb{P}\left(\left|\frac{C_{F,X_\infty}}{N^3 \rho_N^2} R_N\right| \geq x\right) &= \mathbb{P}\left(|R_N| \geq \frac{x N^3 \rho_N^2}{C_{F,X_\infty}}\right) \\ &\leq 2 \exp\left(-\frac{\left(\frac{x N^3 \rho_N^2}{C_{F,X_\infty}}\right)^2}{2 \sum_{k=1}^N (3k^2 + k)^2}\right) = 2 \exp\left(-\frac{x^2 N^6 \rho_N^4}{C_{F,X_\infty}^2 P(N)}\right), \end{aligned}$$

with  $P(N) \underset{N \rightarrow \infty}{\sim} \frac{18}{5} N^5$ . For the choice  $x^2 = \frac{C_{F,X_\infty}^2 P(N)}{N^{6-2\tau} \rho_N^4}$  with  $\tau$  introduced in (4.2.6), that is  $x^2 \propto \frac{1}{N^{1-2\tau} \rho_N^4}$ ,

$$\mathbb{P}\left(\left|\frac{C_{F,X_\infty}}{N^3} R_N\right| \geq \sqrt{\frac{C_{F,X_\infty}^2 P(N)}{N^{6-2\tau} \rho_N^4}}\right) \leq 2 \exp(-N^{2\tau}),$$

which is summable hence by Borel-Cantelli Lemma, there exists  $\mathcal{O} \in \mathcal{F}$  such that  $\mathbb{P}(\mathcal{O}) = 1$  and on  $\mathcal{O}$ , there exists  $\tilde{N} < \infty$  such that if  $N \geq \tilde{N}$ ,  $\left|\frac{C_{F,X_\infty}}{N^3} R_N\right| \leq \sqrt{\frac{C_{F,X_\infty}^2 P(N)}{N^{6-2\tau} \rho_N^4}} \propto \frac{1}{N^{1/2-\tau} \rho_N^2}$ , hence  $\mathbb{P}$ -a.s. for  $N$  large enough

$$\|\gamma_{N,2}\|_2^2 \leq C \left( \frac{1}{N \rho_N^2} + \frac{1}{N^{1-2\tau} \rho_N^4} \right). \quad (4.6.19)$$

Coming back to (4.6.17), combining (4.6.18) and (4.6.19) and a control of  $S_N^{\max}$  from Lemma 4.A.5, we have  $\mathbb{P}$ -a.s. for  $N$  large enough

$$\|\gamma_N(s)\|_2^2 \leq C_F (\|Y_N(s)\|_2^4 + 1) \left( \frac{1}{N^{1/2-\tau} \rho_N^2} + \frac{1}{N \rho_N^2} \right) + C_F \left( \frac{1}{N \rho_N^2} + \frac{1}{N^{1-2\tau} \rho_N^4} \right),$$

hence taking the square root and using (4.6.14),

$$\|\phi_{N,1}(t)\|_2 \leq C_F \int_{t_0}^t e^{-(t-s)\gamma} \|Y_N(s)\|_2^2 ds + G_{N,1},$$

where  $G_{N,1} = C_F \left( \frac{1}{N \rho_N^2} + \frac{1}{N^{1-2\tau} \rho_N^4} + \frac{1}{N^{1/2-\tau} \rho_N^2} \right) \rightarrow 0$  under Hypothesis 4.2.6.

Let us now turn to the sharper control on  $\gamma_N$  defined in (4.6.15) when  $F$  is bounded. Coming back to (4.6.17), we have

$$\begin{aligned}
 \|\gamma_N(s)\|_2^2 &= \int \left( \sum_{i,j=1}^N \frac{1}{N\rho_N} \overline{\xi_{ij}} F(X_{N,j}(s-), \eta_s(x_j)) \mathbf{1}_{B_{N,i}}(x) \right)^2 dx \\
 &= \frac{1}{N} \sum_{i=1}^N \left( \sum_{j=1}^N \frac{1}{N\rho_N} \overline{\xi_{ij}} F(X_{N,j}(s-), \eta_s(x_j)) \right)^2 \\
 &= \frac{1}{N^3 \rho_N^2} \sum_{i,j,k=1}^N \overline{\xi_{ij}} \overline{\xi_{ik}} F(X_{N,j}(s-), \eta_s(x_j)) F(X_{N,k}(s-), \eta_s(x_k)) \\
 &= \left( \frac{\|F\|_\infty}{N\rho_N} \right)^2 \frac{1}{N} \sum_{j,k=1}^N \alpha_{jk} F_j F_k,
 \end{aligned}$$

with  $\alpha_{jk} := \sum_{i=1}^N \overline{\xi_{ij}} \overline{\xi_{ik}}$  and  $F_j := \frac{F(X_{N,j}(s-), \eta_s(x_j))}{\|F\|_\infty}$ . Grothendieck inequality (see Theorem 4.A.1) gives then that there exists  $K > 0$  such that

$$\begin{aligned}
 \|\gamma_N(s)\|_2^2 &\leq K \frac{1}{N} \left( \frac{\|F\|_\infty}{N\rho_N} \right)^2 \sup_{s_j, t_k = \pm 1} \sum_{j,k} \alpha_{jk} s_j t_k \\
 &\leq \frac{C_F}{N^3 \rho_N^2} \sup_{s_j, t_k = \pm 1} \sum_{i,j,k=1}^N \overline{\xi_{ij}} \overline{\xi_{ik}} s_j t_k.
 \end{aligned}$$

Fix some vectors of signs  $s = (s_i)_{1 \leq i \leq N}$  and  $t = (t_j)_{1 \leq j \leq N}$ . Let  $A = (\overline{\xi_{ij}})_{1 \leq i, j \leq N}$ ,

then  $\sum_{i,j,k=1}^N \overline{\xi_{ij}} \overline{\xi_{ik}} s_j t_k = \langle t, A^* A s \rangle$  where  $\langle \cdot, \cdot \rangle$  denotes the scalar product in  $\mathbb{R}^N$  and

$A^*$  the transpose of  $A$ . As for any sign vector  $t$ ,  $\|t\|^2 = \sum_{k=1}^N t_k^2 = N$ , and  $\|A^* A\| = \|A\|_{\text{op}}^2$ , we obtain as  $|\langle t, A^* A s \rangle| \leq \|t\| \|A^* A s\| \leq N \|A\|_{\text{op}}^2$ :

$$\|\gamma_N(s)\|_2^2 \leq \frac{C_F}{N^3 \rho_N^2} N \|A\|_{\text{op}}^2 = \frac{C_F}{N^2 \rho_N^2} \|A\|_{\text{op}}^2.$$

From Theorem 4.A.3, there exist  $C_a$  and  $C_b$  positive constants such that for any  $x \geq C_a$ ,

$$\mathbb{P} \left( \|A\|_{\text{op}} > x\sqrt{N} \right) \leq C_a \exp(-C_b x N).$$

We apply it for  $x = C_a$ , hence, by Borel-Cantelli Lemma as  $\exp(-CN)$  is summable, there exists  $\tilde{\mathcal{O}} \in \mathcal{F}$  such that  $\mathbb{P}(\tilde{\mathcal{O}}) = 1$  and on  $\tilde{\mathcal{O}}$ , there exists  $\tilde{N} < \infty$  such that if  $N \geq \tilde{N}$ ,  $\|A\|_{\text{op}} \leq C_a \sqrt{N}$ . We obtain then that

$$\|\gamma_N(s)\|_2^2 \leq \frac{C_F}{N\rho_N^2}$$

$\mathbb{P}$ -a.s. for  $N$  large enough, which concludes the proof in the bounded case with (4.6.14) as  $\int_{t_0}^t e^{-(t-s)\gamma} ds \leq \frac{1}{\gamma}$ .  $\square$

**Lemma 4.6.5.** *Under Hypothesis 4.2.1, there exists  $C_F > 0$  such that for any  $t > t_0 \geq 0$ ,*

$$\|\phi_{N,2}(t)\|_2 \leq \frac{C_F}{N\rho_N}.$$

*Proof.* Recall the definition of  $\phi_{N,2}(t)$  in (4.6.7) and  $\Theta_{t,i,2}$  in (4.6.2), it directly comes from the Lipschitz continuity of  $F$  and the fact that  $Z_{N,1}, \dots, Z_{N,N}$  do not jump simultaneously.  $\square$

**Lemma 4.6.6.** *Under Hypothesis 4.2.1, for any  $t > t_0 \geq 0$ ,*

$$\|\phi_{N,3}(t)\|_2 \leq C_{F,X_\infty,\eta,W} \int_{t_0}^t e^{-(t-s)\gamma} (\|Y_N(s)\|_2^2 + \delta_s) ds + G_{N,2}, \quad (4.6.20)$$

where  $G_{N,2}$  is explicit in  $N$  and tends to 0 as  $N \rightarrow \infty$ . Moreover, if we suppose  $F$  bounded, we have

$$\|\phi_{N,3}(t)\|_2 \leq C \left( \int_{t_0}^t e^{-(t-s)\gamma} \delta_s ds + \sqrt{R_{N,2}^W} + \frac{1}{N} \right), \quad (4.6.21)$$

with  $R_{N,2}^W$  defined in (4.2.24).

*Proof.* We have, with  $\Theta_{s,i,3}$  defined in (4.6.3),  $\Theta_{s,i,3} \leq e_{s,i,1} + e_{s,i,2} + e_{s,i,3}$  with

$$\begin{aligned} e_{s,i,1} &:= \sum_{j=1}^N \int_{B_{N,j}} (W(x_i, x_j) - W(x_i, y)) \\ &\quad (F(X_N(s, x_j), \eta_s(x_j)) - F(X_\infty(x_j), \eta_s(x_j))) dy \\ e_{s,i,2} &:= \sum_{j=1}^N \int_{B_{N,j}} (W(x_i, x_j) - W(x_i, y)) F(X_\infty(x_j), \eta_s(x_j)) dy \\ e_{s,i,3} &:= \sum_{j=1}^N \int_{B_{N,j}} W(x_i, y) (F(X_N(s, x_j), \eta_s(x_j)) - F(X_N(s, x_j), \eta_s(y))) dy. \end{aligned}$$

We upper-bound each term. We have as  $F$  is Lipschitz continuous that  $e_{s,i,1}$  is upper-bounded by

$$\begin{aligned} &\sum_{j=1}^N |F(X_N(s, x_j), \eta_s(x_j)) - F(X_\infty(x_j), \eta_s(x_j))| \left| \int_{B_{N,j}} (W(x_i, x_j) - W(x_i, y)) dy \right| \\ &\leq \sum_{j=1}^N \|F\|_L |Y_N(s)(x_j)| \left| \int_{B_{N,j}} (W(x_i, x_j) - W(x_i, y)) dy \right|, \\ &\leq C_F \left( \sum_{j=1}^N \left| \int_{B_{N,j}} (W(x_i, x_j) - W(x_i, y)) dy \right| \right)^{\frac{1}{2}} \\ &\quad \left( \sum_{j=1}^N |Y_N(s)(x_j)|^2 \left| \int_{B_{N,j}} (W(x_i, x_j) - W(x_i, y)) dy \right| \right)^{\frac{1}{2}}, \end{aligned}$$

using discrete Jensen's inequality. We have  $N \left| \int_{B_{N,j}} (W(x_i, x_j) - W(x_i, y)) dy \right| \leq C$  as  $W$  is bounded, hence

$$\begin{aligned} e_{s,i,1} &\leq C_{F,W} \left( \sum_{j=1}^N \left| \int_{B_{N,j}} (W(x_i, x_j) - W(x_i, y)) dy \right| \right)^{\frac{1}{2}} \left( \frac{1}{N} \sum_{j=1}^N |Y_N(s)(x_j)|^2 \right)^{\frac{1}{2}} \\ &\leq C_{F,W} \left( \sum_{j=1}^N \left| \int_{B_{N,j}} (W(x_i, x_j) - W(x_i, y)) dy \right| \right)^{\frac{1}{2}} \left\| \tilde{Y}_N(s) \right\|_2. \end{aligned}$$

We have then

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N e_{s,i,1}^2 &\leq \frac{C_{F,W}}{N} \sum_{i=1}^N \left( \sum_{j=1}^N \left| \int_{B_{N,j}} (W(x_i, x_j) - W(x_i, y)) dy \right| \right) \left\| \tilde{Y}_N(s) \right\|_2^2 \\ &\leq C_{F,W} R_{N,1}^W \left\| \tilde{Y}_N(s) \right\|_2^2, \end{aligned}$$

where  $R_{N,1}^W$  is defined in (4.2.24).

For the second term, we have as  $x \mapsto \sup_s F(X_\infty(x), \eta_s(x))$  is bounded

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N e_{s,i,2}^2 &= \frac{1}{N} \sum_{i=1}^N \left( \sum_{j=1}^N \int_{B_{N,j}} (W(x_i, x_j) - W(x_i, y)) F(X_\infty(x_j), \eta_s(x_j)) dy \right)^2 \\ &\leq \frac{C_F}{N} \sum_{i=1}^N \sum_{j=1}^N \int_{B_{N,j}} |W(x_i, x_j) - W(x_i, y)|^2 dy \leq C_F R_{N,2}^W, \end{aligned}$$

where  $R_{N,2}^W$  is defined in (4.2.24).

For the third term, as  $F$  is Lipschitz continuous

$$\begin{aligned} e_{s,i,3} &\leq \sum_{j=1}^N \int_{B_{N,j}} W(x_i, y) \|F\|_L |\eta_s(x_j) - \eta_s(y)| dy \\ &\leq \sum_{j=1}^N \int_{B_{N,j}} W(x_i, y) \|F\|_L (|\eta_s(x_j) - \eta_\infty(x_j)| + |\eta_\infty(x_j) - \eta_\infty(y)|) dy \\ &\leq C_{F,X,W} \left( \delta_s + \frac{1}{N} \right). \end{aligned}$$

We obtain then with (4.6.9)

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \Theta_{s,i,2}^2 &\leq \frac{3}{N} \sum_{i=1}^N (e_{s,i,1}^2 + e_{s,i,2}^2 + e_{s,i,3}^2) \\ &\leq C_{F,X_\infty,X,W} \left( R_{N,1}^W (1 + \|Y_N(s)\|_2^4) + R_{N,2}^W + \delta_s^2 + \frac{1}{N^2} \right). \end{aligned}$$

With (4.6.7) and Proposition 4.2.10,  $\|\phi_{N,3}(t)\|_2 \leq \int_{t_0}^t e^{-(t-s)\gamma} \left\| \sum_{i=1}^N \Theta_{s,i,3} \mathbf{1}_{B_{N,i}} \right\|_2 ds$ , and as  $\left\| \sum_{i=1}^N \Theta_{s,i,3} \mathbf{1}_{B_{N,i}} \right\|_2^2 = \frac{1}{N} \sum_{i=1}^N \Theta_{s,i,3}^2$ , the result follows with

$$G_{N,2} = \sqrt{R_{N,1}^W + R_{N,2}^W} + \frac{1}{N},$$

and Lemma 4.6.2.

When  $F$  is bounded, similarly we show that

$$\frac{1}{N} \sum_{i=1}^N \Theta_{s,i,3}^2 \leq C_{F,X_\infty,\eta,W} \left( R_{N,2}^W + \delta_s^2 + \frac{1}{N^2} \right),$$

hence the result.  $\square$

**Lemma 4.6.7.** *Under Hypothesis 4.2.1, for any  $t > t_0 \geq 0$ ,*

$$\|\phi_{N,4}(t)\|_2 \leq C_{F,X_\infty,W} \int_{t_0}^t e^{-(t-s)\gamma} (\|Y_N(s)\|_2^2 + \delta_s) ds + G_{N,3}, \quad (4.6.22)$$

where  $G_{N,3}$  is explicit in  $N$  and tends to 0 as  $N \rightarrow \infty$ . Moreover, if we suppose  $F$  bounded, we have

$$\sup_{t \geq 0} \|\phi_{N,4}(t)\|_2 \leq \sqrt{S_N^W}, \quad (4.6.23)$$

where  $S_N^W$  is defined in (4.2.25).

*Proof.* We have

$$\begin{aligned} \left\| \sum_{i=1}^N \Theta_{s,i,4} \mathbf{1}_{B_{N,i}} \right\|_2^2 &= \int_I \left( \sum_{i=1}^N \Theta_{s,i,4}(x) \mathbf{1}_{B_{N,i}}(x) \right)^2 dx \\ &= \sum_{i=1}^N \int_{B_{N,i}} \left( \int_I (W(x_i, y) - W(x, y)) F(X_N(s, y), \eta_s(y)) dy \right)^2 dx \\ &\leq \sum_{i=1}^N \int_{B_{N,i}} \left( \int_I (W(x_i, y) - W(x, y))^2 dy \right) \left( \int_I (F(X_N(s, y), \eta_s(y)))^2 dy \right) dx, \end{aligned}$$

with Cauchy Schwarz's inequality. We can recognize  $S_N^W$  defined in (4.2.25), and we have that, as  $F$  is Lipschitz continuous and  $y \mapsto F(X_\infty(y), \eta_\infty(y))$  is bounded,

$$\begin{aligned} &\int_I F(X_N(s, y), \eta_s(y))^2 dy \\ &\leq \int_I (F(X_N(s, y), \eta_s(y)) - F(X_\infty(y), \eta_s(y)))^2 dy + \int_I F(X_\infty(y), \eta_s(y))^2 dy \\ &\leq \|F\|_L^2 \int_I Y_N(s)(y)^2 dy + \|F(X_\infty, \eta_\infty)\|_\infty^2 \leq C_{F,W} (\|Y_N(s)\|_2^2 + 1) \\ &\leq C_{F,W} (\|Y_N(s)\|_2^4 + 1). \end{aligned}$$

As done before, (4.6.7) and Proposition 4.2.10 give that

$$\begin{aligned} \|\phi_{N,4}(t)\|_2 &\leq \int_{t_0}^t e^{-(t-s)\gamma} \left\| \sum_{i=1}^N \Theta_{s,i,4} \mathbf{1}_{B_{N,i}} \right\|_2 ds \text{ and} \\ \left\| \sum_{i=1}^N \Theta_{s,i,4} \mathbf{1}_{B_{N,i}} \right\|_2^2 &\leq C_{F,W} S_N^W (\|Y_N(s)\|_2^4 + 1), \end{aligned}$$

hence the result with  $G_{N,3} = C_{F,W} \sqrt{R_{N,3}}$  and Lemma 4.6.2. When  $F$  is bounded,

we directly have  $\left\| \sum_{i=1}^N \Theta_{s,i,4} \mathbf{1}_{B_{N,i}} \right\|_2^2 \leq S_N^W$  hence (4.6.23) as  $\int_{t_0}^t e^{-(t-s)\gamma} ds \leq \frac{1}{\gamma}$ .  $\square$

### 4.6.3 Proof of Proposition 4.4.3

Proposition 4.4.3 is then a direct consequence of (4.6.6) and (4.6.7), of the controls given by Lemmas 4.6.3, 4.6.4, 4.6.5, 4.6.6 and 4.6.7, with  $G_N = G_{N,1} + G_{N,2} + G_{N,3}$ , and of Lemma 4.6.2 to have  $G_N \rightarrow 0$ .

## 4.7 About the finite time behavior (proofs)

In this section, we prove Proposition 4.2.11.

### 4.7.1 Main technical results

In the following, we denote by  $\widehat{Y}_N(t) := X_N(t) - X_t$ .

*Proof of Proposition 4.2.11.* Let  $t \leq T$ . Recall the definition of  $X_N(t)$  in (4.1.11) and  $X_t$  in (4.1.12). Proceeding exactly as in the proof of Proposition 4.4.1, and recalling the definition of  $M_N(t)$  in (4.4.5), we have

$$\begin{aligned} d\widehat{Y}_N(t) + \alpha\widehat{Y}_N(t)dt - dM_N(t) &= \sum_{i,j=1}^N \mathbf{1}_{B_{N,i}} \frac{w_{ij}}{N} F(X_{N,j}(t-), \eta_t(x_j)) dt - T_W F(X_t, \eta_t) dt \\ &= \sum_{k=1}^4 \sum_{i=1}^N \Theta_{t,i,k} \mathbf{1}_{B_{N,i}} dt + T_W (F(X_{N,j}(t-), \eta_t(x_j)) - F(X_t, \eta_t)) dt, \end{aligned}$$

with the notations introduced in (4.6.1) - (4.6.4). It gives then, as  $\widehat{Y}_N(0) = 0$ ,

$$\widehat{Y}_N(t) = \int_0^t e^{-\alpha(t-s)} \widehat{r}_N(s) ds + \int_0^t e^{-\alpha(t-s)} dM_N(s) =: \widehat{\phi}_N(t) + \widehat{\zeta}_N(t)$$

with

$$\widehat{r}_N(t) = \sum_{k=1}^4 \sum_{i=1}^N \Theta_{t,i,k} \mathbf{1}_{B_{N,i}} + T_W (F(X_N(t-), \eta_t) - F(X_t, \eta_t)).$$

Note that we obtain a similar expression as for  $Y_N$  in Proposition 4.4.1, but with  $e^{-\alpha t}$  instead of the semi-group  $e^{t\mathcal{L}}$ . We use then the two following results, similar to Propositions 4.4.2 and 4.4.3.

**Proposition 4.7.1.** *Let  $T > 0$ . Under Hypothesis 4.2.1, there exists a constant  $C = C(T, F, \|\eta\|_\infty) > 0$  such that  $\mathbb{P}$ -almost surely for  $N$  large enough:*

$$\mathbf{E} \left[ \sup_{s \leq T} \|\widehat{\zeta}_N(s)\|_2 \right] \leq \frac{C}{\sqrt{N\rho_N}}.$$

**Proposition 4.7.2.** *Under Hypotheses 4.2.1 and 4.2.6, for any  $t > 0$ ,*

$$\|\widehat{\phi}_N(t)\|_2 \leq C \left( \int_0^t e^{-\alpha(t-s)} \|\widehat{Y}_N(s)\|_2 ds + \widehat{G}_N \right), \quad (4.7.1)$$

where  $\widehat{G}_N$  is an explicit quantity to be found in the proof that tends to 0 as  $N \rightarrow \infty$ .

Their proofs are postponed to the following subsection. Hence we obtain

$$\left\| \widehat{Y}_N(t) \right\|_2 \leq C \left( \widehat{G}_N + \left\| \widehat{\zeta}_N(t) \right\|_2 + \int_0^t e^{-\alpha(t-s)} \left\| \widehat{Y}_N(s) \right\|_2 ds \right),$$

which gives with Grönwall lemma

$$\sup_{t \leq T} \left\| \widehat{Y}_N(t) \right\|_2 \leq C \left( \widehat{G}_N + \sup_{t \leq T} \left\| \widehat{\zeta}_N(t) \right\|_2 \right).$$

With Proposition 4.7.1, it leads to

$$\mathbf{E} \left[ \sup_{t \leq T} \left\| \widehat{Y}_N(t) \right\|_2 \right] \leq C \left( \widehat{G}_N + \frac{1}{\sqrt{N\rho_N}} \right),$$

hence the result (4.2.16) as (4.2.6) implies  $\frac{1}{\sqrt{N\rho_N}} \rightarrow 0$  and  $\widehat{G}_N \rightarrow 0$ .  $\square$

#### 4.7.2 Proofs of Propositions 4.7.1 and 4.7.2

*Proof of Proposition 4.7.1.* We do as for Proposition 4.4.2, and apply Itô's formula on

$$\widehat{\zeta}_N(t) = \sum_{j=1}^N \int_0^t \int_0^\infty e^{-\alpha(t-s)} \chi_j(s, z) \tilde{\pi}_j(ds, dz).$$

The term  $I_0(t)$  in (4.5.3) becomes  $-\alpha \int_0^t \left\| \widehat{\zeta}_N(s) \right\|_2 ds$  which is still non-positive. About  $I_1(t)$  and  $I_2(t)$ , the proof remains the same aside from the fact that we now consider  $\widehat{\zeta}_N$  instead of  $\zeta_N$ .  $\square$

To prove 4.7.2, we introduce an auxilliary quantity as in Lemma 4.6.1.

**Lemma 4.7.3.** *Let  $\bar{Y}_N(s)(v) := \widehat{Y}_N(s) \left( \frac{[Nv]}{N} \right)$ . Then for any  $T \geq 0$*

$$\sup_{0 \leq s \leq T} \left\| \bar{Y}_N(s) - \widehat{Y}_N(s) \right\|_2 \xrightarrow{N \rightarrow \infty} 0. \quad (4.7.2)$$

*Proof.* It plays the role of  $\tilde{Y}_N(s)$  introduced in Lemma 4.6.1. Similarly to what has been done before, we have

$$\begin{aligned} \left\| \widehat{Y}_N(s) - \bar{Y}_N(s) \right\|_2^2 &= \sum_{j=1}^N \int_{B_{N,j}} \left( \widehat{Y}_N(s)(y) - \bar{Y}_N(s)(y) \right)^2 dy \\ &= \sum_{j=1}^N \int_{B_{N,j}} \left( X_s(x_j) - X_s(y) \right)^2 dy, \end{aligned}$$

which tends to 0 by uniform continuity of  $X$  on  $[0, T] \times I$ . It still holds under the hypotheses of Section 4.2.3.4 by decomposing the sum on each interval  $C_k$ .  $\square$

*Proof of Proposition 4.7.2.* We divide  $\widehat{\phi}$  as in (4.6.7) and study each contribution. About  $\widehat{\phi}_{N,0}(t) := \int_0^t e^{-\alpha(t-s)} T_W (F(X_N(s), \eta_s) - F(X_s, \eta_s)) ds$ , we have

$$\begin{aligned} \|T_W (F(X_N(s), \eta_s) - F(X_s, \eta_s))\|_2^2 &\leq C_{W,F} \left( \int_I \|F\|_L |X_N(s)(y) - X_s(y)| dy \right)^2 \\ &\leq C_{W,F} \left\| \widehat{Y}_N(s) \right\|_2^2, \end{aligned}$$

which gives

$$\left\| \widehat{\phi}_{N,0}(t) \right\|_2 \leq C_{W,F} \int_0^t e^{-\alpha(t-s)} \left\| \widehat{Y}_N(s) \right\|_2 ds.$$

About  $\widehat{\phi}_{N,1}(t) := \int_0^t e^{-\alpha(t-s)} \sum_{i=1}^N \frac{\Theta_{s,i,1}}{N} \mathbf{1}_{B_{N,i}} ds$ , we do as in Lemma 4.6.4: we insert the terms  $F(X_s(x_j), \eta_s(x_j))$  (instead of the terms  $F(X_\infty(x_j), \eta_\infty(x_j))$  in (4.6.17)), that is

$$\begin{aligned} \gamma_N(s) &\leq \sum_{i,j=1}^N \frac{1}{N} \kappa_{N,i} \overline{\xi}_{ij} (F(X_{N,j}(s-), \eta_s(x_j)) - F(X_s(x_j), \eta_s(x_j))) \mathbf{1}_{B_{N,i}} \\ &\quad + \sum_{i,j=1}^N \frac{1}{N} \kappa_{N,i} \overline{\xi}_{ij} F(X_s(x_j), \eta_s(x_j)) \mathbf{1}_{B_{N,i}} =: \widehat{\gamma}_{N,1}(s) + \widehat{\gamma}_{N,2}(s). \end{aligned}$$

The treatment of  $\widehat{\gamma}_{N,1}$  is similar of  $\gamma_{N,1}$ : we make  $\overline{Y}_N(s-)$  appear instead of  $\check{Y}_N$  and obtain  $\|\widehat{\gamma}_{N,1}(s-)\|_2^2 \leq C_F \left( \left\| \widehat{Y}_N(s) \right\|_2^2 + 1 \right) \left( \frac{S_N^{\max}}{\rho_N^2} + \frac{1}{N\rho_N^2} \right)$  with (4.7.2). About  $\widehat{\gamma}_{N,2}$ , we do as  $\gamma_{N,2}$  as  $\sup_{t \in [0, T], x \in I} F(X_t(x), \eta_t(x)) < \infty$  and obtain that  $\mathbb{P}$ -almost surely if  $N$  is large enough,  $\|\widehat{\gamma}_{N,2}\|_2^2 \leq C \left( \frac{1}{N\rho_N^2} + \frac{1}{N^{1-2\tau}\rho_N^4} \right)$ . We have then that,  $\mathbb{P}$ -almost surely if  $N$  is large enough,

$$\left\| \widehat{\phi}_{N,1}(t) \right\|_2 \leq C_F \int_{t_0}^t e^{-\alpha(t-s)} \left\| \widehat{Y}_N(s) \right\|_2 ds + G_{N,1},$$

where  $G_{N,1} \rightarrow 0$ .

About  $\widehat{\phi}_{N,2}(t)$ , we proceed as Lemma 4.6.5 to show that  $\|\widehat{\phi}_{N,2}(t)\|_2 \leq \frac{C_F}{N\rho_N}$ .

About  $\widehat{\phi}_{N,k}(t) := \int_0^t e^{-\alpha(t-s)} \sum_{i=1}^N \frac{\Theta_{s,i,k}}{N} \mathbf{1}_{B_{N,i}} ds$  for  $k \in \{3, 4\}$ , we proceed similarly, doing as in 4.6.6 and 4.6.7 but instead of inserting the terms  $F(X_\infty(x_j), \eta_\infty(x_j))$  we insert the terms  $F(X_s(x_j), \eta_s(x_j))$ : then there is no  $\delta_s$  terms. We obtain then

$$\begin{aligned} \|\widehat{\phi}_{N,3}(t)\|_2 &\leq C \int_{t_0}^t e^{-\alpha(t-s)} \left\| \widehat{Y}_N(s) \right\|_2 ds + G_{N,2}, \text{ and} \\ \|\widehat{\phi}_{N,4}(t)\|_2 &\leq C \int_{t_0}^t e^{-\alpha(t-s)} \left\| Y_N(s) \right\|_2 ds + G_{N,3}, \end{aligned}$$

where both  $G_{N,2}$  and  $G_{N,3}$  tends to 0. Note that we can obtain better bounds when  $F$  is bounded. By putting all the terms  $\widehat{\phi}_{N,k}$  together, we get (4.7.1).  $\square$



## 4.A Auxiliary results

### 4.A.1 Concentration results

**Theorem 4.A.1** (Grothendieck’s inequality as in [Coppini \(2022\)](#)). *Let  $\{a_{ij}\}_{i,j=1,\dots,n}$  be a  $n \times n$  real matrix such that for all  $s_i, t_j \in \{-1, 1\}$*

$$\sum_{i,j=1}^n a_{ij} s_i t_j \leq 1.$$

*Then, there exists a constant  $K_R > 0$ , such that for every Hilbert space  $(H, \langle \cdot, \cdot \rangle_H)$  and for all  $S_i$  and  $T_j$  in the unit ball of  $H$*

$$\sum_{i,j=1}^n a_{ij} \langle S_i, T_j \rangle_H \leq K_R.$$

**Theorem 4.A.2** (Azuma–Hoeffding inequality). *Let  $(M_n)$  be a martingale with  $M_0 = 0$ . Assume that for all  $1 \leq k \leq n$ ,  $|\Delta M_k| \leq c_k$  a.s. for some constants  $(c_k)$ . Then for all  $x \geq 0$*

$$\mathbb{P}(|M_n| \geq x) \leq 2 \exp\left(-\frac{x^2}{2 \sum_{k=1}^n c_k^2}\right). \quad (4.A.1)$$

**Theorem 4.A.3** (Upper tail estimate for iid ensembles, Corollary 2.3.5 of [Tao \(2012\)](#)). *Suppose that  $M = (m_{ij})_{1 \leq i,j \leq n}$ , where  $n$  is a (large) integer and the  $m_{ij}$  are independent centered random variables uniformly bounded in magnitude by 1. Then there exist absolute constants  $C, c > 0$  such that*

$$\mathbb{P}(\|M\|_{op} > x\sqrt{n}) \leq C \exp(-c x n)$$

for any  $x \geq C$ .

**Lemma 4.A.4.** *Under Hypothesis 4.2.6, we have  $\mathbb{P}$ -almost surely if  $N$  is large enough:*

$$\sup_{1 \leq j \leq N} \left( \sum_{i=1}^N \frac{\xi_{ij}^{(N)}}{N \rho_N} \right) \leq 2, \quad \sup_{1 \leq i \leq N} \left( \sum_{j=1}^N \frac{\xi_{ij}^{(N)}}{N \rho_N} \right) \leq 2. \quad (4.A.2)$$

*Proof.* It is a direct consequence of a previous work ([Agathe-Nerine, 2022](#), Corollary 8.2), that is Corollary 3.8.2 in the present thesis, in the case  $w_N = \rho_N$ ,  $\kappa_N = \frac{1}{\rho_N}$ ,  $W_N(x_i, x_j) = \rho_N W(x_i, x_j)$  with  $W$  bounded.  $\square$

**Lemma 4.A.5.** *Let  $N \geq 1$ , for  $j \neq j'$  in  $\llbracket 1, N \rrbracket$ , let  $S_{jj'} := \frac{1}{N} \sum_{i=1}^N \overline{\xi_{ij}} \overline{\xi_{ij'}}$  with  $\xi$  defined in Definition 4.1.1, and  $S_N^{max} := \sup_{1 \leq j \neq j' \leq N} |S_{jj'}|$ . Then, under Hypothesis 4.2.6,  $\mathbb{P}$ -a.s.*

$$\limsup_{N \rightarrow \infty} S_N^{max} \leq N^{\tau - \frac{1}{2}} \quad (4.A.3)$$

where  $\tau \in (0, \frac{1}{2})$  comes from Hypothesis 4.2.6.

*Proof.* When  $j$  and  $j'$  are fixed and  $j \neq j'$ ,  $(X_i := \overline{\xi_{ij}} \overline{\xi_{ij'}})_{1 \leq i \leq N}$  is a family of independent random variables with  $|X_i| \leq 1$ ,  $\mathbf{E}[X_i] = 0$  and  $\mathbf{E}[X_i^2] \leq 1$ . Bernstein's inequality gives then for any  $t > 0$

$$\mathbf{P} \left( \left| \sum_{i=1}^N \overline{\xi_{ij}} \overline{\xi_{ij'}} \right| > t \right) \leq 2 \exp \left( -\frac{1}{2} \frac{t^2}{N + \frac{t}{3}} \right)$$

hence for the choice  $t = N^{\frac{1}{2}+\tau}$  with  $\tau \in (0, \frac{1}{2})$ ,

$$\mathbf{P} \left( \left| \sum_{i=1}^N \overline{\xi_{ij}} \overline{\xi_{ij'}} \right| > N^{\frac{1}{2}+\tau} \right) \leq 2 \exp \left( -\frac{1}{2} \frac{N^{2\tau}}{1 + \frac{1}{3}N^{-\frac{1}{2}+\tau}} \right) \leq 2 \exp \left( -\frac{1}{4}N^{2\tau} \right)$$

as  $1 + \frac{1}{3}N^{-\frac{1}{2}+\tau} \leq 2$ . With an union bound

$$\mathbf{P} \left( \sup_{j \neq j'} |S_{jj'}| > \frac{1}{N^{\frac{1}{2}-\tau}} \right) \leq 2N^2 \exp \left( -\frac{1}{4}N^{2\tau} \right).$$

We apply then Borel Cantelli's lemma and obtain (4.A.3).  $\square$

**Lemma 4.A.6.** *Fix  $N > 1$  and  $(Y_l)_{l=1, \dots, n}$  real valued random variables defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . Suppose that there exists  $\nu > 0$  such that, almost surely, for all  $l = 1, \dots, n-1$ ,  $Y_l \leq 1$ ,  $\mathbb{E}[Y_{l+1} | Y_l] = 0$  and  $\mathbb{E}[Y_{l+1}^2 | Y_l] \leq \nu$ . Then*

$$\mathbb{P} \left( n^{-1}(Y_1 + \dots + Y_n) \geq x \right) \leq \exp \left( -n \frac{x^2}{2\nu} B \left( \frac{x}{\nu} \right) \right)$$

for all  $x \geq 0$ , where

$$B(u) := u^{-2} ((1+u) \log(1+u) - u). \quad (4.A.4)$$

*Proof.* A direct application of (Dembo & Zeitouni, 1998, Corollary 2.4.7) gives that

$$\mathbb{P} \left( n^{-1}(Y_1 + \dots + Y_n) \geq x \right) \leq \exp \left( -nH \left( \frac{x+\nu}{1+\nu} \middle| \frac{\nu}{1+\nu} \right) \right),$$

where  $H(p|q) := p \log(p/q) + (1-p) \log((1-p)/(1-q))$  for  $p, q \in [0, 1]$ . Then, the inequality  $H \left( \frac{x+\nu}{1+\nu} \middle| \frac{\nu}{1+\nu} \right) \geq \frac{x^2}{2\nu} B \left( \frac{x}{\nu} \right)$  (see (Dembo & Zeitouni, 1998, Exercise 2.4.21)) gives the result.  $\square$

**Corollary 4.A.7.** *Let  $(Z_{ij})_{i,j}$  be a family of independent Bernoulli variables, with  $\mathbb{E}[Z_{ij}] = m_{ij}$ . Let  $(\beta_{ij})_{i,j}$  be a sequence such that for any  $i, j$ ,  $\beta_{i,j} \in (0, 1]$ . Then, for all  $x \geq 0$*

$$\mathbb{P} \left( \frac{1}{N^2} \sum_{i,j=1}^N \beta_{ij} ((Z_{ij} - m_{ij})^2 - \mathbb{E}(Z_{ij} - m_{ij})^2) \geq x \right) \leq \exp \left( -\frac{N^2 x^2}{2} B(x) \right).$$

*Proof.* Fix a bijection  $\phi_N : \llbracket 1, N^2 \rrbracket \rightarrow \llbracket 1, N \rrbracket \times \llbracket 1, N \rrbracket$ . For any  $k \in \llbracket 1, N^2 \rrbracket$  and  $(i, j) = \phi_N(k)$ , let  $R_k = \beta_{ij} ((Z_{ij} - m_{ij})^2 - \mathbb{E}(Z_{ij} - m_{ij})^2)$ . As the  $(m_{ij})_{i,j}$  are independent, the family of random variables  $(R_k)_{1 \leq k \leq N^2}$  is also independent. As  $R_k \leq 1$  a.s.,  $\mathbb{E}[R_{k+1} | R_k] = 0$  and  $\mathbb{E}[R_{k+1}^2 | R_k] \leq 1$ , Lemma 4.A.6 implies that for any  $x \geq 0$ ,

$$\mathbb{P} \left( \frac{1}{N^2} \sum_{k=1}^{N^2} R_k \geq x \right) \leq \exp \left( -\frac{N^2 x^2}{2} B(x) \right)$$

where  $B$  is defined in (4.A.4).  $\square$

### 4.A.2 Other technical results

**Lemma 4.A.8.** *Let  $K$  be a kernel from  $I^2 \rightarrow \mathbb{R}_+$  such that  $\sup_{x \in I} \int_I K(x, y)^2 dy < \infty$ . Let  $T_K : g \mapsto T_K g := (x \rightarrow \int_I K(x, y) dy)$  be the operator associated to  $K$ , that can be defined from  $L^2(I) \rightarrow L^2(I)$  and from  $L^\infty(I) \rightarrow L^\infty(I)$ . We assume that  $T_K^2 : L^2(I) \rightarrow L^2(I)$  is compact. Then*

$$r_2(T_K) = r_\infty(T_K).$$

*Proof.* First note that for both  $r = r_2$  and  $r = r_\infty$ , we have, for all  $p \geq 1$ ,  $r(T_K^p)^{\frac{1}{p}} = \left( \lim_{n \rightarrow \infty} \|T_K^{pn}\|^{\frac{1}{n}} \right)^{\frac{1}{p}} = \lim_{n \rightarrow \infty} \|T_K^{pn}\|^{\frac{1}{pn}} = r(T_K)$ , so that  $r(T_K^p) = r(T_K)^p$ . Hence  $r_2(T_K^2) = r_\infty(T_K^2)$  gives  $r_2(T_K) = r_\infty(T_K)$ .

Denote by  $\sigma_\infty(T_K^2)$  and  $\sigma_2(T_K^2)$  the corresponding spectrum of  $T_K^2$  (in  $L^\infty(I)$  and  $L^2(I)$  respectively). Let us prove that  $r_2(T_K^2) = r_\infty(T_K^2)$  by proving  $\sigma_\infty(T_K^2) = \sigma_2(T_K^2)$ . To do so, first note that  $T_K^2 : L^\infty(I) \rightarrow L^\infty(I)$  is compact: consider  $(f_n)_n$  a bounded sequence of  $L^\infty(I)$ . It is then also bounded in  $L^2(I)$ , and as  $T_K : L^2(I) \rightarrow L^2(I)$  is compact, there exists a subsequence  $(f_{\phi(n)})$  such that  $T_K f_{\phi(n)}$  converges in  $L^2(I)$  to a certain  $g$ . Then for any  $x \in I$ ,

$$|T_K^2 f_{\phi(n)} - T_K^2 g|(x) \leq \int_I K(x, y) |T_K f_{\phi(n)}(y) - g(y)| dy \leq C_K \|T_K f_{\phi(n)} - g\|_2 \xrightarrow{n \rightarrow \infty} 0,$$

thus  $T_K^2 : L^\infty(I) \rightarrow L^\infty(I)$  is compact. Now we prove that  $\sigma_\infty(T_K^2) = \sigma_2(T_K^2)$ : let  $\mu \in \sigma_2(T_K^2) \setminus \{0\}$ , there exists  $g \in L^2(I)$  such that  $\mu g = T_K^2 g$ . As

$$|T_K^2 g(x)| = \left| \int_I K(x, y) \int_I K(y, z) g(z) \nu(dz) \nu(dy) \right| \leq C_K \|g\|_2 < \infty,$$

$g = \frac{1}{\mu} T_K^2 g \in L^\infty(I)$  and  $\mu \in \sigma_\infty(T_K^2)$ . Conversely, let  $\mu \in \sigma_\infty(T_K^2) \setminus \{0\}$ , there exists  $g \in L^\infty(I)$  such that  $\mu g = T_K^2 g$ . As  $L^\infty(I) \subset L^2(I)$ ,  $\mu \in \sigma_2(T_K^2)$ . Hence  $r_2(T_K^2) = r_\infty(T_K^2)$  and (4.2.2) follows.  $\square$

**Lemma 4.A.9** (Quadratic Grönwall's lemma). *Let  $f$  be a non-negative function piecewise continuous with finite number of distinct jumps of size inferior to  $\theta$  on  $[t_0, T]$ , let  $g$  be a non-negative continuous function and  $h \in L^1([t_0, T])$ . For any  $t \in [t_0, T]$ , assume  $f$  satisfies*

$$f(t) \leq f(t_0) + g(t) + \int_{t_0}^t h(t-s) f(s)^2 ds.$$

*Then, for  $\delta < \frac{1}{9\|h\|_1}$ , if  $\theta \leq \frac{\delta}{2}$  and if  $\sup_{t \in [t_0, T]} g(t) \leq \delta$ , we have*

$$\sup_{t \in [t_0, T]} f(t) \leq f(t_0) + 3\delta.$$

*Proof.* Let  $A = \{t \in [t_0, T], f(t) > f(t_0) + 3\delta\}$ , suppose  $A \neq \emptyset$ . Let  $t^* = \inf\{t \in [t_0, T], f(t) > f(t_0) + 3\delta\}$ . If there is no jump at  $t_0$ , by the initial conditions  $t^* > t_0$ , and if there is a jump,  $f(t_0^+) \leq f(t_0) + \frac{\delta}{2}$  hence we also have  $t^* > t_0$ . Moreover,

for all  $t \in [t_0, t^*)$ ,  $f(t) \leq f(t_0) + \delta + 9\delta^2 \int_{t_0}^t h(t-s)ds \leq f(t_0) + 2\delta$ . If there is a jump at  $t^*$ , it is of amplitude  $\theta \leq \frac{\delta}{2}$  hence  $f(t^*) \leq f(t_0) + \frac{5\delta}{2} < f(t_0) + 3\delta$  which is a contradiction. If there is no jump at  $t^*$ , by local continuity we have  $f(t^*) \leq f(t_0) + \delta + 9\delta^2 \int_{t_0}^{t^*} h(t-s)ds \leq f(t_0) + 2\delta$  which is also a contradiction. We conclude then that  $\sup_{t \in [t_0, T]} f(t) \leq f(t_0) + 3\delta$ .  $\square$



# Chapter 5

## Stability of wandering bumps for Hawkes processes interacting on the circle

*This chapter consists in a modified version of my article [Agathe-Nerine \(2023b\)](#) (expanded with some computations), with minor wording and notation changes for harmonization purposes. It has been submitted.*

In this Chapter, we study our main model (2.3.7) in a particular context. The synaptic function  $f$  is fixed as a sigmoid, the interaction terms  $(w_{ij}^{(N)})$  come from a cosine kernel and the memory kernel  $h$  is chosen as a decreasing exponential  $h(t) = e^{-t}$ . There is no spontaneous activity ( $\mu_t = 0$ ) but we consider an initial activity  $v_0(x) = \rho(x)$  that leaks at rate 1. These choices were made to exhibit particular phenomena in the large population and time limit.

### Abstract

We consider a population of Hawkes processes modeling the activity of  $N$  interacting neurons. The neurons are regularly positioned on the circle  $[-\pi, \pi]$ , and the connectivity between neurons is given by a cosine kernel. The firing rate function is a sigmoid. The large population limit admits a locally stable manifold of stationary solutions. The main result of the paper concerns the long-time proximity of the synaptic voltage of the population to this manifold in polynomial times in  $N$ . We show in particular that the phase of the voltage along this manifold converges towards a Brownian motion on a time scale of order  $N$ .

### Contents

---

5.1	Introduction . . . . .	174
5.1.1	Hawkes Processes and Neural Field Equation . . . . .	174
5.1.2	Notations and definition . . . . .	178
5.2	Stability of wandering bumps for interacting Hawkes processes	180
5.2.1	Main results . . . . .	180

5.2.2	Link with the literature . . . . .	184
5.2.3	Strategy of proof of the long time behavior . . . . .	189
5.3	Stationary solutions (proofs) . . . . .	<b>191</b>
5.3.1	Stability . . . . .	192
5.3.2	Projections on the manifold . . . . .	193
5.4	Long time behavior (proofs) . . . . .	<b>198</b>
5.4.1	Main structure of the proof of Theorem 5.2.9 . . . . .	198
5.4.2	About the mild formulation . . . . .	203
5.4.3	About the initialisation . . . . .	204
5.5	Fluctuations on the manifold (proofs) . . . . .	<b>208</b>
5.5.1	Main structure of the proof of Theorem 5.2.11 . . . . .	208
5.5.2	About the decomposition of Proposition 5.5.2 . . . . .	210
5.5.3	Control of the terms of the decomposition . . . . .	211
5.A	On the stationary solutions to the Neural Field Equation . . . . .	<b>221</b>
5.A.1	When $f$ is the Heaviside function . . . . .	221
5.A.2	When $f$ is a sigmoid . . . . .	222
5.B	Some computations . . . . .	<b>224</b>
5.B.1	Control of the noise perturbation . . . . .	224
5.B.2	Analysis complements . . . . .	226
5.B.3	Auxilliary lemmas . . . . .	227

---

## 5.1 Introduction

### 5.1.1 Hawkes Processes and Neural Field Equation

In the present paper we study the large time behavior of a population of interacting and spiking neurons indexed by  $i = 1, \dots, N$ ,  $N \geq 1$ , as the size of the population  $N$  tends to infinity. We model the activity of a neuron by a point process where each point represents the time of a spike: for  $i = 1, \dots, N$ ,  $Z_{N,i}(t)$  counts the number of spikes during the time interval  $[0, t]$  of the  $i$ th neuron of the population. Denoting  $\lambda_{N,i}(t)$  as the conditional intensity of  $Z_{N,i}$  at time  $t$ , that is

$$\mathbf{P}(Z_{N,i} \text{ jumps between } (t, t + dt) | \mathcal{F}_t) = \lambda_{N,i}(t)dt,$$

where  $\mathcal{F}_t := \sigma(Z_{N,i}(s), s \leq t, 1 \leq i \leq N)$ , we want to account for the dependence of the activity of a neuron on the past of the whole population : the spike of one neuron can trigger other spikes. *Hawkes processes* are then a natural choice to emphasize this interdependency and we take here

$$\lambda_{N,i}(t) = f_{\kappa,\varrho} \left( \rho(x_i)e^{-t} + \frac{2\pi}{N} \sum_{j=1}^N \cos(x_i - x_j) \int_0^{t-} e^{-(t-s)} dZ_{N,j}(s) \right), \quad i = 1, \dots, N. \tag{5.1.1}$$

The neurons are located on the circle  $S = (-\pi, \pi]$  with positions  $(x_i)_{1 \leq i \leq N}$  regularly distributed, that is  $x_i = \frac{\pi}{N}(2i - N)$ . We subdivide  $S$  into  $N$  intervals of length  $2\pi/N$  denoted by

$$B_{N,i} = (x_{i-1}, x_i] \quad \text{for } 1 \leq i \leq N, \text{ with } x_0 := -\pi. \quad (5.1.2)$$

The function  $f_{\kappa,\varrho} : \mathbb{R} \rightarrow \mathbb{R}_+$  models the synaptic integration of neuron  $i$  with respect to the input of the other neurons  $j$  in the population, modulated by the spatial kernel  $\cos(x_i - x_j)$ . It is chosen as a sigmoid with parameters  $(\kappa, \varrho)$ ,  $\kappa > 0$ ,  $\varrho \in (0, 1)$ , that is

$$f_{\kappa,\varrho}(u) := (1 + e^{-(u-\varrho)/\kappa})^{-1}. \quad (5.1.3)$$

The function  $\rho : S \rightarrow \mathbb{R}$  represents the initial inhomogeneous voltage of the population and leaks at rate 1. The exponential term  $e^{-(t-s)}$  in the integral in (5.1.1) quantifies how a jump lying back  $t - s$  time units in the past affects the present (at time  $t$ ) intensity: each neuron tends to forget progressively its past. The main object of interest of the paper is the synaptic voltage

$$U_{N,i}(t) = \rho(x_i)e^{-t} + \frac{2\pi}{N} \sum_{j=1}^N \cos(x_i - x_j) \int_0^t e^{-(t-s)} dZ_{N,j}(s) =: \rho(x_i)e^{-t} + X_{N,i}(t), \quad (5.1.4)$$

(i.e.  $\lambda_{N,i}(t) = f_{\kappa,\varrho}(U_{N,i}(t-))$ ) and more precisely the random profile defined for all  $x \in S$  by:

$$U_N(t)(x) := \sum_{i=1}^N U_{N,i}(t) \mathbf{1}_{x \in B_{N,i}}. \quad (5.1.5)$$

The specific form of (5.1.1) originates from the so-called ring model introduced by Shriki *et al.* (2003), modelling the activity of neurons in the visual cortex on a mesoscopic scale. Here each position  $x \in S$  represents a preferred orientation for each neuron, see the biological works of Georgopoulos *et al.* (1982); Bosking *et al.* (1997) and the mathematical works of Veltz & Faugeras (2010); MacLaurin & Bressloff (2020) amongst others. We are looking here at the microscopic counterpart of this model. It means that neurons that prefer close orientation tend to excite each others, whereas neurons with opposite orientation inhibit each others. Making  $\kappa \rightarrow 0$  in (5.1.3), we see that  $f_{\kappa,\varrho}$  converges towards  $H_\varrho$  the Heaviside function

$$H_\varrho(u) = \mathbf{1}_{u \geq \varrho}. \quad (5.1.6)$$

Hence for  $\kappa$  small, a neuron can spike only when it has a high potential with rate approximately 1, and with rate approximately 0 otherwise.

This model (5.1.1) is a specific case of a larger class of mean-field Hawkes processes for which one can write the intensity in the form

$$\lambda_{N,i}(t) = \mu_t(x_i) + f \left( v_t(x_i) + \frac{1}{N} \sum_{j=1}^N w_{ij}^{(N)} \int_0^{t-} h(t-s) dZ_{N,j}(s) \right), \quad i = 1, \dots, N. \quad (5.1.7)$$



The current model (5.1.1) corresponds to the choice  $h(t) = e^{-t}$  and  $w_{ij}^{(N)} = 2\pi \cos(x_i - x_j)$ . In (5.1.7), the neurons are placed in a spatial domain  $I$  endowed with  $\nu$  a probability measure that describes the macroscopic distribution of the positions. The parameter function  $\mu_t : I \rightarrow \mathbb{R}_+$  represents a spontaneous activity of the neuron at time  $t$ ,  $v_t : I \rightarrow \mathbb{R}$  a past activity,  $h$  is the memory kernel of the system,  $f : \mathbb{R} \rightarrow \mathbb{R}^+$  and  $w_{ij}^{(N)}$  represents the interaction between neurons  $i$  and  $j$ . For a suitable class of connectivity sequence  $(w_{ij}^{(N)})$  that can be approximated by some macroscopic interaction kernel  $w(x, y)$  as  $N \rightarrow \infty$  (see [Chevallier et al. \(2019\)](#); [Agathe-Nerine \(2022\)](#) for precise statements), a usual propagation of chaos result as  $N \rightarrow \infty$  (see ([Delattre et al., 2016a](#), Theorem 8), ([Chevallier et al., 2019](#), Theorem 1), ([Agathe-Nerine, 2022](#), Theorem 3.10)) may be stated as follows: for fixed  $T > 0$ , there exists some  $C(T) > 0$  such that

$$\sup_{1 \leq i \leq N} \mathbf{E} \left( \sup_{s \in [0, T]} |Z_{N,i}(s) - \bar{Z}_i(s)| \right) \leq \frac{C(T)}{\sqrt{N}}, \quad (5.1.8)$$

where the limiting process  $(\bar{Z}_i, i = 1, \dots, N)$  consists of independent copies of inhomogeneous Poisson process suitably coupled to  $Z_{N,i}$  with intensity  $(\lambda_t(x_i))_{t \geq 0}$  solving

$$\lambda_t(x) = \mu_t(x) + f \left( v_t(x) + \int_I w(x, y) \int_0^t h(t-s) \lambda_s(y) ds \nu(dy) \right) \quad (5.1.9)$$

(see the above references for details on this coupling, or Definition 3.3.1 of Chapter 3), where  $\nu$  is the macroscopic distribution of the positions. Moreover, for the specific choice  $h(t) = e^{-t}$ , denoting the macroscopic potential of a neuron (the synaptic current) with position  $x$  at time  $t$  by

$$u_t(x) := v_t(x) + \int_I w(x, y) \int_0^t h(t-s) \lambda_s(y) ds \nu(dy), \quad (5.1.10)$$

an easy computation (see [Chevallier et al. \(2019\)](#)) gives that, when  $v_t(x) = \rho(x)e^{-t}$ ,  $u$  solves the *Neural Field Equation* (NFE)

$$\frac{\partial u_t(x)}{\partial t} = -u_t(x) + \int_I w(x, y) f(u_t(y)) \nu(dy), \quad t \geq 0, \quad (5.1.11)$$

with initial condition  $u_0 = \rho$ . The NFE that first appears in [Wilson & Cowan \(1972\)](#) has been extensively studied in the literature, mostly from a phenomenological perspective ([Amari, 1977](#)), and is an important example of macroscopic neural dynamics with non-local interactions (we refer to [Bressloff \(2014\)](#) for an extensive review on the subject). Let us mention here an important point: whereas the analysis of [Chevallier et al. \(2019\)](#) requires the measure  $\nu$  in (5.1.11) to be a probability measure on  $I$ , the historical version of the NFE was originally studied when  $\nu(dy) = dy$  is the Lebesgue measure on  $\mathbb{R}$ . In this last case, thanks to its translation invariance of the Lebesgue measure, one can show the existence of travelling waves solutions to (5.1.11), see [Ermentrout & McLeod \(1993\)](#); [Lang & Stannat \(2016\)](#) for details. The same analysis when  $\nu(dy) = dy$  is replaced by a

probability measure fails, as translation invariance of (5.1.11) is then broken. In this respect, the present choice of  $I = S$  and  $\nu(dy) = \frac{\mathbf{1}_{[-\pi,\pi]}}{2\pi} dy$  combines the two previous advantages:  $\nu$  is a probability measure (hence the previous analysis when  $N \rightarrow \infty$  applies) and translation invariance is preserved in the present periodic case. It can be shown (Kilpatrick & Ermentrout (2013)) that (5.1.11) exhibits localized patterns (*wandering bumps*) which are stationary pulse solutions.

We are interested in this paper in the long time behavior of the microscopic system (5.1.1) and its proximity to these wandering bumps. Before focusing on the microscopic scale, we say a few words on the behavior of the macroscopic system (5.1.9)/(5.1.10). In the pure mean-field case (when  $w_{ij}^{(N)} = 1$  for all  $i, j$ ), the spatial dependency is no longer relevant and (5.1.9) reduces to the scalar nonlinear convolution equation  $\lambda_t = \mu_t + f(v_t + \int_0^t h(t-s)\lambda_s ds)$ . An easy instance concerns the so-called linear case where  $f(x) = x$ ,  $\mu_t = \mu$  and  $v_t = 0$ : in this situation the behavior of  $\lambda_t$  as  $t \rightarrow \infty$  is well known. There is a phase transition (see (Delattre *et al.*, 2016a, Theorems 10,11)) depending on the memory kernel  $h$ : when  $\|h\|_1 = \int_0^\infty h(t)dt < 1$  (the *subcritical case*),  $\lambda_t \xrightarrow[t \rightarrow \infty]{} \frac{\mu}{1 - \|h\|_1}$ , whereas when  $\|h\|_1 > 1$  (the *supercritical case*),  $\lambda_t \xrightarrow[t \rightarrow \infty]{} \infty$ . This phase transition was extended to the inhomogeneous case in Agathe-Nerine (2022) (and more especially where the interaction is made through the realization of weighted random graphs), and the existence of such a phase transition now reads in terms of  $\|h\|_1 r_\infty < 1$  (then  $\lambda_t(x) \rightarrow \ell(x)$  the unique solution of  $\ell(x) = \mu(x) + \int_I w(x,y)\|h\|_1 \ell(y)\nu(dy)$ ) and  $\|h\|_1 r_\infty > 1$  (then  $\|\lambda_t\|_2 \rightarrow \infty$ ), where  $r_\infty$  is the spectral radius of the interaction operator  $T_W g(x) \mapsto \int_I w(x,y)g(y)\nu(dy)$ . In the fully inhomogeneous case and nonlinear case ( $f$  no longer equal to  $Id$ ), a sufficient condition for convergence of  $\lambda_t$  is given in Agathe-Nerine (2023a): whenever

$$\|f'\|_\infty \|h\|_1 r_\infty < 1, \quad (5.1.12)$$

$\lambda_t$  converges to  $\ell$  as  $t \rightarrow \infty$ ,  $\ell$  being the unique solution to

$$\ell = \mu + f(\|h\|_1 T_W \ell). \quad (5.1.13)$$

Note that the present model (5.1.1) obviously does not satisfy (5.1.12), as  $\|f'\|_\infty$  is very large (recall (5.1.3):  $f = f_{\kappa,\varrho}$  is a sigmoid close to the Heaviside function  $H_\varrho$  (5.1.6)). Understanding the longtime behavior of  $\lambda_t$  when (5.1.12) does not hold may be a difficult task for general  $h$ . However the present model is sufficiently simple to be analyzed rigorously: as it was originally noted by Kilpatrick & Ermentrout (2013), the stationary points of (5.1.9) when  $w$  is a cosine can be found by solving an appropriate fixed point relation (see (5.2.2) below) and by invariance by translation, each fixed-point gives rise to a circle of stationary solutions to (5.1.9). One part of the proof will be to show the local stability of these circles (extending the results of Kilpatrick & Ermentrout (2013) when  $f$  is the Heaviside function).

The main concern of the paper is to analyse the microscopic system (5.1.5) on a long time scale. An issue common to all mean-field models (and their

perturbations) is that there is, in general, no possibility to interchange the limits  $N \rightarrow \infty$  and  $t \rightarrow \infty$ . Specifying to Hawkes processes, the constant  $C(T)$  in (5.1.8) is of the form  $\exp(CT)$ , such that (5.1.8) remains only relevant up to  $T \sim c \log N$  with  $c$  sufficiently small. In the linear subcritical case,  $C(T)$  is linear ( $C(T) = CT$ ) so that the mean-field approximation remains relevant up to  $T = o(\sqrt{N})$  (Delattre *et al.* (2016a)). In a previous work Agathe-Nerine (2023a), we showed that, in the subcritical regime defined by (5.1.12) with  $h(t) = e^{-t}$ , the macroscopic intensity (5.1.9) converges to  $\ell$  defined by (5.1.13) and the microscopic intensity (5.1.7) remains close to this limit up to polynomial times in  $N$ , see Chapter 4 and more especially Theorem 4.2.12 for more details. Here, the main difference is that (5.1.10) admits a manifold of stable stationary solutions parameterized by  $S$ , instead of a unique one. We show here that, with some initial condition close to this manifold, our microscopic process (5.1.5) stays close to the manifold up to time horizons that are polynomial in  $N$ , and moreover the dynamics of the microscopic current follows a Brownian motion on the manifold.

**Organization of the paper** The paper is organized as follows: after introducing some notations, we start in Section 5.1.2.2 by introducing the precise mathematical set-up. In Section 5.2, we present the main results of our paper. Section 5.2.1 is divided into three parts: in the first part 5.2.1.1, we present the deterministic dynamics of (5.1.16) and the manifold of stationary solutions  $\mathcal{U}$  defined in (5.2.4). In the second part we introduce two ways of defining some phase reduction along  $\mathcal{U}$ , the variational phase (Proposition 5.2.7) and isochronal phase (Proposition 5.2.8). In the last part, Theorem 5.2.9 ensures that if the system is close to  $\mathcal{U}$ , it stays so for a long time, and with Theorem 5.2.11, we analyze the dynamics of the isochronal phase of  $U_N$  along  $\mathcal{U}$ . Such dynamics are represented in the simulations of Figure 5.2. In Section 5.2.2, we explain how our paper is linked to the present litterature on the subject. In Section 5.2.3, we sketch the strategy of proof we follow. Section 5.3 collects the proofs of the results of Sections 5.2.1.1 and 5.2.1.2, Section 5.4 concerns the proof of the proximity between  $U_N$  and  $\mathcal{U}$  seen in Theorem 5.2.9 and Section 5.5 is devoted to prove the diffusive behavior of  $U_N$  along  $\mathcal{U}$  seen in Theorem 5.2.11. Some technical estimates and computations are gathered in the appendix.

## 5.1.2 Notations and definition

### 5.1.2.1 Notations

We denote by  $C_{\text{parameters}}$  a constant  $C > 0$  which only depends on the parameters inside the lower index. These constants can change from line to line or inside a same equation, and when it is not relevant, we just write  $C$ . For any  $d \geq 1$ , we denote by  $|x|$  and  $x \cdot y$  the Euclidean norm and scalar product of  $x, y \in \mathbb{R}^d$ . For  $(E, \mathcal{A}, \mu)$  a measured space, for a function  $g$  in  $L^p(E, \mu)$  with  $p \geq 1$ , we write  $\|g\|_{E, \mu, p} := \left(\int_E |g|^p d\mu\right)^{\frac{1}{p}}$ . When  $p = 2$ , we denote by  $\langle \cdot, \cdot \rangle$  the Hermitian scalar product in  $L^2(E, \mu)$ . Without ambiguity, we may omit the subscript  $(E, \mu)$  or  $\mu$ .

For a real-valued bounded function  $g$  on a space  $E$ , we write  $\|g\|_\infty := \|g\|_{E,\infty} = \sup_{x \in E} |g(x)|$ .

For  $(E, d)$  a metric space, we denote by  $\|g\|_L = \sup_{x \neq y} |g(x) - g(y)|/d(x, y)$  the Lipschitz seminorm of a real-valued function  $g$  on  $E$ . We denote by  $\mathcal{C}(E, \mathbb{R})$  the space of continuous functions from  $E$  to  $\mathbb{R}$ , and  $\mathcal{C}_b(E, \mathbb{R})$  the space of continuous bounded ones. For any  $T > 0$ , we denote by  $\mathbb{D}([0, T], E)$  the space of càdlàg (right continuous with left limits) functions defined on  $[0, T]$  and taking values in  $E$ . For any integer  $N \geq 1$ , we denote by  $\llbracket 1, N \rrbracket$  the set  $\{1, \dots, N\}$ .

For any  $h, k, l \in E$ , we denote by  $Dg(h)[k] \in S$  the derivative of  $g : E \rightarrow F$  at  $h$  in the direction  $k$ , and similarly for second derivatives  $D^2g(h)[k, l]$ .

### 5.1.2.2 Definition of the model

We define now formally our process of interest. Definition 5.1.1 follows a standard representation of point processes as thinning of independent Poisson measures, see Ogata (1988); Delattre *et al.* (2016a).

**Definition 5.1.1.** Let  $(\pi_i(ds, dz))_{1 \leq i \leq N}$  be a sequence of i.i.d. Poisson random measures on  $\mathbb{R}_+ \times \mathbb{R}_+$  with intensity measure  $dsdz$ . The multivariate counting process  $(Z_{N,1}(t), \dots, Z_{N,N}(t))_{t \geq 0}$  defined by, for all  $t \geq 0$  and  $i \in \llbracket 1, N \rrbracket$ :

$$Z_{N,i}(t) = \int_0^t \int_0^\infty \mathbf{1}_{\{z \leq \lambda_{N,i}(s)\}} \pi_i(ds, dz), \quad (5.1.14)$$

where  $\lambda_{N,i}$  is defined in (5.1.1) is called a *multivariate Hawkes process* with set of parameters  $(N, \kappa, \varrho, \rho)$ .

It has been showed in several works (see e.g. Agathe-Nerine (2022); Delattre *et al.* (2016a) amongst others) that the process defined in (5.1.14) is well posed in the following sense.

**Proposition 5.1.2.** *For a fixed realization of the family  $(\pi_i)_{1 \leq i \leq N}$ , there exists a pathwise unique multivariate Hawkes process (in the sense of Definition 5.1.1) such that for any  $T < \infty$ ,*

$$\sup_{t \in [0, T]} \sup_{1 \leq i \leq N} \mathbf{E}[Z_{N,i}(t)] < \infty.$$

Proposition 5.1.2 can be found in (Agathe-Nerine, 2022, Propositions 2.5). In our framework, the macroscopic intensity (5.1.9) population limits is

$$\lambda_t(x) = f_{\kappa, \varrho} \left( \rho(x) e^{-t} + \int_S \cos(x - y) \int_0^t e^{-(t-s)} \lambda_s(y) ds dy \right), \quad (5.1.15)$$

and the neural field equation (5.1.11) becomes

$$\frac{\partial u_t(x)}{\partial t} = -u_t(x) + \int_S \cos(x - y) f_{\kappa, \varrho}(u_t(y)) dy. \quad (5.1.16)$$

**Proposition 5.1.3.** *Let  $T > 0$ . There exists a unique solution  $(u_t)_{t \in [0, T]}$  in  $\mathcal{C}_b(S, \mathbb{R})$  to (5.1.16) with initial condition  $u_0 = \rho$ .*

Proposition 5.1.3 can be found in (Agathe-Nerine, 2022, Propositions 2.7), and follows from a standard Grönwall estimate. We can then define the flow of (5.1.16) by  $(t, g) \mapsto \psi_t(g)$ , that is the solution at time  $t$  of (5.1.16) starting from  $g$  at  $t = 0$ :

$$\psi_t(g)(x) = e^{-t}g(x) + \int_0^t e^{-(t-s)} \int_S \cos(x-y) f_{\kappa, \varrho}(\psi_s(g)(x)) ds. \quad (5.1.17)$$

## 5.2 Stability of wandering bumps for interacting Hawkes processes

### 5.2.1 Main results

#### 5.2.1.1 Stationary solutions to (5.1.16)

We are concerned here with the stationary solutions to (5.1.16), that is

$$u(x) = \int_{-\pi}^{\pi} \cos(x-y) f(u(y)) dy. \quad (5.2.1)$$

We follow a similar approach to Kilpatrick & Ermentrout (2013), see Appendix 5.A.1.

**Remark 5.2.1.** For a general choice of  $f$ , if  $u$  is solution to (5.2.1), then for any  $\phi$ ,  $x \mapsto u(x + \phi)$  is also solution to (5.2.1) by invariance of  $S$ . Expanding the cosine, (5.2.1) becomes

$$u(x) = \cos(x) \int_{-\pi}^{\pi} \cos(y) f(u(y)) dy + \sin(x) \int_{-\pi}^{\pi} \sin(y) f(u(y)) dy.$$

By translation symmetry, with no loss of generality we can ask  $\int_{-\pi}^{\pi} \sin(y) f(u(y)) dy = 0$  and solving (5.2.1) means finding  $A \geq 0$  such that

$$A = \int_S \cos(y) f(A \cos(y)) dy. \quad (5.2.2)$$

As (5.2.1) is invariant by translation, any  $A$  solution to (5.2.2) gives rise to the set  $\mathcal{U}_A := \{x \mapsto A \cos(x + \phi), \phi \in [-\pi, \pi]\}$  of stationary solutions to (5.2.1).

Recall (5.1.6), when  $f = H_{\varrho}$  the Heaviside function with threshold  $\varrho$ , Kilpatrick & Ermentrout (2013) and Veltz & Faugeras (2010) showed that for  $\varrho \in [-1, 1]$ , the unique solutions to (5.2.2) are

$$A = 0, \quad A_-(0) = \sqrt{1 + \varrho} - \sqrt{1 - \varrho} \quad \text{and} \quad A_+(0) := \sqrt{1 + \varrho} + \sqrt{1 - \varrho}. \quad (5.2.3)$$

This result is recalled in Appendix 5.A.1. One can show that the set  $\mathcal{U}_{A_-(0)}$  is unstable whereas  $\mathcal{U}_{A_+(0)}$  and  $\mathcal{U}_0$  are locally stable. In the following we focus on the largest fixed point  $A_+(0)$  which we rename for  $A(0)$  by convenience. Recall that in the paper, we are under the assumption that  $f = f_{\kappa, \varrho}$  defined in (5.1.3) for a small fixed  $\kappa$ . As  $f_{\kappa, \varrho} \xrightarrow{\kappa \rightarrow 0} H_{\varrho}$ , our first result is that when  $\kappa$  is close enough to 0, we can still find a stationary solution to (5.1.16) of the form  $u = A(\kappa) \cos$  where  $A(\kappa)$  is also close to  $A(0)$ .

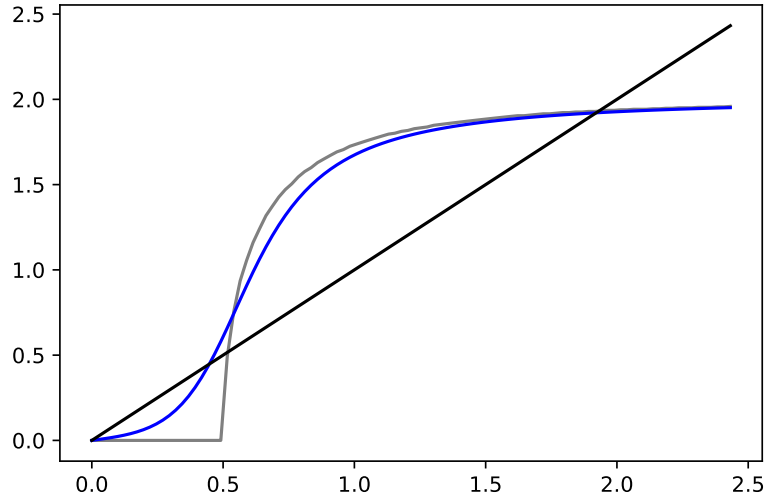


Figure 5.1: Graph of  $G : A \mapsto \int_S A \cos(x) f(A \cos(x)) dx$

We represent the fixed-point function  $G$  appearing in (5.2.2) for the choice  $f = H_\varrho$  defined in (5.1.6) in gray and its smooth version with  $f = f_{\kappa,\varrho}$ , defined in (5.1.3) in blue. We chose  $(\kappa, \varrho) = (\frac{1}{10}, \frac{1}{2})$ . The black line is the graph of  $y = x$  and its intersections with the two other lines give the fixed points of  $G$ . Note that we are interested here on the fixed point on the far right, that is  $A(0)$  for the gray line and  $A(\kappa)$  for the blue line.

**Proposition 5.2.2.** *Assume  $\varrho \in (-1, 1)$ . Then there exists  $\kappa_0 > 0$  and a function  $A : (0, \kappa_0) \rightarrow (|\varrho|, +\infty)$  of class  $C^1$  such that for any  $\kappa \in (0, \kappa_0)$ ,  $u = A(\kappa) \cos$  is a stationary solution to (5.1.16) when  $f = f_{\kappa,\varrho}$  and  $A(\kappa) \xrightarrow{\kappa \rightarrow 0} A(0)$  given in (5.2.3). Moreover, there exists  $\kappa_1 \in (0, \kappa_0)$  such that for any  $\kappa \in (0, \kappa_1)$ ,  $1 < I(1, \kappa) < 2$  for  $I(1, \kappa) := \int_S f'_{\kappa,\varrho}(A(\kappa) \cos(x)) dx$ .*

Proposition 5.2.2 is based on a simple implicit function argument and is proved in the Appendix 5.A.2. An illustration of this Proposition is done in Figure 5.1: we see that for each  $A$  solving (5.2.2) for the Heaviside function, there is indeed another close  $A$  solving (5.2.2) for the sigmoid function with small  $\kappa$ . For the rest of the paper we fix  $\varrho \in (-1, 1)$ ,  $\kappa < \kappa_1$  and  $A = A(\kappa)$  and may omit the indexes  $(\kappa, \varrho)$ . We have then established that

$$\mathcal{U} := (A \cos(\cdot + \phi))_{\phi \in S} =: (u_\phi)_{\phi \in S} \quad (5.2.4)$$

is a set of stationary solutions to (5.1.16), which is a manifold parameterized by the circle  $S$ . To study the stability of these stationary solutions, we introduce linear operators that are also parameterized by the circle  $S$ .

**Definition 5.2.3.** Let  $\phi \in S$ , and define for any function  $\psi \in L^2(S)$

$$T_\phi \psi(x) := \int_S \cos(x - y) f'(u_\phi(y)) \psi(y) dy \quad (5.2.5)$$

$$\mathcal{L}_\phi \psi := -\psi + T_\phi \psi. \quad (5.2.6)$$

Define also  $L_\phi^2 := L_{f'(u_\phi)}^2$ , that is the  $L^2$  weighted space defined by the scalar product

$$\langle g_1, g_2 \rangle_{2,\phi} = \int_S g_1(x) g_2(x) f'(u_\phi(x)) dx.$$

We denote by  $\|\cdot\|_{2,\phi}$  the associated norm. Recall (5.2.4) and define

$$v_\phi := \partial_x u_\phi = -A \sin(\cdot + \phi). \quad (5.2.7)$$

We consider also the orthogonal projection  $P_\phi^\circ$  on  $\text{Span}(v_\phi)$  and its complementary projection  $P_\phi^\perp$ , both defined for any  $g \in L_\phi^2$  by

$$P_\phi^\circ g := \frac{\langle g, v_\phi \rangle_{2,\phi}}{\|v_\phi\|_{2,\phi}} v_\phi =: \alpha_\phi^\circ(g) v_\phi \quad (5.2.8)$$

$$P_\phi^\perp g := g - P_\phi^\circ g. \quad (5.2.9)$$

We will also need the projection on  $\text{Span}(u_\phi)$  hence we define

$$\alpha_\phi^\gamma(g) = \frac{\langle g, u_\phi \rangle_{2,\phi}}{\|u_\phi\|_{2,\phi}}. \quad (5.2.10)$$

**Remark 5.2.4.** Without ambiguity and for a general  $\phi$ , we may write  $\|\cdot\|_\phi$  instead of  $\|\cdot\|_{2,\phi}$  to gain in clarity. Note that by compactness of  $S$ , since  $0 < \inf_{[-A,A]} f' < \sup_{[-A,A]} f' < \infty$ , the norms  $\|\cdot\|_2$  and  $\|\cdot\|_{2,\phi}$  are equivalent: there exists  $C_0, \tilde{C}_0 > 0$  (independent of  $\phi$ ) such that for any  $g \in L^2(S)$ ,

$$\tilde{C}_0 \|g\|_2 \leq \sup_{\phi \in S} \|g\|_{2,\phi} \leq C_0 \|g\|_2. \quad (5.2.11)$$

**Proposition 5.2.5.** *Let  $\phi \in S$ . The operator  $\mathcal{L}_\phi$  defined in (5.2.6) is self-adjoint in  $L_\phi^2$  and has three distinct eigenvalues,  $-1, 0$  and  $\gamma \in (-1, 0)$ . If for  $\iota \in \{-1, \gamma, 0\}$ , we denote by  $\mathcal{E}_\iota$  the eigenspace associated to the eigenvalue  $\iota$ , one has that  $\mathcal{E}_0 = \text{Ker} \mathcal{L}_\phi = \text{Span}(u_\phi)$ ,  $\mathcal{E}_\gamma = \text{Span}(v_\phi)$  and  $\mathcal{E}_{-1} = (\text{Span}(u_\phi, v_\phi))^\perp$ . Moreover,  $\mathcal{E}_0 \perp \mathcal{E}_\gamma$ . Furthermore, there exists  $C_\mathcal{L}, C_P$  such that for any  $\phi \in S$ ,  $\mathcal{L}_\phi$  generates an analytic semigroup of contraction  $(e^{t\mathcal{L}_\phi})$  and for any  $g \in L_\phi^2$ ,  $t \geq 0$ ,*

$$\|e^{t\mathcal{L}_\phi} P_\phi^\perp g\|_{2,\phi} \leq e^{t\gamma} \|P_\phi^\perp g\|_\phi, \quad (5.2.12)$$

$$\|e^{t\mathcal{L}_\phi} g\|_2 \leq C_\mathcal{L} \|g\|_2, \quad (5.2.13)$$

$$\|e^{t\mathcal{L}_\phi} P_\phi^\perp g\|_{2,\phi} \leq C_P \|g\|_{2,\phi}. \quad (5.2.14)$$

Proposition 5.2.5 is proved in Section 5.3.1. A straightforward corollary of Proposition 5.2.5 is the following

**Corollary 5.2.6.** *The manifold  $\mathcal{U}$  is locally stable under the flow (5.1.16): there exists  $\varepsilon_0 > 0$  such that, for any  $g \in L^2(S)$  satisfying  $\text{dist}_{L^2}(g, \mathcal{U}) \leq \varepsilon_0$ , we have  $\lim_{t \rightarrow \infty} \text{dist}_{L^2}(\psi_t(g), \mathcal{U}) = 0$  where  $\psi$  is defined in (5.1.17). We denote by  $B(\mathcal{U}, \varepsilon_0) := \{g \in L^2(I), \text{dist}_{L^2}(g, \mathcal{U}) \leq \varepsilon_0\}$ .*

### 5.2.1.2 Representation on the manifold

Recall that we are interested in the behaviour of the process (5.1.5), when the initial condition  $U_N(0)$  to (5.1.5) is close to the manifold  $\mathcal{U}$  introduced in (5.2.4). We need a way to define a proper phase reduction of  $U_N$  along  $\mathcal{U}$ . We have two ways to do so that we use in our results that are well explained in the recent work Adams & MacLaurin (2022), which takes the NFE as a good class of examples and motivation. The first one is via the *variational phase*, defined in the following Proposition 5.2.7:

**Proposition 5.2.7** (Variational phase). *There exists  $\varpi > 0$  such that, for any  $g \in L^2(S)$  satisfying  $\text{dist}_{L^2(S)}(g, \mathcal{U}) \leq \varpi$ , there exists a unique phase  $\phi := \text{proj}_{\mathcal{U}}(g) \in S$  such that  $P_\phi^\circ(g - u_\phi) = 0$  and the mapping  $g \mapsto \text{proj}_{\mathcal{U}}(g)$  is smooth.*

The second one is via the *isochronal phase*, defined in the following Proposition 5.2.8. In a few words, as the manifold  $\mathcal{U}$  is stable and attractive, a solution to the NFE from a neighborhood of  $\mathcal{U}$  is attracted to  $\mathcal{U}$  and converges to it. As  $t \rightarrow \infty$ , it identifies with one stationary solution of the manifold, we call it its isochron.

**Proposition 5.2.8** (Isochronal phase). *For any  $g \in B(\mathcal{U}, \varepsilon_0)$  (see Corollary 5.2.6), there exists a unique  $\theta(g) \in S$  such that*

$$\|\psi_t(g) - u_{\theta(g)}\|_2 \xrightarrow[t \rightarrow \infty]{} 0, \quad (5.2.15)$$

where  $\psi$  is defined in (5.1.17). Such a map  $\theta : B(\mathcal{U}, \varepsilon_0) \rightarrow S$  is called the *isochronal map* of  $\mathcal{U}$ , and  $\theta(g)$  is the *isochronal phase* of  $g$ . Moreover, it is three times continuously Fréchet differentiable (in fact  $C^\infty$ ), and in particular for  $u_\phi \in \mathcal{U}$ ,  $h, l \in L^2(S)$ , we have

$$D\theta(u_\phi)[h] = \frac{\langle v_\phi, h \rangle_\phi}{\|v_\phi\|_\phi}, \text{ and} \quad (5.2.16)$$

$$\begin{aligned} D^2\theta(u_\phi)[h, l] &= \frac{1}{2A^2} (\alpha_\phi^\circ(h)\beta_\phi(v_\phi, l) + \alpha_\phi^\circ(l)\beta_\phi(v_\phi, h) + \beta_\phi(h, l)) \\ &\quad + \frac{1 + \gamma}{2A^2(1 - \gamma)} (\alpha_\phi^\gamma(h)\beta_\phi(u_\phi, l) + \alpha_\phi^\gamma(l)\beta_\phi(u_\phi, h)) \\ &\quad - \frac{(2 - \gamma)(1 + \gamma)}{2(1 - \gamma)} (\alpha_\phi^\circ(h)\alpha_\phi^\circ(l) + \alpha_\phi^\gamma(h)\alpha_\phi^\gamma(l)), \end{aligned} \quad (5.2.17)$$

where  $\alpha_\phi^\circ$  and  $\alpha_\phi^\gamma$  are respectively defined in (5.2.8) and (5.2.10), and

$$\beta_\phi(h, l) := \int_S f''(u_\phi(y))v_\phi(y)h(y)l(y)dy. \quad (5.2.18)$$

Note that in particular, as  $u_{\theta(g)} \in \mathcal{U}$  and  $\mathcal{U}$  consists in stationary points,  $\psi_t(u_{\theta(g)}) = u_{\theta(g)}$ . Propositions 5.2.7 and 5.2.8 are proved in Section 5.3.2.

### 5.2.1.3 Long time behavior

The first result uses the variational phase to ensure that  $(U_N(t))$  defined in (5.1.5) reaches a neighborhood of  $\mathcal{U}$  in time of order  $\log(N)$  and stays inside it for arbitrary polynomial times in  $N$ .

**Theorem 5.2.9.** *Suppose that  $\rho \in B(\mathcal{U}, \varepsilon_0)$  and*

$$\|U_N(0) - \rho\|_2 \xrightarrow[N \rightarrow \infty]{} 0. \quad (5.2.19)$$

Let  $\alpha, \tau_f > 0$ . There exists some  $C > 0$  such that, defining for any  $N \geq 1$ ,  $T_0(N) := C \log(N)$ , for any  $\varepsilon > 0$ ,

$$\mathbf{P} \left( \sup_{t \in [T_0(N), N^\alpha \tau_f]} \text{dist}_{L^2}(U_N(t), \mathcal{U}) \leq \varepsilon \right) \xrightarrow[N \rightarrow \infty]{} 1. \quad (5.2.20)$$



**Remark 5.2.10.** In fact, we show a more precise result than (5.2.20) that will be useful for the proof of Theorem 5.2.11: we prove that for any fixed  $\eta \in (0, \frac{1}{4})$ , we have with some constant  $C > 0$

$$\mathbf{P} \left( \sup_{t \in [T_0(N), N^\alpha \tau_f]} \text{dist}_{L^2}(U_N(t), \mathcal{U}) \leq CN^{\eta-1/2} \right) \xrightarrow{N \rightarrow \infty} 1.$$

Theorem 5.2.9 is proved in Section 5.4. The second main result of the paper is the analysis of the behavior of  $U_N$  along  $\mathcal{U}$  when  $\alpha = 1$ .

**Theorem 5.2.11.** *Let  $\rho \in B(\mathcal{U}, \varepsilon_0)$ . Suppose (5.2.19). Let  $\tau_f > 0$ . There exist a deterministic  $\theta_0 \in S$  and for every  $N$  some  $\tau_0(N) \propto \frac{\log(N)}{N}$  and a càdlàg process  $(W_N(t))_{t \in (\tau_0(N), \tau_f)}$  that converges weakly in  $\mathbb{D}([0, \tau_f], S)$  towards a standard Brownian such that for every  $\varepsilon > 0$ ,*

$$\lim_{N \rightarrow \infty} \mathbf{P} \left( \sup_{\tau \in (\tau_0(N), \tau_f)} \|U_N(N\tau) - u_{\theta_0 + \sigma W_N(\tau)}\|_2 \leq \varepsilon \right) = 1, \quad (5.2.21)$$

where

$$\sigma := \left( 2\pi \int_S \sin^2(x) f(A \cos(x)) dx \right)^{\frac{1}{2}}, \quad (5.2.22)$$

with  $A = A(\kappa)$  defined with Proposition 5.2.2.

Theorem 5.2.11 is proved in Section 5.5. We have run several simulations to illustrate our results, see Figure 5.2. We represent the evolution of the current  $U_N(t, x)$  for  $t \in [0, T_{\max}]$  where the time is on the x-axis and spatial position on the y-axis. The different values taken are scaled with a color bar. We can see the *wandering bumps* evolving in Figure 5.2a, whereas in Figure 5.2b the initialization is too far from the manifold and the system is no longer attracted to  $\mathcal{U}$ .

## 5.2.2 Link with the literature

Hawkes processes have been introduced in Hawkes (1971) to model earthquakes and have been thoroughly studied since, see e.g. Brémaud & Massoulié (1996). The seminal work of Delattre *et al.* (2016a) has renewed the interest for large population of interacting Hawkes processes, which have proven to be particularly useful in a neuroscience context to model the mutually exciting properties of a population of neurons, see for instance Ditlevsen & Löcherbach (2017); Chevallier *et al.* (2019).

In this respect, a common setting for the modelling of interacting neurons is the mean-field framework. For instance, in Baladron *et al.* (2012), the authors describe the propagation of chaos in networks of Hodgkin-Huxley and FitzHugh-Nagumo neurons. Another popular model is the integrate-and-fire dynamics, first introduced in the seminal work of Lapicque (1907), and still studied mathematically, as e.g. in Masi *et al.* (2014); Delarue *et al.* (2015) and also Cormier *et al.* (2020).

Several works have extended the mean-field framework to take into account the presence of a macroscopic spatial structure in the interaction, originally for diffusion models (see Touboul (2014); Luçon & Stannat (2016)), as well as for

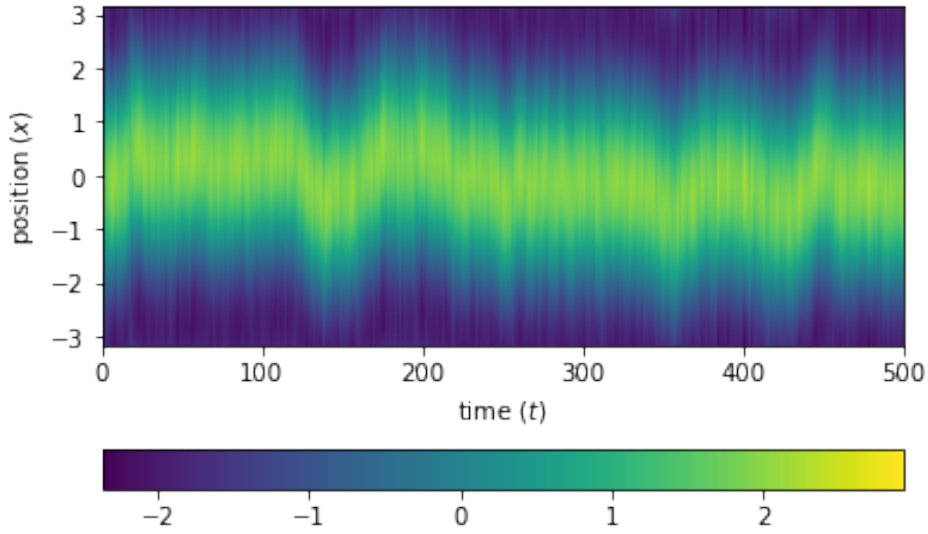
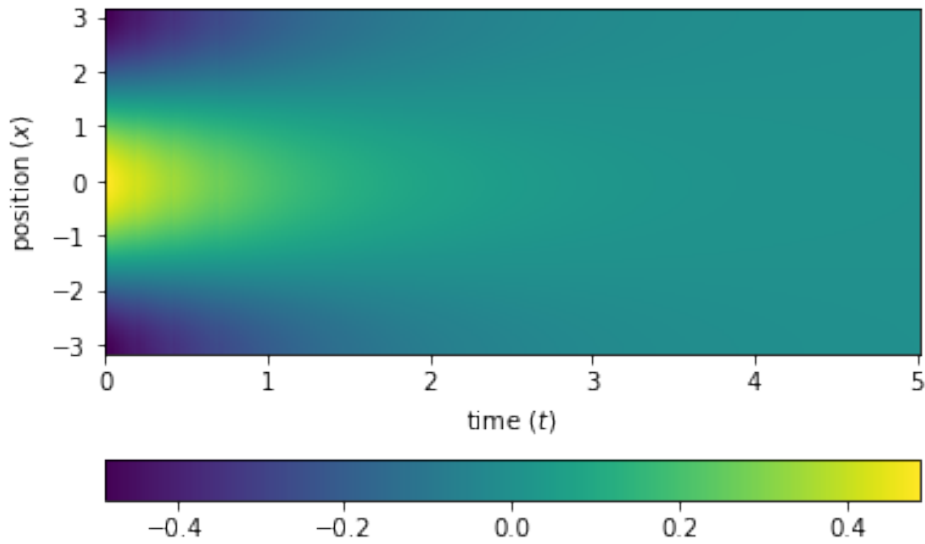

 (a) The initialization in the vicinity of  $\mathcal{U}$  leads to wandering bumps

 (b) The initialization far from the vicinity of  $\mathcal{U}$  does not trigger the wandering bumps

 Figure 5.2: Evolution of the voltage  $U_N(t)(x)$ 

We chose  $(\kappa, \varrho) = (\frac{1}{20}, \frac{1}{2})$  and run simulations of  $N = 500$  neurons following (5.1.1). We represent the evolution of the current  $U_N(t, x_i)$  obtained for two different simulations where we changed the initial profile  $\rho$ . In 5.2a, we start in a vicinity of  $\mathcal{U}$  as we take for initialization  $\rho(x) = A(\kappa) \cos(x) + \cos(2x)$ , where  $A(\kappa)$  solving (5.2.2) for  $f = f_{\kappa, \varrho}$  is found by a numerical root finding method, with a final time  $T_{\max} = 500$  (of the same order that the size of the population). In 5.2b, we initialize the system with  $\rho(x) = \frac{1}{4} A(\kappa) \cos(x)$ . It is too far from the manifold  $\mathcal{U}$  and we can see that the dynamics is attracted to  $\mathcal{U}_A$  where  $A$  is the smallest solution of (5.2.2) (in Figure 5.1 it corresponds to the far left intersection of the black and blue lines) which is approximately 0, hence we only run the simulation with a final time  $T_{\max} = 5$ .

Hawkes processes (see [Ditlevsen & Löcherbach \(2017\)](#); [Chevallier \*et al.\* \(2019\)](#)). The main difficulty with this extension is that we lose the exchangeability specific to homogeneous mean-field models as in [Sznitman \(1991\)](#); [Delattre \*et al.\* \(2016a\)](#). Concerning our present model, [Chevallier \*et al.\* \(2019\)](#) was the first to provide with a rigorous mesoscopic interpretation of the neural field equation (5.1.11) in terms of the limit of spatially extended Hawkes processes interacting through a mesoscopic spatial kernel. The recent work [Agathe-Nerine \(2022\)](#) extends this result for Hawkes processes interacting on inhomogeneous random graphs. Another possibility to circumvent the exchangeability issue would have been to use replica mean-field models as [Davydov \(2022\)](#) and describe the propagation of chaos for an infinite number of replicas. Note however that this description keeps the size  $N$  of the population fixed, whereas we want to have  $N \rightarrow \infty$ .

Note also that the present model include interaction that may be negative: this reflects some inhibitive effect among neurons with opposite orientations. Modelling the inhibition present in the brain has been historically difficult. For Hawkes processes, a common approach is to allow the synaptic kernel  $h$  in (5.1.7) to take negative values. This is however impossible for linear Hawkes processes as the intensity cannot be negative. To circumvent this, one has to choose a non-negative and nonlinear function  $f$  to preserve the non-negativity of the intensity. A classic choice is to take  $f(x) = \max(0, \mu + x)$  (see for instance [Bonnet \*et al.\* \(2021\)](#) for estimation model or [Costa \*et al.\* \(2020\)](#); [Cattiaux \*et al.\* \(2022\)](#) with  $h$  in (5.1.7) signed and with compact support). One can also introduce inhibition through a signed multiplying factor (that may depend or not on the neuron), see for instance [Duarte \*et al.\* \(2019\)](#); [Ditlevsen & Löcherbach \(2017\)](#); [Pfaffelhuber \*et al.\* \(2022\)](#). Some works have also parted the whole population into two subclasses of neurons, the excitatory ones and the inhibitory ones ([Raad & Löcherbach, 2020](#); [Duval \*et al.\*, 2022](#)). In the latter, the inhibition is made thanks to a (small) multiplicative factor onto the intensity of the excitatory population. The present work is another contribution concerning models with inhibition, as it is present thanks to the cosine interaction kernel that takes negative values. This choice is essential to our dynamics as the balance between excitation and inhibition within the population of neurons allows to have a stable manifold of stationary solutions to (5.1.16).

The analysis of mean-field interacting processes on long time scales has a significant history in the case of interacting diffusions, in particular in the case of phase oscillators as the Kuramoto model ([Kuramoto \(1975\)](#), see [Giacomin & Poquet \(2015\)](#) and references therein for a comprehensive review on the subject). The techniques used in the present work have some formal similarities to the ones used for diffusions, the main difference being that with Hawkes processes, the noise is Poissonian (rather Brownian) and multiplicative (rather than additive). The so-called uniform propagation of chaos concerns situations where estimates such as (5.1.8) are uniform in time. Such estimates are commonly met in reversible situations (e.g. granular type media diffusions in [Bolley \*et al.\* \(2013\)](#)). See also the recent paper of [Colombani & Bris \(2022\)](#), where the authors studies a uniform propagation of chaos on the FitzHugh-Nagumo diffusive model.

Let us comment on the analysis of the Kuramoto model as it presents some informal proximity with our model. One is here interested in the longtime behavior

of the empirical measure  $\mu_{N,t} := \frac{1}{N} \sum_{i=1}^N \delta_{\theta_{i,t}}$  of the system of interacting diffusions  $(\theta_1, \dots, \theta_N)$  solving the system of coupled SDEs

$$d\theta_{i,t} = -\frac{K}{N} \sum_{j=1}^N \sin(\theta_{i,t} - \theta_{j,t}) dt + dB_{i,t},$$

with  $(B_i)$  i.i.d. Brownian motions. Standard propagation of chaos techniques show that  $\mu_N$  converges weakly on a bounded time interval  $[0, T]$  to the solution  $\mu_t$  to the nonlinear Fokker-Planck (NFP) equation

$$\partial_t \mu_t = \frac{1}{2} \partial_\theta^2 \mu_t + K \partial_\theta \left( \mu_t (\sin * \mu_t) \right), \quad (5.2.23)$$

(to compare with our microscopic current  $U_{N,i}$  in (5.1.4) converging towards  $u_t$  solution to the NFE (5.1.16)). One can easily prove the existence of a phase transition for (5.2.23): when  $K \leq 1$ ,  $\mu \equiv \frac{1}{2\pi}$  is the only (stable) stationary point of (5.2.23) (subcritical case), whereas it coexists with a stable circle of synchronised profiles when  $K > 1$  (supercritical case). A series of papers have analysed the longtime behavior of the empirical measure  $\mu_N$  of the Kuramoto model (and extensions) in both the subcritical and supercritical cases, the first one being Bertini *et al.* (2014), followed by Giacomini *et al.* (2012); Luçon & Poquet (2017); Coppini (2022); Delarue & Tse (2021). The main arguments of the mentioned papers lie in a careful analysis of two contradictory phenomena that arise on a long-time scale: the stability of the deterministic dynamics around stationary points (that forces  $\mu_N$  to remain in a small neighborhood of these points) and the presence of noise in the microscopic system (which makes  $\mu_N$  diffuse around these points).

We are here in a similar situation to the supercritical case: the deterministic dynamics of the spatial profile  $U_N$  (given by (5.1.5)) has a stationary manifold  $\mathcal{U}$  (defined in (5.2.4)) which possesses sufficient stability properties, see Corollary 5.2.6. The point of the analysis relies then on a time discretization and some careful control on the diffusive influence of noise that competes with the deterministic dynamics. In a previous work Agathe-Nerine (2023a), we have analysed in depth the case where (5.1.13) has a unique solution, that would be comparable to the subcritical case of the Kuramoto model.

The first main result of the paper is to show that once  $U_N(0)$  is close to the stationary manifold  $\mathcal{U}$ , it stays so for a long time, see Theorem 5.2.9. The next step is to find a way to describe the projection of the dynamics onto  $\mathcal{U}$ . A convenient tool for this is the use of isochronicity, we refer to Guckenheimer (1975) for a precise approach on the subject, and to Giacomini *et al.* (2018) for their use of isochronicity to study the proximity between the noisy trajectory of interacting particles and the limit cycle in a finite dimensional setting. See also Luçon & Poquet (2021a) where the microscopic system is a diffusion and the large population limit admits a stable periodic solution: they show that the empirical measure stays close to the periodic solution with a random dephasing. The isochron map in this case helps to describe the dephasing as a Brownian motion with a constant drift.

Going back to Hawkes processes, several other works have already complemented the propagation of chaos result mentioned in (5.1.8) and studied, mostly at the level of fluctuations, finite approximations of the NFE. Central Limit Theorems (CLT) have been obtained in [Delattre \*et al.\* \(2016a\)](#); [Ditlevsen & Löcherbach \(2017\)](#) for homogeneous mean-field Hawkes processes (when both time and  $N$  go to infinity) or with age-dependence in [Chevallier \(2017\)](#). One should also mention the functional fluctuation result recently obtained in [Heesen & Stannat \(2021\)](#), also in a pure mean-field setting. A result closer to our case with spatial extension is [Chevallier & Ost \(2020\)](#), where a functional CLT is obtained for the spatial profile  $U_N$  around its limit. Note here that all of these works provide approximation results of quantities such that  $\lambda_N$  or  $U_N$  that are either valid on a bounded time interval  $[0, T]$  or under strict growth condition on  $T$  (see in particular the condition  $\frac{T}{N} \rightarrow 0$  for the CLT in [Ditlevsen & Löcherbach \(2017\)](#)), whereas we are here concerned with time-scales that grow polynomially with  $N$ .

One alternative to study large time behavior is to use a Brownian approximation of the dynamics of  $U_N$ , see the initial work of [Ditlevsen & Löcherbach \(2017\)](#). However this approximation is based on the comparison of the corresponding semigroups and is not uniform in time. Nevertheless, let us comment on this diffusive approximation in large population regime on bounded time intervals that can be found in both [Ditlevsen & Löcherbach \(2017\)](#); [Chevallier & Ost \(2020\)](#). A second order approximation of the NFE was proposed in [Chevallier & Ost \(2020\)](#) with (adapted to the notations of the present article)

$$dU_N(t) = -U_N(t)dt + w * f(U_N(t))dt + C \int_S w(x, y) \frac{\sqrt{f(U_N(t)(y))}}{\sqrt{N}} W(dt, dy), \quad (5.2.24)$$

where  $W$  is a Gaussian white noise. This approximating diffusion process (5.2.24) is a noisy NFE, it can be seen as an intermediate modeling between the microscopic scale given by the Hawkes process and the macroscopic scale given by the NFE. In our framework with a cosine kernel, the infinitesimal increment of the noise in (5.2.24) can be expanded as

$$\begin{aligned} C \cos(x) \int_S \cos(y) \frac{\sqrt{f(U_N(t)(y))}}{\sqrt{N}} W(dt, dy) \\ + C \sin(x) \int_S \sin(y) \frac{\sqrt{f(U_N(t)(y))}}{\sqrt{N}} W(dt, dy). \end{aligned}$$

To compare with our result, let us informally project the last quantity on  $\text{Ker}(\mathcal{L}_0)$  introduced in Proposition 5.2.5. The scalar product  $\langle \cdot, v_0 \rangle_{2,0}$  with  $v_0 = -A \sin(\cdot)$  gives that the cosine term becomes zero and the noise left is a random variable of the form

$$\begin{aligned} -CA \int_S \sin^2 f'(A \cos) \int_S \sin(y) \frac{\sqrt{f(U_N(t)(y))}}{\sqrt{N}} W(dt, dy) \\ = -C \int_S \sin(y) \frac{\sqrt{f(U_N(t)(y))}}{\sqrt{N}} W(dt, dy), \end{aligned}$$

using (5.3.2). The infinitesimal noise that effectively drives the dynamics of (5.2.24) along  $\mathcal{U}$  is then Gaussian with variance proportional to  $\int_S \sin^2(y) \frac{f(U_N(t)(y))}{N} dy dt$ , which is exactly the variance found in (5.2.22), rescaled by  $\frac{1}{N}$  and where  $U_N(t)$  has been replaced by the limit  $u_t$ . This analogy remains informal, but shows that our results are compatible to the computations of Ditlevsen & Löcherbach (2017) and Chevallier & Ost (2020): one could see the present result as a rigorous justification that the approximation introduced by Ditlevsen & Löcherbach (2017); Chevallier & Ost (2020) can be extended for polynomial times in  $N$ .

Approximation between Hawkes and Brownian dynamics has also been studied in Chevallier *et al.* (2021); Erny (2023), based on Komlós, Major and Tusnády (KMT) coupling techniques (see Ethier & Kurtz (1986)). Recently, Prodhomme (2023) used similar KMT coupling techniques applied to finite dimensional Markov chains and found Gaussian approximation to remain precise for very large periods of time. However these results are valid for  $\mathbb{Z}^d$ -valued continuous-time Markov chains, it is unclear how they can be applied in our situation (with infinite dimension and space extension). The proof we propose is direct and does not rely on such Brownian coupling.

The question of Stochastic Neural Field Equations has also been considered directly from a macroscopic perspective at multiple times. It consists in considering the NFE (5.1.11) with an additive or multiplicative spatio-temporal noise, see for instance Bressloff & Webber (2012); Krüger & Stannat (2014). Existence and uniqueness results have been obtained for various expressions of the noise, see Faugeras & Inglis (2015); Inglis & MacLaurin (2016). Let us mention in particular Kilpatrick & Ermentrout (2013); MacLaurin & Bressloff (2020); Cihak *et al.* (2022) who propose a heuristical derivation of the diffusion coefficient of the wandering bumps in a setting similar to ours (the ring model with  $f$  the Heaviside function). See also MacLaurin (2023) where the author studies the effect of the added noise on patterns such that traveling waves and oscillations thanks to the use of some projection of the dynamics, to obtain long time stability. Whereas all of the previous results are concerned with a macroscopic approach concerning stochastic perturbation of the NFE, we provide here a rigorous and microscopic interpretation of this phenomenon.

## 5.2.3 Strategy of proof of the long time behavior

### 5.2.3.1 About Theorem 5.2.9

Section 5.4 is devoted to prove the proximity result of Theorem 5.2.9. This in particular requires some spectral estimates on the operators  $\mathcal{L}_\phi$  introduced in Definition 5.2.3 and the stability of stationary solutions to (5.1.16), results that are gathered in Section 5.2.1.1 and proved in Section 5.3. The main lines of proof for Theorem 5.2.9 are given in Section 5.4. The strategy of proof is sketched here, and follows the one used in a previous work Agathe-Nerine (2023a).

First we show in Proposition 5.4.1 that one can find some initial time  $T_0(N) \propto$

$\log(N)$  for which  $\text{dist}_{L^2}(U_N(T_0(N)), \mathcal{U}) \leq \frac{N^{2\eta}}{\sqrt{N}}$ , with  $0 < \eta < \frac{1}{4}$ . This essentially boils down to following the predominant deterministic dynamics of the NFE. Let  $T_f(N) = N^\alpha$ , we discretize the interval of interest  $[T_0(N), T_f(N)]$  into  $n_f$  intervals of same length  $T$  denoted by  $[T_i, T_{i+1}]$ ,  $T$  chosen sufficiently large below. On each subinterval, we can decompose the dynamics of  $U_N(t)$  in terms of, at first order, the linearized dynamics of (5.1.16) around any stationary solution, modulo some drift terms coming from the mean-field approximation, some noise term coming from the underlying Poisson measure, and some quadratic remaining error coming from the nonlinearity of  $f$ . It gives a semimartingale decomposition of  $U_N(t) - u_{\text{proj}(U_N(T_i))}$  for  $t \in [T_i, T_{i+1}]$ , detailed in Section 5.4.2.

Provided one has some sufficient control on each of these terms in the semimartingale expansion on a bounded time interval, we do an iterative procedure that works as follows: the point is to see that provided  $U_N$  is initially close to  $u_{\text{proj}(U_N(T_i))} \in \mathcal{U}$ , it will remain close to it for a time interval of length  $T$  for some sufficiently large deterministic  $T > 0$  so that the deterministic dynamics prevails upon the other contributions. The time horizon at which one can pursue this recursion is controlled by moment estimates on the noise in Proposition 5.4.3.

### 5.2.3.2 About Theorem 5.2.11

Section 5.5 is devoted to prove the analysis of the behavior of  $U_N$  along  $\mathcal{U}$  seen in Theorem 5.2.11. We sketch here the strategy of proof. First we use the semimartingale decomposition of  $U_N$

$$dU_N(t) = B_N(t)dt + dM_N(t)$$

(with  $B_N$  some drift and  $M_N$  a martingale defined in (5.4.30)) and Itô formula to write the semimartingale decomposition of  $\theta(U_N(t))$  on the interval  $[T_0(N), N\tau_f]$ . As in Theorem 5.2.9, one can show a careful control on each of the terms appearing in the semimartingale decomposition, as done in Section 5.5.3. The difficulty here is to show rigorously that there is no macroscopic drift appearing on this time scale (this point is essentially due to the invariance by rotation of the whole problem). After rescaling the time by  $N$ , we identify the noise with a Brownian motion thanks to Aldous' tightness criterion and Lévy's characterization so that the result of Theorem 5.2.11 follows.

### 5.2.3.3 Extensions

**On the interaction kernel** - Note that Theorem 5.2.11 is of local nature: stability holds provided the initial condition  $\rho$  is sufficiently close to  $\mathcal{U}$ . Following Kilpatrick & Ermentrout (2013), it would be possible to consider the more general interaction kernel

$$w(x, y) = \sum_{k=0}^n A_k \cos(k(x - y)),$$

with more than one Fourier mode. The fixed point equation (5.2.2) becomes a more complicated system of equations

$$A_k = \int_S \cos(kx) f \left( \sum_{k=0}^n A_k \cos(kx) \right) dx. \quad (5.2.25)$$

The exact number of solutions to (5.2.25) remain unclear but if one can solve (5.2.25) and show local stability of the solutions  $u_\phi(x) = \sum_{k=0}^N A_k \cos(k(x + \phi))$ , the same strategy should apply: we would obtain local stability provided one starts sufficiently close to these structures.

**Oscillatory behavior** - Note that  $\mathcal{U}$  consists of stationary points. We claim that a similar strategy should apply also to situations where (5.1.1) admits generic oscillations, see [Giacomin \*et al.\* \(2014\)](#) in a context of diffusion. We have in particular in mind the framework proposed in [Ditlevsen & Löcherbach \(2017\)](#): the authors study interacting Hawkes processes with Erlang memory kernel. The population is divided into classes, and the classes interact with a cycling feedback system, so that the large population limit is attracted to non-constant periodic orbits. It is reasonable to think that our techniques can be transposed to this situation, to show that the microscopic system is closed to the limit cycle under their hypotheses in large times and without using the approximating diffusion process.

### 5.3 Stationary solutions (proofs)

Let us first define for any function  $r \in L^2(S)$

$$\mathcal{I}(r) := \int_S r(y) f'(u_0(y)) dy, \quad (5.3.1)$$

where  $u_0$  is defined in (5.2.4). We start by giving a computation Lemma that will be useful in the whole paper.

**Lemma 5.3.1.** *We have*

$$\mathcal{I}(\sin^2) = 1, \quad \mathcal{I}(\cos^2) = \mathcal{I}(1) - 1 \quad \text{and} \quad \mathcal{I}(\sin \cos) = 0.$$

*Proof.* Recall that  $u_0 = A \cos$ , as  $A$  solves (5.2.2) by integrating by parts we obtain

$$A = \int_S \cos(y) f(A \cos(y)) dy = A \int_S \sin^2(y) f'(u_0(y)) dy = A \mathcal{I}(\sin^2), \quad (5.3.2)$$

and as  $A > 0$  it implies  $\mathcal{I}(\sin^2) = 1$ . By integrating by parts we also have

$$-A \mathcal{I}(\cos \sin) = \int_{-\pi}^{\pi} \sin(y) f(A \cos(y)) dy.$$

Since  $y \rightarrow \sin(y) f(A \cos(y))$  is odd, we obtain that  $\mathcal{I}(\cos \sin) = 0$ . As  $\cos^2 = 1 - \sin^2$  and  $\mathcal{I}$  is linear, we have  $\mathcal{I}(\cos^2) = \mathcal{I}(1) - \mathcal{I}(\sin^2) = \mathcal{I}(1) - 1$ .  $\square$



### 5.3.1 Stability

Here we prove Proposition 5.2.5.

*Proof.* Let  $\phi \in S$ . Let us first show that the operator  $\mathcal{L}_\phi$  is indeed self-adjoint in  $L_\phi^2$ . Let  $g_1, g_2 \in L_\phi^2$ , we have by Fubini's theorem and recalling Definition 5.2.3

$$\begin{aligned} & \langle \mathcal{L}_\phi g_1, g_2 \rangle_\phi \\ &= - \int_S g_1 g_2 f'(u_\phi) + \int_S \left( \int_S \cos(x-y) f'(u_\phi(y)) g_1(y) dy \right) g_2(x) f'(u_\phi(x)) dx \\ &= - \int_S g_1 g_2 f'(u_\phi) + \int_S f'(u_\phi(y)) g_1(y) \left( \int_S \cos(x-y) g_2(x) f'(u_\phi(x)) dx \right) dy \\ &= \langle g_1, \mathcal{L}_\phi g_2 \rangle_\phi, \end{aligned}$$

hence  $\mathcal{L}_\phi$  is self-adjoint in  $L_\phi^2$ .

We focus now on its spectrum, we want to prove that it has three distinct eigenvalues,  $-1$ ,  $0$  and  $\gamma \in (-1, 0)$ . The following arguments follow the same procedure of the one that can be found in Kilpatrick & Ermentrout (2013). First note that  $T_\phi$  is compact in  $L_\phi^2$  (in fact, with finite range). Hence it has a discrete spectrum consisting of eigenvalues. Let  $\lambda$  be an eigenvalue of  $\mathcal{L}_\phi$  and  $\psi$  an associated eigenvector, that is  $\mathcal{L}_\phi \psi = \lambda \psi$  hence  $(\lambda + 1)\psi = T_\phi \psi$  with Definition 5.2.3. As seen in Remark 5.2.1,  $\lambda$  does not depend on  $\phi$  and if  $\psi$  is an eigenvector for  $\phi = 0$ , then  $\psi(\cdot - \phi)$  is an eigenvector for  $\phi$ . Hence, in the following, we focus on the case  $\phi = 0$ . We have

$$T_0 \psi(x) = A_0(\psi) \cos(x) + B_0(\psi) \sin(x), \quad (5.3.3)$$

with

$$A_0(\psi) := \int_S \cos(y) f'(u_0(y)) \psi(y) dy, \quad B_0(\psi) := \int_S \sin(y) f'(u_0(y)) \psi(y) dy. \quad (5.3.4)$$

The eigenvalue  $-1$  is spanned by functions  $\psi \in L^2$  such that  $A_0(\psi) = B_0(\psi) = 0$ . Recall (5.3.1), we have that, since  $(\lambda + 1)\psi = T_0 \psi$ ,

$$\begin{aligned} (\lambda + 1)A_0(\psi) &= \int_S \cos(y) (\lambda + 1)\psi(y) f'(u_0(y)) dy \\ &= \int_S \cos(y) (A_0(\psi) \cos(y) + B_0(\psi) \sin(y)) f'(u_0(y)) dy \\ &= A_0(\psi) \mathcal{I}(\cos^2) + B_0(\psi) \mathcal{I}(\sin \cos), \end{aligned} \quad (5.3.5)$$

and similarly,

$$(\lambda + 1)B_0(\psi) = A_0(\psi) \mathcal{I}(\sin \cos) + B_0(\psi) \mathcal{I}(\sin^2). \quad (5.3.6)$$

See Lemma 5.3.1 for the computations of  $\mathcal{I}(\cos^2)$ ,  $\mathcal{I}(\sin^2)$  and  $\mathcal{I}(\sin \cos)$ . Putting these computations into (5.3.5) and (5.3.6) implies that  $(\lambda, \psi)$  solves  $\mathcal{L}_0 \psi = \lambda \psi$  if and only if

$$\begin{cases} (\lambda + 1)A_0(\psi) &= (\mathcal{I}(1) - 1) A_0(\psi) \\ (\lambda + 1)B_0(\psi) &= B_0(\psi). \end{cases}$$

With no loss of generality, one can suppose that  $\psi$  is such that  $(A_0(\psi), B_0(\psi)) \neq (0, 0)$  (and thus  $\lambda \neq -1$ ). Then  $(\lambda, \psi)$  solves the previous system if and only if, either  $\lambda = 0$  with  $A_0(\psi) = 0$  and  $B_0(\psi) \neq 0$  (and hence we see from (5.3.3) that the eigenvalue 0 is spanned by  $\sin \propto v_0$ ) or  $\lambda = \gamma$  given by

$$\gamma := \mathcal{I}(1) - 2 = \int_S f'(A \cos(x)) dx - 2, \quad (5.3.7)$$

with  $A_0(\psi) \neq 0$  and  $B_0(\psi) = 0$ , so that the eigenspace related to  $\gamma$  is one-dimensional, spanned by  $\cos \propto u_0$ . The fact that  $\langle u_\phi, v_\phi \rangle_\phi = 0$  follows immediately from the fact that  $u_\phi$  is even and  $v_\phi$  is odd. The last eigenvalue  $\lambda = -1$  is spanned by  $\psi$  such that  $A(\psi) = B(\psi) = 0$ .

To conclude the proof of Proposition 5.2.5, it remains to prove the inequalities (5.2.12), (5.2.13) and (5.2.14). We come back to a general  $\phi \in S$ . By definition of the projection  $P_\phi^\circ$  in (5.2.8), we have that  $\mathcal{L}_\phi P_\phi^\circ = 0$ . Moreover, by definition of  $P_\phi^\perp$  in (5.2.9), we have that for any  $g \in L_\phi^2$ ,  $P_\phi^\perp g$  belongs in the orthogonal of  $\text{Ker}(\mathcal{L}_\phi)$  in  $L_\phi^2$ . Then  $\mathcal{L}_\phi P_\phi^\perp = \mathcal{L}_\phi (Id - P_\phi^\circ)$  generates a contraction semigroup on  $L^2(S)$  and (5.2.12) follows then from functional analysis (see e.g. (Pazy, 1974, Theorem 3.1)). For the two last inequalities, we use Remark 5.2.4. From the definition of the projection  $P_\phi^\circ$  in (5.2.8), we have that

$$e^{t\mathcal{L}_\phi} P_\phi^\circ g = \frac{\langle g, v_\phi \rangle_\phi}{\|v_\phi\|_\phi} e^{t\mathcal{L}_\phi} v_\phi = \frac{\langle g, v_\phi \rangle_\phi}{\|v_\phi\|_\phi} v_\phi,$$

as  $v_\phi \in \text{Ker}(\mathcal{L}_\phi)$ . We obtain then  $\|e^{t\mathcal{L}_\phi} P_\phi^\circ g\|_\phi \leq \|g\|_\phi \|v_\phi\|_\phi$ . From (5.2.12) we have  $\|e^{t\mathcal{L}_\phi} P_\phi^\perp g\|_\phi \leq e^{\gamma t} \|P_\phi^\perp g\|_\phi \leq C_P \|g\|_\phi$  for some  $C_P > 0$ , that is exactly (5.2.14). As  $\|e^{t\mathcal{L}_\phi} g\|_2 \leq \|e^{t\mathcal{L}_\phi} P_\phi^\circ g\|_2 + \|e^{t\mathcal{L}_\phi} P_\phi^\perp g\|_2$ , (5.2.13) follows for the choice  $C_{\mathcal{L}} = C_1 C_2 \max(\sup_{\phi \in S} \|v_\phi\|_\phi, C_P)$ .  $\square$

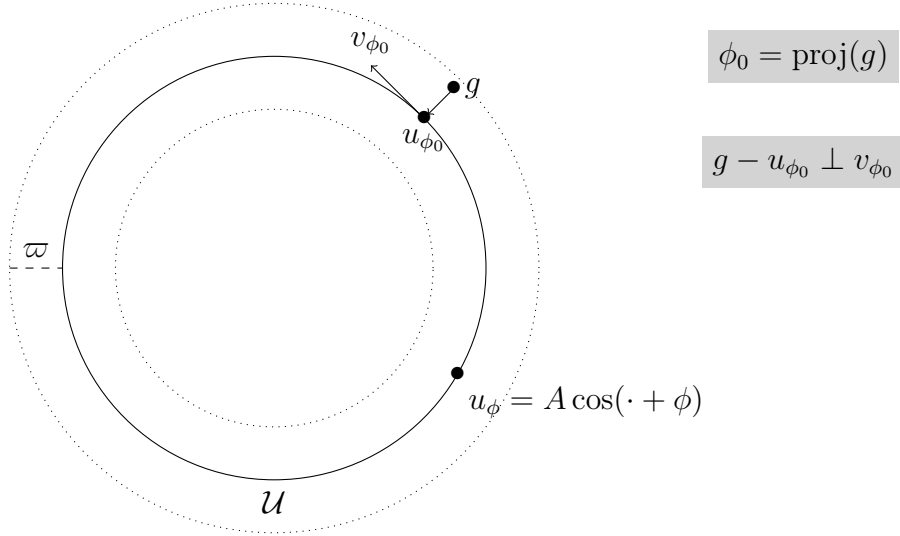
### 5.3.2 Projections on the manifold

We prove that both the variational phase seen in Proposition 5.2.7 and isochronal phase seen in Proposition 5.2.8 are well defined.

*Proof of Proposition 5.2.7.* (similar to (Luçon & Poquet, 2017, Lemma 2.8)) Define for any  $(g, \phi) \in L^2(S) \times S$ :

$$F(g, \phi) := \int_S (g(x) - u_\phi(x)) v_\phi(x) f'(u_\phi(x)) dx = \langle g - u_\phi, v_\phi \rangle_\phi.$$

We have for any fixed  $\phi_0$ ,  $F(u_{\phi_0}, \phi_0) = 0$ . Note that  $F$  is smooth in both variables as it can be written  $F(g, \phi) = -A \int_S (g(x) - A \cos(x + \phi)(x)) \sin(x + \phi) f'(u_\phi(x)) dx$ . Moreover,  $\partial_\phi F(u_{\phi_0}, \phi_0) = -\langle v_{\phi_0}, v_{\phi_0} \rangle_{\phi_0} = -A^2 \mathcal{I}_{\phi_0}(\sin^2)$  with  $\mathcal{I}_\phi(r) := \int_S r(y + \phi) f'(u_\phi(y)) dy$ . By invariance on the circle  $\mathcal{I}_{\phi_0}(\sin^2) = \mathcal{I}(\sin^2)$  defined in (5.3.1) and Lemma 5.3.1 implies then that  $\partial_\phi F(u_{\phi_0}, \phi_0) = -A^2 = -A(\kappa)^2 \neq 0$  with Proposition 5.2.2. By the implicit function theorem, for any  $\phi_0$  there exists a neighborhood  $\mathcal{V}(u_{\phi_0})$  of  $u_{\phi_0}$  such that the projection is well defined (i.e. for any


 Figure 5.3: Projection of  $g \in L^2(S)$  on  $\mathcal{U}$ 

$g \in \mathcal{V}(u_{\phi_0})$ , there exists a unique  $\phi$  such that  $F(g, \phi) = 0$  and  $g \mapsto \text{proj}_{\mathcal{U}}(g)$  is smooth). By compactness of  $\mathcal{U}$ , the existence of  $\varpi$  and the result of Proposition 5.2.7 follow. The situation can be summarized by the following Figure 5.3.  $\square$

*Proof of Proposition 5.2.8.* We reproduce the argument of (Adams, 2023, Theorem 3.1) that establishes the existence and regularity of the isochron map in a more general context than here.

Let  $g \in B(\mathcal{U}, \varepsilon_0)$  and  $(\varepsilon_n)_n$  a sequence decreasing to 0. The first step is to prove that  $\theta(g)$  satisfying (5.2.15) exists. To do so, using the stability of  $\mathcal{U}$  proved in Corollary 5.2.6, one can find an increasing sequence of times  $(t_n)$  and a sequence of closed non-empty sets  $\Phi_n \subset \mathcal{U}$  such that for all  $n \in \mathbb{N}$  and  $\theta \in \Phi_n$ ,  $\|\psi_{t_n}(g) - u_\theta\|_2 \leq C\varepsilon_n$  for some constant  $C > 0$ . It gives in particular that the diameter of  $\Phi_n$  tends to zero as  $n \rightarrow \infty$ , hence the existence of a unique  $\theta(g)$  such that  $\bigcap_{n \in \mathbb{N}} \Phi_n = \{u_{\theta(g)}\}$  by Cantor's Intersection Theorem. The second step is to prove the regularity of  $\theta : B(\mathcal{U}, \varepsilon_0) \rightarrow S$ . As  $\mathcal{U}$  is parameterized by  $S$ , we can define  $\pi(u)$  for  $u \in \mathcal{U}$  as the unique  $\phi \in S$  such that  $u = u_\phi$ . As the flow  $\psi$  is  $\mathcal{C}^\infty$ , the map  $g \mapsto \lim_{t \rightarrow \infty} \psi_t(g)$  is well defined and  $\mathcal{C}^\infty$ , and we have also  $\lim_{t \rightarrow \infty} \psi_t(g) = u_{\theta(g)}$ . Then  $\theta(g)$  can be written as  $\pi(\lim_{t \rightarrow \infty} \psi_t(g))$ , hence  $g \mapsto \theta(g)$  is indeed  $\mathcal{C}^\infty$ .

We focus now on the derivatives of  $g \mapsto \theta(g)$ . Define  $\Gamma : g \in B(\mathcal{U}, \varepsilon_0) \mapsto \Gamma(g) = \lim_{t \rightarrow \infty} \Psi_t g = u_{\theta(g)} \in \mathcal{U}$ . From Proposition 5.2.8,  $\Gamma$  is smooth and is differentiable, and for  $g, h \in L^2(S)$ ,  $D\Gamma(g)[h] = u'_{\theta(g)} D\theta(g)[h] = v_{\theta(g)} D\theta(g)[h] \in L^2$ . Applied for  $g = u_\phi$  and taking the scalar product with  $v_\phi$ , one obtains

$$\langle D\Gamma(u_\phi)[h], v_\phi \rangle = D\theta(u_\phi)[h] \|v_\phi\|^2. \quad (5.3.8)$$

Let us focus on  $D\Psi_t g[h]$ . Let  $g_t$  be the solution of (5.1.16) with  $g_0 = g$ , that is  $g_t = \Psi_t(g)$ , and  $h_t$  the solution of (5.1.16) with  $h_0 = g + h$ , that is  $h_t = \Psi_t(g + h)$ .

Then with Taylor's formula we have

$$\begin{aligned}\partial_t(h_t - g_t) &= -(h_t - u_t) + \cos * (f(h_t) - f(g_t)) \\ &= -(h_t - g_t) + \cos * (f'(g_t)(h_t - g_t)) + r_t,\end{aligned}$$

where  $r_t := \cos * \left( (h_t - g_t)^2 \int_0^1 (1-s) f''(g_t + s(h_t - g_t)) ds \right) = o(\|h\|)$ . We have then that  $D\Psi_t(g)[h] =: w_t$  with

$$\partial_t w_t = -w_t + \cos * (f'(\Psi_t g)(w_t)), \quad w_0 = h. \quad (5.3.9)$$

In particular for the choice  $g = u_\phi$ ,  $D\Psi_t(u_\phi)[h] = e^{t\mathcal{L}_\phi} h$  where  $\mathcal{L}_\phi$  is defined in (5.2.6). Moreover we can write with the operators defined in Definition 5.2.3

$$e^{t\mathcal{L}_\phi} h = e^{t\mathcal{L}_\phi} (P_\phi^\circ h + P_\phi^\perp h) = \frac{\langle h, v_\phi \rangle_\phi}{\|v_\phi\|_\phi} v_\phi + e^{t\mathcal{L}_\phi} P_\phi^\perp h.$$

From (5.2.12),  $\|e^{t\mathcal{L}_\phi} P_\phi^\perp h\|_\phi \leq e^{t\gamma} \|P_\phi^\perp h\|_\phi$  hence  $\lim_{t \rightarrow \infty} e^{t\mathcal{L}_\phi} h = \frac{\langle h, v_\phi \rangle_\phi}{\|v_\phi\|_\phi} v_\phi$ . As

$\Gamma(u_\phi) = \lim_{t \rightarrow \infty} \Psi_t u_\phi = u_\phi$  and  $\lim_{t \rightarrow \infty} D\Psi_t(u_\phi)[h] = \frac{\langle h, v_\phi \rangle_\phi}{\|v_\phi\|_\phi} v_\phi$ , we obtain that

$$D\Gamma(u_\phi)[h] = D \left( \lim_{t \rightarrow \infty} \Psi_t u_\phi \right) [h] = \lim_{t \rightarrow \infty} D\Psi_t(u_\phi)[h] = \lim_{t \rightarrow \infty} e^{t\mathcal{L}_\phi} h = \frac{\langle h, v_\phi \rangle_\phi}{\|v_\phi\|_\phi} v_\phi,$$

which gives with (5.3.8) the result (5.2.16).

We focus now on  $D^2\theta$ . Recall  $\Gamma$ , for  $g, h, l \in B(\mathcal{U}, \varepsilon_0)$ ,

$$D^2\Gamma(g)[h, l] = -D\theta(g)[h]D\theta(g)[l]u_{\theta(g)} + D^2\theta(g)[h, l]v_{\theta(g)}.$$

Applied for  $g = u_\phi$ , it gives with (5.2.16)

$$D^2\Gamma(u_\phi)[h, l] = -\frac{\langle v_\phi, h \rangle_\phi \langle v_\phi, l \rangle_\phi}{\|v_\phi\|_\phi^2} u_\phi + D^2\theta(u_\phi)[h, l]v_\phi.$$

Taking the scalar product with  $v_\phi$ , as  $\langle u_\phi, v_\phi \rangle_\phi = 0$  we obtain

$$D^2\theta(u_\phi)[h, l] = \frac{\langle D^2\Gamma(u_\phi)[h, l], v_\phi \rangle_\phi}{\|v_\phi\|_\phi^2}. \quad (5.3.10)$$

Let us focus on  $D^2\Psi_t g[h, l]$ . We have that  $D\Psi_t(g)[h] = w_t$ , recall that it solves (5.3.9). Let  $D\Psi_t(g+l)[h] := \tilde{w}_t$ , it solves

$$\partial_t \tilde{w}_t = -\tilde{w}_t + \cos * (f'(\Psi_t(g+l))\tilde{w}_t), \quad \tilde{w}_0 = h.$$

As done before, we obtain that  $\zeta_t := \tilde{w}_t - w_t$  solves with  $\zeta_0 = 0$

$$\begin{aligned}\partial_t \zeta_t &= -\zeta_t + \cos * [f'(\Psi_t(g+l))\tilde{w}_t - f'(\Psi_t g)w_t] \\ &= -\zeta_t + \cos * [f'(\Psi_t(g+l))(\zeta_t + w_t) - f'(\Psi_t g)w_t] \\ &= -\zeta_t + \cos * [f'(\Psi_t(g+l))\zeta_t] + \cos * [(f'(\Psi_t(g+l)) - f'(\Psi_t g))w_t].\end{aligned}$$

From Taylor's expansion in  $l$ ,

$$\begin{aligned} f'(\Psi_t(g+l)) &= f' \left( \Psi_t(g) + D\Psi_t(g)[l] + \int_0^1 (1-s)D^2\Psi_t(g)[l]^2 ds \right) \\ &= f'(\Psi_t(g)) + f''(\Psi_t(g))D\Psi_t(g)[l] + o(\|l\|) \end{aligned}$$

hence

$$\cos * [f'(\Psi_t(g+l))\zeta_t] = \cos * (f'(\Psi_t(g))\zeta_t) + O(\|l\|),$$

and

$$\begin{aligned} \cos * [(f'(\Psi_t(g+l)) - f'(\Psi_t(g)))w_t] &= \cos * (f''(\Psi_t(g))D\Psi_t(g)[l]w_t) + o(\|l\|) \\ &= \cos * (f''(\Psi_t(g))D\Psi_t(g)[l]D\Psi_t(g)[h]) + o(\|l\|). \end{aligned}$$

We obtain then after linearizing that  $D^2\Psi_t(g)[h, l] = \xi_t$  is solution of

$$\partial_t \xi_t = -\xi_t + \cos * (f'(\Psi_t(g))\xi_t) + \cos * (f''(\Psi_t(g))D\Psi_t(g)[l]D\Psi_t(g)[h]), \quad \xi_0 = 0.$$

In particular, for the choice  $g = u_\phi$ ,

$$\partial_t \xi_t = \mathcal{L}_\phi \xi_t + \cos * [f''(u_\phi)(e^{t\mathcal{L}_\phi}h)(e^{t\mathcal{L}_\phi}l)], \quad \xi_0 = 0,$$

hence it solves the mild equation

$$\xi_t = \int_0^t e^{(t-s)\mathcal{L}_\phi} (\cos * (f''(u_\phi)(e^{s\mathcal{L}_\phi}h)(e^{s\mathcal{L}_\phi}l))) ds.$$

Recall (5.3.10), hence we focus now on  $\langle \xi_t, v_\phi \rangle_\phi$ . From Proposition 5.2.5,  $\mathcal{L}_\phi$  is self-adjoint hence

$$\begin{aligned} \langle \xi_t, v_\phi \rangle_\phi &= \int_0^t \langle e^{(t-s)\mathcal{L}_\phi} (\cos * (f''(u_\phi)(e^{s\mathcal{L}_\phi}h)(e^{s\mathcal{L}_\phi}l))), v_\phi \rangle_\phi ds \\ &= \int_0^t \langle \cos * (f''(u_\phi)(e^{s\mathcal{L}_\phi}h)(e^{s\mathcal{L}_\phi}l)), e^{(t-s)\mathcal{L}_\phi} v_\phi \rangle_\phi ds \\ &= \int_0^t \langle \cos * (f''(u_\phi)(e^{s\mathcal{L}_\phi}h)(e^{s\mathcal{L}_\phi}l)), v_\phi \rangle_\phi ds \end{aligned}$$

as  $v_\phi \in \text{Ker } \mathcal{L}_\phi$ . Recall (5.2.8) and (5.2.10). By the spectral decomposition of  $\mathcal{L}_\phi$  along its eigenvalues 0,  $\gamma$  and  $-1$ , one has with Proposition 5.2.5, for  $s \geq 0$ ,

$$\begin{aligned} e^{s\mathcal{L}_\phi}h &= \alpha_\phi^\circ(h)v_\phi + e^{s\gamma}\alpha_\phi^\gamma(h)u_\phi + e^{-s}(h - \alpha_\phi^\circ(h)v_\phi - \alpha_\phi^\gamma(h)u_\phi) \\ &= e^{-s}h + \alpha_\phi^\circ(h)(1 - e^{-s})v_\phi + \alpha_\phi^\gamma(h)(e^{s\gamma} - e^{-s})u_\phi, \end{aligned}$$

so that one obtains

$$(e^{s\mathcal{L}_\phi}h)(e^{s\mathcal{L}_\phi}l) = \alpha_\phi^\circ(h)\alpha_\phi^\circ(l)(1 - e^{-s})^2 v_\phi^2 + \alpha_\phi^\gamma(h)\alpha_\phi^\gamma(l)(e^{s\gamma} - e^{-s})^2 u_\phi^2 \quad (5.3.11)$$

$$+ e^{-s}(1 - e^{-s})\{\alpha_\phi^\circ(h)l + \alpha_\phi^\circ(l)h\}v_\phi \quad (5.3.12)$$

$$+ e^{-s}(e^{s\gamma} - e^{-s})\{\alpha_\phi^\gamma(h)l + \alpha_\phi^\gamma(l)h\}u_\phi \quad (5.3.13)$$

$$+ (1 - e^{-s})(e^{s\gamma} - e^{-s})\{\alpha_\phi^\circ(h)\alpha_\phi^\gamma(l) + \alpha_\phi^\circ(l)\alpha_\phi^\gamma(h)\}u_\phi v_\phi \quad (5.3.14)$$

$$+ e^{-2s}hl. \quad (5.3.15)$$

We compute now  $\langle \xi_t, v_\phi \rangle_\phi$  based on the previous decomposition. Fix some generic test functions  $h$  and  $l$ . Then

$$\langle \cos * (f''(u_\phi)hl), v_\phi \rangle_\phi = \int_S v_\phi(x) f'(u_\phi(x)) \int_S \cos(x-y) f''(u_\phi(y)) h(y) l(y) dy dx.$$

Expanding the cosine within the convolution and as  $\int_S v_\phi(x) f'(u_\phi(x)) \cos(x+\phi) dx = 0$ , we have with Lemma 5.3.1

$$\begin{aligned} & \langle \cos * (f''(u_\phi)hl), v_\phi \rangle_\phi \\ &= \left( \int_S v_\phi(x) f'(u_\phi(x)) \sin(x+\phi) dx \right) \int_S \sin(y+\phi) f''(u_\phi(y)) h(y) l(y) dy, \\ &= -A\mathcal{I}(\sin^2) \int_S \sin(y+\phi) f''(u_\phi(y)) h(y) l(y) dy = \int_S f''(u_\phi(y)) v_\phi(y) h(y) l(y) dy. \end{aligned}$$

If now we take  $h = l = v_\phi$  or  $h = l = u_\phi$ , we see that the two terms of (5.3.11) give a zero contribution to  $\langle \xi_t, v_\phi \rangle_\phi$  as the function within the last integral is odd. Taking now  $h = v_\phi$  (resp.  $h = u_\phi$ ) for given  $l$ , we see that the generic term within (5.3.12) (resp. (5.3.13)) gives rise to

$$\begin{aligned} \langle \cos * (f''(u_\phi)lv_\phi), v_\phi \rangle_\phi &= \int_S f''(u_\phi) v_\phi(y)^2 l(y) dy, \\ \langle \cos * (f''(u_\phi)lu_\phi), v_\phi \rangle_\phi &= \int_S f''(u_\phi) v_\phi(y) u_\phi(y) l(y) dy. \end{aligned}$$

Applying finally the last expression for  $l = v_\phi$  gives for (5.3.14), by integration by parts

$$\begin{aligned} \langle \cos * (f''(u_\phi)u_\phi v_\phi), v_\phi \rangle_\phi &= \int_S f''(u_\phi) v_\phi(y)^2 u_\phi(y) dy \\ &= - \int_S \frac{d}{dy} \{u_\phi(y) v_\phi(y)\} f'(u_\phi(y)) dy \\ &= - \int_S v_\phi(y)^2 f'(u_\phi(y)) dy + \int_S u_\phi(y)^2 f'(u_\phi(y)) dy = A^2 \gamma, \end{aligned}$$

where we used (5.3.7). Recall the definition of  $\beta_\phi$  in (5.2.18), putting all these estimates together we obtain

$$\begin{aligned} \langle \xi_t, v_\phi \rangle_\phi &= \int_0^t [e^{-s} (1 - e^{-s}) (\alpha_\phi^\circ(h) \beta_\phi(v_\phi, l) + \alpha_\phi^\circ(l) \beta_\phi(v_\phi, h)) \\ &\quad + e^{-s} (e^{s\gamma} - e^{-s}) (\alpha_\phi^\gamma(h) \beta_\phi(u_\phi, l) + \alpha_\phi^\gamma(l) \beta_\phi(u_\phi, h)) \\ &\quad + (1 - e^{-s}) (e^{s\gamma} - e^{-s}) A^2 \gamma (\alpha_\phi^\circ(h) \alpha_\phi^\circ(l) + \alpha_\phi^\gamma(h) \alpha_\phi^\gamma(l)) + e^{-2s} \beta_\phi(h, l)] ds, \end{aligned}$$

so that

$$\begin{aligned} \lim_{t \rightarrow \infty} \langle \xi_t, v_\phi \rangle_\phi &= \frac{1}{2} (\alpha_\phi^\circ(h) \beta_\phi(v_\phi, l) + \alpha_\phi^\circ(l) \beta_\phi(v_\phi, h)) \\ &\quad + \frac{1+\gamma}{2(1-\gamma)} (\alpha_\phi^\gamma(h) \beta_\phi(u_\phi, l) + \alpha_\phi^\gamma(l) \beta_\phi(u_\phi, h)) \\ &\quad - A^2 \frac{(2-\gamma)(1+\gamma)}{2(1-\gamma)} (\alpha_\phi^\circ(h) \alpha_\phi^\circ(l) + \alpha_\phi^\gamma(h) \alpha_\phi^\gamma(l)) + \frac{1}{2} \beta_\phi(h, l). \end{aligned}$$

As  $D^2\theta(u_\phi)[h, l] = \frac{1}{A^2} \lim_{t \rightarrow \infty} \langle \xi_t, v_\phi \rangle_\phi$ , we obtain (5.2.17).  $\square$

## 5.4 Long time behavior (proofs)

The aim of this section is to prove Theorem 5.2.9.

### 5.4.1 Main structure of the proof of Theorem 5.2.9

First, fix some constant  $\eta$  such that

$$0 < \eta < \frac{1}{4}. \quad (5.4.1)$$

We also look for some  $T > 0$  that verifies

$$C_P C_{\mathcal{L}} e^{T\gamma} \leq 1/4, \quad (5.4.2)$$

where  $C_P$ ,  $C_{\mathcal{L}}$  and  $\gamma$  are introduced in Proposition 5.2.5. We first define the initial time  $T_0(N)$  thanks to the following Proposition, whose proof is postponed to Section 5.4.3.

**Proposition 5.4.1** (Initialisation). *In the framework of Theorem 5.2.9, there exists a deterministic phase  $\theta_0 \in S$ , an event  $B_N$  such that  $\mathbf{P}(B_N) \xrightarrow{N \rightarrow \infty} 1$  and a constant  $C > 0$  such that for all  $\varepsilon > 0$ , for  $N$  sufficiently large, on the event  $B_N$ , the projection  $\psi = \psi_0^N = \text{proj}(U_N(C \log N))$  is well defined and*

$$\|U_N(C \log N) - u_{\psi_0^N}\|_2 \leq \frac{N^{2\eta}}{\sqrt{N}}, \quad (5.4.3)$$

$$|\psi_0^N - \theta_0| \leq \varepsilon. \quad (5.4.4)$$

We define  $T_0(N)$  thanks to Proposition 5.4.1 by  $T_0(N) = C \log(N)$ . Define the time discretisation of the interval  $[T_0(N), N^\alpha \tau_f]$  into subintervalls of length  $T$ ,  $[T_n, T_{n+1}]$ : define  $n_f = \inf\{n \in \mathbb{N}, N^\alpha \tau_f \leq T_0(N) + nT\}$  and for  $n = 0, \dots, n_f - 1$ ,  $T_n = T_0(N) + nT$ . Let  $T_f(N) := T_{n_f}$ , by construction,  $T_f(N) \geq N^\alpha \tau_f$ . We prove in fact a more precise result that Theorem 5.2.9 as stated in Remark 5.2.10: we show that there exists some  $C > 0$  such that we have

$$\mathbf{P} \left( \sup_{t \in [T_0(N), T_f(N)]} \text{dist}_{L^2}(U_N(t), \mathcal{U}) \leq CN^{\eta-1/2} \right) \xrightarrow{N \rightarrow \infty} 1. \quad (5.4.5)$$

We focus on a process  $(V_n(t))_{n \in \llbracket 1, n_f \rrbracket, t \in [0, T]}$  that iteratively compares  $U_N$  and its projection on  $\mathcal{U}$  at each step. We ensure it is correctly defined in the next part, then we give the main proof before the proof of some technical results we also need.

**Discretization** In order to define the projection of  $U_N(T_n)$  into  $\mathcal{U}$ , following Proposition 5.2.7, we need to ensure that  $\text{dist}_{L^2}(U_N(T_n), \mathcal{U}) \leq \varpi$ . In order to do so, we introduce the stopping couple

$$(n_\tau, \tau) := \inf \{(n, t) \in \llbracket 1, n_f \rrbracket \times [0, T] : \text{dist}_{L^2}(U_N(T_{n-1} + t), \mathcal{U}) > \varpi\}, \quad (5.4.6)$$

where the infimum corresponds to the lexicographic order. We introduce then

$$\tau_n := \begin{cases} T & \text{if } n < n_\tau \\ \tau & \text{if } n \geq n_\tau. \end{cases} \quad (5.4.7)$$

The process we consider is then  $(U_N(T_{n \wedge n_\tau-1} + t \wedge \tau_n))_{n \in [1, n_f], t \in [0, T]}$ . The projection of this stopped process is well defined on the whole interval  $[T_0(N), T_f(N)]$  by construction, so that we can now define rigorously the random phases  $\phi_{n-1}$  for  $n = 1, \dots, n_f$  by

$$\phi_{n-1} := \text{proj}(U_N(T_{n \wedge n_\tau-1})). \quad (5.4.8)$$

The object of interest is then the process  $V_n(t)$  of  $L^2(S)$  defined for  $n = 1, \dots, n_f$  and  $t \in [0, T]$  by

$$V_n(t) := U_N(T_{n \wedge n_\tau-1} + t \wedge \tau_n) - u_{\phi_{n-1}}, \quad (5.4.9)$$

as (5.4.5) translates then into

**Proposition 5.4.2.** *There exists an event  $\Omega_N$  with  $\mathbf{P}(\Omega_N) \xrightarrow[N \rightarrow \infty]{} 1$  such that on  $\Omega_N$ ,*

$$\sup_{1 \leq n \leq n_f} \sup_{t \in [0, T]} \|V_n(t)\|_2 = O\left(\frac{N^{2\eta}}{\sqrt{N}}\right), \quad (5.4.10)$$

where the error is uniform on  $\Omega_N$ .

Here are the steps of the proof of Proposition 5.4.2.

**Step 1 -** We show that the process  $(V_n(t))_{n \in [1, n_f], t \in [0, T]}$  satisfies the mild equation

$$V_n(t) = e^{(t \wedge \tau_n) \mathcal{L}_{\phi_{n-1}}} V_n(0) + \int_0^{t \wedge \tau_n} e^{(t \wedge \tau_n - s) \mathcal{L}_{\phi_{n-1}}} R_n(s) ds + \zeta_n(t \wedge \tau_n) \quad (5.4.11)$$

where

$$\zeta_n(t) := \int_0^t e^{(t-s) \mathcal{L}_{\phi_{n-1}}} dM_N(s), \quad (5.4.12)$$

and

$$\begin{aligned} R_n(t) = \cos * & \left( y \mapsto V_n(t)(y)^2 \int_0^1 f''(u_{\phi_{n-1}}(y) + rV_n(t)(y)) (1-r) dr \right) \\ & + \left( \sum_{i,j=1}^N \frac{2\pi \cos(x_i - x_j)}{N} f(U_{N,j}(t-)) \mathbf{1}_{B_{N,i}} - \cos * f(U_N(t)) \right), \end{aligned} \quad (5.4.13)$$

where the notation  $*$  stands for the convolution  $f * g(x) = \int_{-\pi}^{\pi} f(x-y)g(y)dy$ . The rigorous meaning of (5.4.11) is given in Proposition 5.4.4, postponed to Section 5.4.2.



**Step 2** - We show a control of several terms of (5.4.11) with the following Proposition, whose proof is postponed to Section 5.B.1.

**Proposition 5.4.3** (Noise perturbation). *Define the event*

$$A_N := \left\{ \sup_{1 \leq n \leq n_f} \sup_{t \in [0, T]} \|\zeta_n(t)\|_2 \leq \frac{N^\eta}{\sqrt{N}} \right\}. \quad (5.4.14)$$

In the framework of Theorem 5.2.9,  $\mathbf{P}(A_N) \xrightarrow{N \rightarrow \infty} 1$ .

Now recall  $B_N$  from Proposition 5.4.1 and let  $\Omega_N := A_N \cup B_N$ , we have  $\mathbf{P}(\Omega_N) \xrightarrow{N \rightarrow \infty} 1$  with Propositions 5.4.3 and 5.4.1. For the rest of the proof, we place ourselves now on this event  $\Omega_N$ .

**Step 3** - Based on Steps 1 and 2 above, it remains to prove (5.4.10). We proceed by induction. We know (as  $\Omega_N \subset B_N$ ) that  $\|V_1(0)\| \leq N^{2\eta-1/2}$ . Suppose that  $\|V_n(0)\|_2 \leq N^{2\eta-1/2}$  for some  $n \geq 1$ . From the mild formulation satisfied by  $(V_n(t))$  seen in (5.4.11) we get

$$\|V_n(t)\|_2 = \left\| e^{(t \wedge \tau_n) \mathcal{L}_{\phi_{n-1}}} V_n(0) \right\|_2 + \left\| \int_0^{t \wedge \tau_n} e^{(t \wedge \tau_n - s) \mathcal{L}_{\phi_{n-1}}} R_n(s) ds \right\|_2 + \|\zeta_n(t \wedge \tau_n)\|_2.$$

By Proposition 5.2.7 and (5.4.8),  $P_{\phi_{n-1}, 0}(U_N(T_{n \wedge n_{\tau-1}}) - u_{\phi_{n-1}}) = 0$  hence  $V_n(0) = U_N(T_{n \wedge n_{\tau-1}}) - u_{\phi_{n-1}} = P_{\phi_{n-1}, s} V_n(0)$ . Proposition 5.2.5 and more especially (5.2.12) give then, with the induction hypothesis

$$\|e^{(t \wedge \tau_n) \mathcal{L}_{\phi_{n-1}}} V_n(0)\|_{\phi_{n-1}} \leq e^{(t \wedge \tau_n) \gamma} \|V_n(0)\|_{\phi_{n-1}} \leq C_0 e^{(t \wedge \tau_n) \gamma} N^{2\eta - \frac{1}{2}},$$

where  $C_0$  is introduced in (5.2.11). From Proposition 5.2.5, we have

$$\left\| \int_0^{t \wedge \tau_n} e^{(t \wedge \tau_n - s) \mathcal{L}_{\phi_{n-1}}} R_n(s) ds \right\|_2 \leq TC_{\mathcal{L}} \sup_{0 \leq s \leq T} \|R_n(s)\|_2.$$

By definition of  $A_N$ ,  $\sup_{1 \leq n \leq n_f} \sup_{t \in [0, T]} \|\zeta_n(t)\|_2 \leq N^{\eta-1/2}$  as we are on  $\Omega_N$ . We obtain then, for any  $t \in [0, T]$

$$\|V_n(t)\|_2 \leq C_0 e^{(t \wedge \tau_n) \gamma} N^{2\eta - \frac{1}{2}} + TC_{\mathcal{L}} \sup_{0 \leq s \leq T} \|R_n(s)\|_2 + N^{\eta-1/2}. \quad (5.4.15)$$

For any  $t \in [0, T]$ , recalling (5.4.13),

$$\begin{aligned} & \sup_{0 \leq s \leq t} \|R_n(s)\|_2 \\ & \leq \sup_{0 \leq s \leq t} \left\| \cos * \left( y \mapsto V_n(s)(y)^2 \int_0^1 f''(u_{\phi_{n-1}}(y) + rV_n(s)(y)) (1-r) dr \right) \right\|_2 \\ & + \sup_{0 \leq s \leq t} \left\| \sum_{i,j=1}^N \frac{2\pi \cos(x_i - x_j)}{N} f(U_{N,j}(s-)) \mathbf{1}_{B_{N,i}} - \cos * f(U_N(s)) \right\|_2 = (A) + (B). \end{aligned} \quad (5.4.16)$$

Using Young's inequality  $\|u * v\|_2 \leq \|u\|_1 \|v\|_2$  and the boundedness of  $f''$ , we have

$$\begin{aligned} (A) &\leq \sup_{0 \leq s \leq t} \left( \|\cos\|_2 \int_S \left| V_n(s)(y)^2 \int_0^1 f''(u_{\phi_{n-1}}(y) + rV_n(s)(y)) (1-r) dr \right| dy \right) \\ &\leq C \sup_{0 \leq s \leq t} \|V_n(s)\|_2^2 \end{aligned}$$

for some positive  $C$ . For the second term (B) of (5.4.16), we introduce

$$\begin{aligned} \Upsilon_{1,i,s} &= \frac{2\pi}{N} \sum_{j=1}^N \cos(x_i - x_j) (f(U_{N,j}(s-)) - f(U_{N,j}(s))) \\ \Upsilon_{2,i,s} &= \frac{2\pi}{N} \sum_{j=1}^N \cos(x_i - x_j) f(U_{N,j}(s)) - \int_S \cos(x_i - y) f(U_N(s)(y)) dy \\ \Upsilon_{3,i,s}(x) &= \int_S (\cos(x_i - y) - \cos(x - y)) f(U_N(s)(y)) dy, \quad x \in S. \end{aligned} \quad (5.4.17)$$

From the Lipschitz continuity of  $f$  and the fact that the processes  $Z_{N,1}, \dots, Z_{N,N}$  do not jump simultaneously,  $|\Upsilon_{1,i,s}| \leq \frac{C}{N}$  hence  $\left\| \sum_{i=1}^N \Upsilon_{1,i,s} \mathbf{1}_{B_{N,i}} \right\|_2^2 = O\left(\frac{1}{N^2}\right)$ .

As  $\mathbf{1}_{B_{N,i}} \mathbf{1}_{B_{N,j}} \equiv 0$  for  $i \neq j$ , for any  $0 \leq s \leq t$  we have that  $\left\| \sum_{i=1}^N \Upsilon_{2,i,s} \mathbf{1}_{B_{N,i}} \right\|_2^2$  is equal to

$$\frac{2\pi}{N} \sum_{i=1}^N \left( \sum_{j=1}^N \int_{B_{N,j}} (\cos(x_i - x_j) - \cos(x_i - y)) f(U_N(s)(y)) dy \right)^2.$$

As  $f$  is bounded (by 1) and  $\cos$  is 1-Lipschitz continuous, we obtain

$$\left\| \sum_{i=1}^N \Upsilon_{2,i,s} \mathbf{1}_{B_{N,i}} \right\|_2^2 \leq \frac{2\pi}{N} \sum_{i=1}^N \left( \sum_{j=1}^N \int_{B_{N,j}} |x_j - y| dy \right)^2 \leq \frac{8\pi^5}{N^2}. \quad (5.4.18)$$

Similarly,

$$\begin{aligned} \left\| \sum_{i=1}^N \Upsilon_{3,i,s} \mathbf{1}_{B_{N,i}} \right\|_2^2 &= \int_S \sum_{i=1}^N \Upsilon_{3,i,s}(x)^2 \mathbf{1}_{B_{N,i}}(x) dx \\ &= \sum_{i=1}^N \int_{B_{N,i}} \left( \int_S (\cos(x_i - y) - \cos(x - y)) f(U_N(s)(y)) dy \right)^2 dx \\ &\leq \sum_{i=1}^N \int_{B_{N,i}} \left( \int_S |x_i - x| dy \right)^2 dx \leq \frac{8\pi^5}{N^2}. \end{aligned} \quad (5.4.19)$$

Hence we have for some positive  $C_{R,1}$

$$\sup_{0 \leq s \leq t} \|R_n(s)\|_2 \leq C_{R,1} \left( \sup_{0 \leq s \leq t} \|V_n(s)\|_2^2 + \frac{1}{N} \right). \quad (5.4.20)$$

Define then  $t^*$  as

$$t^* := \inf \left\{ t \in [0, T] : \|V_n(t)\|_2 \geq 2C_0 \frac{N^{2\eta}}{\sqrt{N}} \right\}. \quad (5.4.21)$$

Note that with no loss of generality, one can assume that  $C_0 > 1$ . Since by assumption  $\|V_n(0)\|_2 \leq \frac{N^{2\eta}}{\sqrt{N}} < C_0 \frac{N^{2\eta}}{\sqrt{N}}$ , we have  $\|V_n(t)\|_2 \leq 2C_0 \frac{N^{2\eta}}{\sqrt{N}}$  at least for  $t < t_1$  where  $t_1$  is the first jump among  $(Z_{N,1}, \dots, Z_{N,N})$ . Hence  $t^* > 0$ . If  $t \leq t^*$ ,  $\sup_{0 \leq s \leq t} \|R_n(s)\|_2 \leq C_{R,2} N^{4\eta-1}$  (as  $\eta > 0$ ,  $N^{-1} \ll N^{4\eta-1}$ ). Coming back to (5.4.15), we obtain that (for some positive constant  $C_R$ )

$$\|V_n(t)\|_2 \leq C_0 e^{(t \wedge \tau_n)\gamma} N^{2\eta-\frac{1}{2}} + TC_R N^{4\eta-1} + N^{\eta-1/2}. \quad (5.4.22)$$

Since  $0 < \eta < \frac{1}{4}$ ,  $N^{4\eta-1} \ll N^{2\eta-1/2}$  hence for  $N$  large enough  $TC_R N^{4\eta-1} + N^{\eta-1/2} \leq C_0 N^{2\eta-1/2}$  thus as  $\gamma < 0$ ,  $t^* = T$ . By construction of the stopping time  $\tau_n$  in (5.4.7), we have then that  $\tau_n = T$ , hence

$$\sup_{0 \leq t \leq T} \|V_n(t)\|_2 \leq 2C_0 N^{2\eta-1/2}. \quad (5.4.23)$$

To conclude the induction, we need to show that  $\|V_{n+1}(0)\|_2 \leq N^{2\eta-1/2}$ . By definition (5.4.9) and as  $\tau_n = T$ ,  $V_{n+1}(0) = U_N(T_n) - u_{\phi_n}$  and  $V_n(T) = U_N(T_n) - u_{\phi_{n-1}}$  hence  $V_{n+1}(0) = V_n(T) + u_{\phi_{n-1}} - u_{\phi_n}$ . Moreover, as  $V_{n+1}(0) = P_{\phi_n}^\perp V_{n+1}(0)$  since by definition  $V_{n+1}(0) \in \text{Ker}(\mathcal{L}_{\phi_n})^\perp$  (recall Proposition 5.2.5), we obtain

$$\begin{aligned} V_{n+1}(0) &= P_{\phi_n}^\perp (V_n(T) + u_{\phi_{n-1}} - u_{\phi_n}) \\ &= (P_{\phi_n}^\perp - P_{\phi_{n-1}}^\perp) V_n(T) + P_{\phi_{n-1}}^\perp V_n(T) + P_{\phi_n}^\perp (u_{\phi_{n-1}} - u_{\phi_n}). \end{aligned} \quad (5.4.24)$$

We are going to control each term of (5.4.24). First, using the smoothness of the phase projection from Proposition 5.2.7,

$$\begin{aligned} |\phi_{n-1} - \phi_n| &= |\text{proj}(U_N(T_{(n-1) \wedge n_{\tau-1}})) - \text{proj}(U_N(T_{n-1} \wedge n_{\tau-1}))| \\ &\leq C_{\text{proj}} \|U_N(T_{(n-1) \wedge n_{\tau-1}}) - U_N(T_{n-1} \wedge n_{\tau-1})\|_2 \\ &\leq C_{\text{proj}} \|V_{n-1}(0) - V_{n-1}(T)\|_2 \leq CN^{2\eta-1/2}, \end{aligned} \quad (5.4.25)$$

using (5.4.23). Recall (5.2.4) and (5.2.7), we have for any  $x \in S$

$$\begin{aligned} u_{\phi_{n-1}}(x) - u_{\phi_n}(x) &= A \cos(x + \phi_{n-1}) - A \cos(x + \phi_n) \\ &= -2A \sin(\phi_{n-1} - \phi_n) \sin\left(x + \phi_n + \frac{\phi_{n-1} - \phi_n}{2}\right) \\ &= 2 \sin(\phi_{n-1} - \phi_n) \left( \cos\left(\frac{\phi_{n-1} - \phi_n}{2}\right) v_{\phi_n}(x) - \sin\left(\frac{\phi_{n-1} - \phi_n}{2}\right) u_{\phi_n}(x) \right) \end{aligned}$$

thus, as  $P_{\phi_n}^\perp v_{\phi_n} = 0$ ,

$$P_{\phi_n}^\perp (u_{\phi_{n-1}} - u_{\phi_n}) = -2 \sin(\phi_{n-1} - \phi_n) \sin\left(\frac{\phi_{n-1} - \phi_n}{2}\right) P_{\phi_n}^\perp u_{\phi_n}.$$

As  $u_{\phi_n}$  is bounded and  $\sin$  is Lipschitz continuous, we obtain with (5.4.25) a control of the third term of (5.4.24)

$$\|P_{\phi_n}^\perp (u_{\phi_{n-1}} - u_{\phi_n})\|_2 \leq C (\phi_{n-1} - \phi_n)^2 = O(N^{4\eta-1}). \quad (5.4.26)$$

Similarly, recall (5.2.9),  $\phi \mapsto P_\phi^\perp$  is smooth, hence for some  $C > 0$

$$\|(P_{\phi_n}^\perp - P_{\phi_{n-1}}^\perp) V_n(T)\|_2 \leq C |\phi_{n-1} - \phi_n| \|V_n(T)\| = O(N^{4\eta-1}). \quad (5.4.27)$$

Combining (5.4.26) and (5.4.27) in (5.4.24), using (5.4.23) at time  $t = T$  and recalling Proposition 5.2.5, we obtain for  $N$  large enough

$$\|V_{n+1}(0)\|_2 \leq \|P_{\phi_{n-1}}^\perp V_n(T)\|_2 + O(N^{4\eta-1}) \leq 2C_P C_0 e^{T\gamma} N^{2\eta-1/2} + O(N^{4\eta-1}).$$

From the choice of  $T$  satisfying (5.4.2), the fact that  $\|V_{n+1}(0)\|_2 \leq N^{2\eta-1/2}$  follows and the recursion is concluded, so that Theorem 5.2.9 follows.

## 5.4.2 About the mild formulation

**Step 1** of Section 5.4.1 is a direct consequence of the following proposition.

**Proposition 5.4.4.** *Fix  $\phi \in S$  and  $0 < t_a < t_b$ . Recall the definition of  $U_N$  in (5.1.5), and define, for any  $t \in [t_a, t_b]$ ,*

$$\tilde{U}_{N,\phi}(t) = U_N(t) - u_\phi. \quad (5.4.28)$$

The process  $\left(\tilde{U}_{N,\phi}(t)\right)_{t \in [t_a, t_b]}$  satisfies the following semimartingale decomposition in  $D([t_a, t_b], L^2(S))$ , written in a mild form: for any  $t_a \leq t \leq t_b$

$$\tilde{U}_{N,\phi}(t) = e^{(t-t_a)\mathcal{L}_\phi} \tilde{U}_{N,\phi}(t_a) + \int_{t_a}^t e^{(t-s)\mathcal{L}_\phi} r_{N,\phi}(s) ds + \int_{t_a}^t e^{(t-s)\mathcal{L}_\phi} dM_N(s), \quad (5.4.29)$$

with

$$M_N(t) = \sum_{i=1}^N \sum_{j=1}^N \frac{2\pi \cos(x_i - x_j)}{N} \left( Z_{N,j}(t) - \int_0^t \lambda_{N,j}(s) ds \right) \mathbf{1}_{B_{N,i}} \quad (5.4.30)$$

and

$$\begin{aligned} r_{N,\phi}(t) = & \cos * \left( y \mapsto \tilde{U}_{N,\phi}(t)(y)^2 \int_0^1 f'' \left( u_\phi(y) + r \tilde{U}_{N,\phi}(t)(y) \right) (1-r) dr \right) \\ & + \left( \sum_{i,j=1}^N \frac{2\pi \cos(x_i - x_j)}{N} f(U_{N,j}(t-)) \mathbf{1}_{B_{N,i}} - \cos * f(U_N(t)) \right). \end{aligned} \quad (5.4.31)$$

*Proof of Proposition 5.4.4.* From (5.1.4), we obtain that  $U_N$  verifies

$$dU_N(t) = -U_N(t)dt + \sum_{i,j=1}^N \frac{2\pi \cos(x_i - x_j)}{N} dZ_{N,j}(t) \mathbf{1}_{B_{N,i}}. \quad (5.4.32)$$

The centered noise  $M_N$  defined in (5.4.30) verifies

$$dM_N(t) := \sum_{i=1}^N \sum_{j=1}^N \frac{2\pi \cos(x_i - x_j)}{N} (dZ_{N,j}(t) - f(U_{N,j}(t-))dt) \mathbf{1}_{B_{N,i}},$$

and is a martingale in  $L^2(S)$ . Thus recalling that  $u_\phi$  solves (5.2.1) and by inserting the terms  $\sum_{i=1}^N \sum_{j=1}^N \frac{2\pi \cos(x_i - x_j)}{N} f(U_{N,j}(t-))dt \mathbf{1}_{B_{N,i}}$  and  $u_\phi$ , we obtain

$$\begin{aligned} d\tilde{U}_{N,\phi}(t) &= -\tilde{U}_{N,\phi}(t)dt + dM_N(t) \\ &+ \left( \sum_{i,j=1}^N \frac{2\pi \cos(x_i - x_j)}{N} f(U_{N,j}(t-)) \mathbf{1}_{B_{N,i}} - \int_{-\pi}^{\pi} \cos(\cdot - y) f(u_\phi(y)) dy \right) dt. \end{aligned}$$

A Taylor's expansion gives that for any  $y \in S$ ,

$$\begin{aligned} f(U_N(t)(y)) - f(u_\phi(y)) &= f'(u_\phi(y)) \tilde{U}_{N,\phi}(t)(y) \\ &+ \int_0^1 f''(u_\phi(y) + r\tilde{U}_{N,\phi}(t)(y)) (1-r) dr \tilde{U}_{N,\phi}(t)(y)^2, \end{aligned}$$

hence identifying the operator  $\mathcal{L}_\phi$  defined in (5.2.6) we have

$$\begin{aligned} d\tilde{U}_{N,\phi}(t) &= \mathcal{L}_\phi \tilde{U}_{N,\phi}(t)dt + dM_N(t) \\ &+ \int_{-\pi}^{\pi} \cos(\cdot - y) \int_0^1 f''(u_\phi(y) + r\tilde{U}_{N,\phi}(t)(y)) (1-r) dr \tilde{U}_{N,\phi}(t)(y)^2 dy dt \\ &+ \left( \sum_{i,j=1}^N \frac{2\pi \cos(x_i - x_j)}{N} f(U_{N,j}(t-)) \mathbf{1}_{B_{N,i}} - \int_{-\pi}^{\pi} \cos(\cdot - y) f(U_N(t)(y)) dy \right) dt, \end{aligned}$$

and recognizing  $r_{N,\phi}$  defined in (5.4.31) we have

$$d\tilde{U}_{N,\phi}(t) = \mathcal{L}_\phi \tilde{U}_{N,\phi}(t)dt + r_{N,\phi}(t)dt + dM_N(t). \quad (5.4.33)$$

Then the mild formulation (5.4.29) is a direct consequence of (Zhu *et al.*, 2017, Lemma 3.2): the unique strong solution to (5.4.33) is indeed given by (5.4.29).  $\square$

### 5.4.3 About the initialisation

We prove here Proposition 5.4.1, that we use to define the initial time  $T_0(N)$  and in the second part of **Step 2** of Section 5.4.1.

*Proof of Proposition 5.4.1.* To prove Proposition 5.4.1, we proceed in several steps, as done in (Luçon & Poquet, 2017, Proposition 2.9).

- a. We rely on the convergence in finite time of  $U_N$  to its large population limit, that is  $u_t$  solving (5.1.16) with initial condition  $\rho$ . From the deterministic behavior of  $u_t$  and the stability of  $\mathcal{U}$ ,  $U_N$  approaches  $\mathcal{U}$  in a  $2\varepsilon_0$ -neighborhood; and this takes a time interval of order  $|\log \varepsilon_0|$ .

- b. We rely on the stability of  $\mathcal{U}$  and the control of the noise to show that, from a  $2\varepsilon_0$ -neighborhood,  $U_N$  approaches  $\mathcal{U}$  in a  $N^{2\eta-1/2}$ -neighborhood; and this takes a time interval of order  $\log N$ .
- c. We ensure that  $U_N$  stays at distance  $N^{2\eta-1/2}$  from  $\mathcal{U}$  at time  $T_0(N)$ .

**Step a.** We focus first on  $\psi_t(\rho)$ , solution to (5.1.16) with initial condition  $\rho \in B(\mathcal{U}, \varepsilon_0)$ . Thanks to Corollary 5.2.6, we have that it converges as  $t \rightarrow \infty$  towards some  $u_{\theta_0} \in \mathcal{U}$ . Thus, there exists a time  $s_1 \geq 0$  such that  $\|u_{s_1} - u_{\theta_0}\|_2 \leq \varepsilon_0$ , and this time is of order  $\frac{1}{\gamma} \log \varepsilon_0$ . We focus then on the random profile  $U_N$ . We use a mild formulation similar to the one used in Proposition 5.4.4: one can obtain, with  $u_t$  solving (5.1.16)

$$d(U_N(t) - u_t) = -(U_N(t) - u_t) dt + dM_N(t) + \left( \sum_{i,j=1}^N \frac{2\pi \cos(x_i - x_j)}{N} f(U_{N,j}(t-)) \mathbf{1}_{B_{N,i}} - \int_{-\pi}^{\pi} \cos(\cdot - y) f(u_t(y)) dy \right) dt,$$

where  $M_N$  is defined in (5.4.30). We have then for any  $t \geq 0$

$$U_N(t) - u_t = e^{-t} (U_N(0) - \rho) + \int_0^t e^{-(t-s)} dM_N(s) + \int_0^t e^{-(t-s)} r_N(s) ds$$

with

$$\begin{aligned} r_N(s) &:= \sum_i \mathbf{1}_{B_{N,i}} \sum_j \frac{2\pi \cos(x_i - x_j)}{N} (f(U_{N,j}(s-)) - f(U_{N,j}(s))) \\ &+ \sum_i \mathbf{1}_{B_{N,i}} \left( \sum_j \frac{2\pi \cos(x_i - x_j)}{N} f(U_{N,j}(s)) - \int_{-\pi}^{\pi} \cos(x_i - y) f(U_N(s)(y)) dy \right) \\ &+ \sum_i \mathbf{1}_{B_{N,i}} \int_{-\pi}^{\pi} (\cos(x_i - y) - \cos(\cdot - y)) f(U_N(s)(y)) dy \\ &+ \int_{-\pi}^{\pi} \cos(\cdot - y) (f(U_N(s)(y)) - f(u_s(y))) dy \\ &= \sum_{i=1}^N \mathbf{1}_{B_{N,i}} (\Upsilon_{1,i,s} + \Upsilon_{2,i,s} + \Upsilon_{3,i,s}) + \Upsilon_{4,s}. \end{aligned} \quad (5.4.34)$$

As done for  $\Upsilon_{1,i,s}$ ,  $\Upsilon_{2,i,s}$  and  $\Upsilon_{3,i,s}$  (5.4.17) in Proposition 5.4.2, we have for some  $C > 0$

$$\left\| \sum_{i=1}^N \mathbf{1}_{B_{N,i}} (\Upsilon_{1,i,s} + \Upsilon_{2,i,s} + \Upsilon_{3,i,s}) \right\|_2^2 \leq \frac{C}{N^2}.$$

Moreover an immediate computation gives, as  $f$  is Lipschitz continuous

$$\|\Upsilon_{4,s}\|_2 \leq C \|U_N(s) - u_s\|_2.$$

Then we have for any  $t \in [0, s_1]$  with  $\zeta_N(s) := \int_0^s e^{-(s-u)} dM_N(u)$ ,

$$\|U_N(t) - u_t\|_2 \leq \|U_N(0) - \rho\|_2 + \|\zeta_N(t)\|_2 + \frac{C}{N} + \int_0^t e^{-(t-s)} \|U_N(s) - u_s\|_2 ds. \quad (5.4.35)$$

Take  $N$  sufficiently large so that  $\|U_N(0) - \rho\|_2 \leq \frac{\varepsilon_0}{2}$ . We place ourselves on the event

$$C_N := \left\{ \sup_{t \in [0, s_1]} \|\zeta_N(t)\|_2 \leq N^{\eta-1/2} \right\}. \quad (5.4.36)$$

As done in Proposition 5.4.3,  $\mathbf{P}(C_N) \xrightarrow{N \rightarrow \infty} 1$ . Going back to (5.4.35), we have on  $C_N$

$$\|U_N(t) - u_t\|_2 \leq \frac{\varepsilon_0}{2} + N^{\eta-1/2} + \frac{C}{N} + \int_0^t e^{-(t-s)} \|U_N(s) - u_s\|_2 ds.$$

We deduce with Grönwall lemma that for  $N$  large enough,  $\|U_N(s_1) - u_{s_1}\|_2 \leq \varepsilon_0$  on  $C_N$ , which means that  $\|U_N(s_1) - u_{\theta_0}\|_2 \leq 2\varepsilon_0$  hence  $\text{dist}(U_N(s_1), \mathcal{U}) \leq 2\varepsilon_0$ . Choosing  $\varepsilon_0$  small enough so that  $2\varepsilon_0 < \varpi$  (recall Proposition 5.2.7), we can define  $\psi_0^1 = \text{proj}(U_N(s_1))$  and  $|\psi_0^1 - \theta_0| \leq C\varepsilon_0$ .

**Step b.** Since we know that  $\text{dist}_{L^2}(U_N(s_1), \mathcal{U}) \leq 2\varepsilon_0$  with increasing probability as  $N \rightarrow \infty$ , we show that  $U_N$  approaches  $\mathcal{U}$  up to a distance  $N^{2\eta-1/2}$  doing a similar iteration as in Proposition 5.4.2. Define the sequence  $(h_n)$  such that  $h_1 = 2\varepsilon_0$  and  $h_{n+1} = h_n/2$ , and let  $\tilde{n}_f := \inf \{n \geq 1, h_n \leq N^{2\eta-1/2}\}$ . Note that such  $\tilde{n}_f$  is of order  $O(\log N)$ . Fix  $\tilde{T}$  satisfying

$$C_P C_0 e^{\tilde{T}\gamma} \leq 1/4, \quad (5.4.37)$$

and define then for any  $n \in \llbracket 1, \tilde{n}_f \rrbracket$  the times  $\tilde{T}_n = s_1 + (n-1)\tilde{T}$ . As in (5.4.6) and (5.4.7), define

$$(\tilde{n}_\tau, \tilde{\tau}) := \inf \left\{ (n, t) \in \llbracket 1, \tilde{n}_f \rrbracket \times [0, \tilde{T}] : \text{dist}_{L^2}(U_N(\tilde{T}_{n-1} + t), \mathcal{U}) > \varpi \right\}, \quad (5.4.38)$$

and

$$\tilde{\tau}_n := \begin{cases} \tilde{T} & \text{if } n < \tilde{n}_\tau \\ \tilde{\tau} & \text{if } n \geq \tilde{n}_\tau. \end{cases} \quad (5.4.39)$$

The process we consider is then  $\left( U_N \left( \tilde{T}_{n \wedge \tilde{n}_\tau - 1} + t \wedge \tilde{\tau}_n \right) \right)_{n \in \llbracket 1, \tilde{n}_f \rrbracket, t \in [0, \tilde{T}]}$ , which is exactly  $(U_N(t))_{t \in [s_1, \tilde{T}_{\tilde{n}_f}]}$  unless the process has been stopped. The projection of this stopped process is well defined on the whole interval, so that we can now define rigorously the random phases  $\tilde{\phi}_{n-1}$  for  $n = 1, \dots, \tilde{n}_f$  by

$$\tilde{\phi}_{n-1} := \text{proj}(U_N(\tilde{T}_{n \wedge \tilde{n}_\tau - 1})). \quad (5.4.40)$$

The object of interest is then the process  $\tilde{V}_n(t)$  of  $L^2(S)$  defined for  $n = 1, \dots, \tilde{n}_f$  and  $t \in [0, \tilde{T}]$  by

$$\tilde{V}_n(t) := U_N(\tilde{T}_{n \wedge \tilde{n}_\tau - 1} + t \wedge \tilde{\tau}_n) - u_{\tilde{\phi}_{n-1}}. \quad (5.4.41)$$

It satisfies the mild equation

$$\tilde{V}_n(t) = e^{(t \wedge \tilde{\tau}_n) \mathcal{L}_{\tilde{\phi}_{n-1}}} \tilde{V}_n(0) + \int_0^{t \wedge \tilde{\tau}_n} e^{(t \wedge \tilde{\tau}_n - s) \mathcal{L}_{\tilde{\phi}_{n-1}}} \tilde{R}_n(s) ds + \tilde{\zeta}_n(t \wedge \tilde{\tau}_n) \quad (5.4.42)$$

where

$$\tilde{\zeta}_n(t) := \int_0^t e^{(t-s) \mathcal{L}_{\tilde{\phi}_{n-1}}} dM_N(s), \quad (5.4.43)$$

and

$$\begin{aligned} \tilde{R}_n(t) = & \cos * \left( y \mapsto \tilde{V}_n(t)(y)^2 \int_0^1 f'' \left( u_{\tilde{\phi}_{n-1}}(y) + r \tilde{V}_n(t)(y) \right) (1-r) dr \right) \\ & + \left( \sum_{i,j=1}^N \frac{2\pi \cos(x_i - x_j)}{N} f(U_{N,j}(t-)) \mathbf{1}_{B_{N,i}} - \cos * f(U_N(t)) \right). \end{aligned} \quad (5.4.44)$$

Define the event

$$B_N := C_N \cap \left\{ \sup_{n \in \llbracket 1, \tilde{n}_f \rrbracket} \sup_{t \in [0, \tilde{T}]} \|\tilde{\zeta}_n(t)\|_2 \leq N^{\eta-1/2} \right\}. \quad (5.4.45)$$

As done in Proposition 5.4.3,  $\mathbf{P}(B_N) \rightarrow 1$  and from now on we work under  $B_N$ . We want to show by induction that on  $B_N$ , for all  $n \in \llbracket 1, \tilde{n}_f \rrbracket$ ,  $\tilde{V}_n(0) \leq h_n$ . The first step of the proof ensures that on  $C_N$ ,  $\tilde{V}_1(0) \leq h_1$ . Assume for some  $n < \tilde{n}_f$ ,  $\tilde{V}_n(0) \leq h_n$ . From the mild formulation (5.4.42) we obtain (as done in (5.4.22))

$$\|\tilde{V}_n(t)\|_2 \leq C_0 e^{(t \wedge \tilde{\tau}_n) \gamma} h_n + \tilde{T} C_{\mathcal{L}} \sup_{0 \leq s \leq \tilde{T}} \|\tilde{R}_n(s)\|_2 + N^{\eta-1/2}. \quad (5.4.46)$$

Define then  $t^*$  as

$$\tilde{t}^* := \inf \left\{ t \in [0, \tilde{T}] : \|\tilde{V}_n(t)\|_2 \geq 2C_0 h_n \right\}. \quad (5.4.47)$$

We have  $\tilde{t}^* > 0$ , and if  $t \leq \tilde{t}^*$ ,  $\sup_{0 \leq s \leq t} \|\tilde{R}_n(s)\|_2 \leq C_{R_2}(h_n^2 + N^{-1})$ , as done in (5.4.20). Coming back to (5.4.46), we obtain that (for some positive constant  $C_{\tilde{R}}$ )

$$\|\tilde{V}_n(t)\|_2 \leq C_0 e^{(t \wedge \tilde{\tau}_n) \gamma} h_n + \tilde{T} C_{\tilde{R}}(h_n^2 + N^{-1}) + N^{\eta-1/2}. \quad (5.4.48)$$

Since  $n < \tilde{n}_f$ ,  $2\varepsilon_0 \geq h_n > N^{2\eta-1/2}$  hence for  $N$  large enough,  $N^{\eta-1/2}$ ,  $N^{-1}$  are negligible with respect to  $h_n$ , same for  $h_n^2$  thus  $\tilde{t}^* \geq \tilde{T}$ . To conclude the induction, we need to show that  $\|\tilde{V}_{n+1}(0)\| \leq h_{n+1} = \frac{h_n}{2}$ . As shown in (5.4.24),  $\tilde{V}_{n+1}(0) = \left( P_{\tilde{\phi}_n}^\perp - P_{\tilde{\phi}_{n-1}}^\perp \right) \tilde{V}_n(\tilde{T}) + P_{\tilde{\phi}_{n-1}}^\perp \tilde{V}_n(\tilde{T}) + P_{\tilde{\phi}_n}^\perp \left( u_{\tilde{\phi}_{n-1}} - u_{\tilde{\phi}_n} \right)$ . From the similar controls (5.4.26) and (5.4.27) and using (5.4.48) for  $t = \tilde{T}$ , we have for  $N$  large enough,

$$\|\tilde{V}_{n+1}(0)\|_2 \leq \|P_{\tilde{\phi}_{n-1}}^\perp \tilde{V}_n(\tilde{T})\|_2 + O(h_n^2) \leq 2C_P C_0 e^{\tilde{T} \gamma} h_n + O(h_n^2).$$

Recall (5.4.37) and  $\gamma < 0$ , the fact that  $\|\tilde{V}_{n+1}(0)\|_2 \leq h_{n+1}$  follows then and the iteration is concluded. Thus, we have constructed a time  $s_2 = s_1 + (\tilde{n}_f - 1)\tilde{T}$  such that, on  $B_N$  for  $N$  large enough, setting  $\psi_0^2 := \text{proj}(U_N(s_2))$ , we have  $\|U_N(s_2) - u_{\psi_0^2}\|_2 \leq N^{2\eta-1/2}$  and  $|\psi_0^2 - \psi_0^1| \leq C\varepsilon_0$ , which gives  $|\psi_0^2 - \theta_0| \leq C'\varepsilon_0$  for some  $C' > 0$ .



**Step c.** So far, we have constructed a time  $s_2 = C(|\log \varepsilon_0| + \log N)$  for which we have  $\text{dist}_{L^2}(U_N(s_2), \mathcal{U}) \leq N^{2\eta-1/2}$ . We want some  $s_3 = \tilde{C} \log N \geq s_2$ ,  $\tilde{C} = C + 1$ , independent of  $\varepsilon_0$  such that with  $\psi_0^3 := \text{proj}(U_N(s_3))$ ,  $\|U_N(s_3) - u_{\psi_0^3}\| \leq N^{2\eta-1/2}$ . For this, it suffices to decompose the dynamics on  $[s_2, s_3]$  in a same way as before in both Steps 1 and 2. This induces a drift  $|\psi_0^3 - \psi_0^2| \leq CN^{2\eta-1/2} \log(N) \leq \varepsilon_0$  for  $N$  large enough. This last step concludes the proof with  $T_0(N) = s_3$ .  $\square$

## 5.5 Fluctuations on the manifold (proofs)

The aim of this section is to prove Theorem 5.2.11. We start by giving an auxiliary lemma.

**Lemma 5.5.1.** *There exists some  $C > 0$  such that for any  $g \in B(\mathcal{U}, \varepsilon_0)$ ,*

$$\text{dist}_{L^2}(g, \mathcal{U}) \leq \|g - u_{\theta(g)}\|_2 \leq C \text{dist}_{L^2}(g, \mathcal{U}).$$

*Proof.* Let  $g \in B(\mathcal{U}, \varepsilon_0)$ . The first inequality directly comes from the definition of  $\text{dist}_{L^2}(g, \mathcal{U})$ . By compactness of  $\mathcal{U}$ , there exists some  $y \in \mathcal{U}$  such that  $\text{dist}_{L^2}(g, \mathcal{U}) = \|g - y\|_2$  (and  $y = u_{\theta(y)}$ ). Then

$$\|g - u_{\theta(g)}\|_2 \leq \|g - y\|_2 + \|u_{\theta(y)} - u_{\theta(g)}\|_2,$$

and as  $\phi \mapsto u_\phi$  and  $\theta$  are Lipschitz continuous (recall  $u_\phi = A \cos(\cdot + \phi)$  and  $\theta$  is  $\mathcal{C}^2$  from Proposition 5.2.8),  $\|u_{\theta(y)} - u_{\theta(g)}\|_2 \leq \hat{C} \|g - y\|_2$  for some  $\hat{C} > 0$  (independent of the choice of  $g$ ).  $\square$

### 5.5.1 Main structure of the proof of Theorem 5.2.11

First, Theorem 5.2.9 and Lemma 5.5.1 give that one can find an event  $\Omega_N$  such that  $\mathbf{P}(\Omega_N) \xrightarrow{N \rightarrow \infty} 1$  and on this event

$$\sup_{t \in [T_0(N), T_f(N)]} \|U_N(t) - u_{\theta(U_N(t))}\|_2 = O(N^{\eta-1/2}), \quad (5.5.1)$$

with  $T_0(N) = C \log(N)$  and  $T_f(N) = N\tau_f$ . It remains to study the behavior of the isochron map of the process, that is  $\theta(U_N(t))$ . We do a change of variables and introduce  $\tau_0(N) := \frac{T_0(N)}{N}$ , we define for any  $\tau \in [\tau_0(N), \tau_f]$  the rescaled process

$$\hat{\theta}_N(\tau) = \theta(U_N(N\tau)). \quad (5.5.2)$$

In the proof, we keep the notation  $t$  for the microscopic time variable, that is when  $t \in [T_0(N), T_f(N)]$  and  $\tau$  for the macroscopic time variable, when  $\tau \in [\tau_0(N), \tau_f]$ . Theorem 5.2.11 relies on the following decomposition of  $\hat{\theta}_N$ , obtained by Itô's lemma.

**Proposition 5.5.2.** *For any initial condition  $\tau_0 \geq \tau_0(N)$ , for any  $\tau \geq \tau_0$ ,  $\widehat{\theta}_N(\tau)$  can be written as*

$$\widehat{\theta}_N(\tau) = \widehat{\theta}_N(\tau_0) + \vartheta_N(\tau_0, \tau) + \Theta_N(\tau_0, \tau), \quad (5.5.3)$$

where

$$\sup_{\tau_0(N) \leq \tau_0 \leq \tau \leq \tau_f} \mathbf{E} (|\vartheta_N(\tau_0, \tau)|) \xrightarrow{N \rightarrow \infty} 0$$

and  $\Theta_N(\tau_0, \tau)$  is a real martingale with quadratic variation

$$[\Theta_N]_\tau = \frac{1}{N} \sum_{j=1}^N \int_{\tau_0}^{\tau} \Phi(x_j, \widehat{\theta}_N(s)) f(u_{\widehat{\theta}_N(s)}(x_j)) ds \quad (5.5.4)$$

with

$$\Phi(x, \theta) := 4\pi^2 \sin^2(x + \theta). \quad (5.5.5)$$

The proof of Proposition 5.5.2 is postponed to Section 5.5.2. The remaining of the proof of Theorem 5.2.11 is to prove the tightness of  $(\widehat{\theta}_N(t))$  and to identify its limit. We apply Aldous criterion: note first that for any  $\tau \in [\varepsilon, \tau_f]$ ,  $\widehat{\theta}_N(\tau) \in S$  a compact set. Let  $(\tau_N)_N$  be a bounded sequence of  $\widehat{\theta}_N$ -optional times, let  $(h_N)$  be a sequence of positive constants such that  $h_N \rightarrow 0$ . From Proposition 5.5.2, we have

$$\widehat{\theta}_N(\tau_N + h_N) - \widehat{\theta}_N(\tau_N) = \vartheta_N(\tau_N, \tau_N + h_N) + \Theta_N(\tau_N, \tau_N + h_N),$$

where  $\vartheta_N(\tau_N, \tau_N + h_N) \xrightarrow{L^1, N \rightarrow \infty} 0$  and  $\Theta_N$  has the quadratic variation

$$[\Theta_N]_{\tau_N + h_N} = \frac{1}{N} \sum_{j=1}^N \int_{\tau_N}^{\tau_N + h_N} \Phi(x_j, \widehat{\theta}_N(s)) f(u_{\widehat{\theta}_N(s)}(x_j)) ds.$$

Using Burkholder-Davis-Gundy inequality, as  $\Phi$  and  $f$  are bounded, we have that

$$\mathbf{E} [\Theta_N(\tau_N, \tau_N + h_N)^2] \leq C \mathbf{E} [[\Theta_N]_{\tau_N + h_N}] \leq Ch_N$$

for some positive constants  $C$ . We obtain then that  $\widehat{\theta}_N(\tau_N + h_N) - \widehat{\theta}_N(\tau_N) \xrightarrow{L^1, N \rightarrow \infty} 0$  hence the convergence in probability: for all  $\varepsilon > 0$ ,

$$\mathbf{P} \left( \left| \widehat{\theta}_N(\tau_N + h_N) - \widehat{\theta}_N(\tau_N) \right| > \varepsilon \right) \xrightarrow{N \rightarrow \infty} 0.$$

We can then use Aldous criterion (see (Billingsley, 1968, Theorem 16.8)): the process  $(\tau \in [\varepsilon_N, \tau_f] \mapsto \widehat{\theta}_N(\tau))_N$  is tight. Let  $\tau \mapsto \widehat{\theta}(\tau)$  be a limit in distribution of any subsequence of  $(\tau \mapsto \widehat{\theta}_N(\tau))_N$  (by convenience renamed  $\widehat{\theta}_N$ ), that is  $\widehat{\theta}_N \xrightarrow[N \rightarrow \infty]{law} \widehat{\theta}$ . By Skorokhod's representation theorem, we can represent this convergence on a common probability space such that  $\widehat{\theta}_N \xrightarrow[N \rightarrow \infty]{a.s.} \widehat{\theta}$ . Using this in (5.5.4), we obtain that for any  $\tau \in [0, \tau_f]$ , as  $N$  goes to infinity, the quadratic variation of  $\widehat{\theta}$  is

$$[\widehat{\theta}]_\tau = 2\pi \int_0^\tau \int_S \sin^2(x + \widehat{\theta}(s)) f(A \cos(x + \widehat{\theta}(s))) dx ds = \sigma^2 \tau,$$

with  $\sigma$  defined in (5.2.22). We conclude by Lévy's characterization theorem and obtain (5.2.21).

### 5.5.2 About the decomposition of Proposition 5.5.2

*Proof of Proposition 5.5.2.* To show (5.5.3), we study  $(\theta(U_N(t)))_{t \in [T_0(N), T_f(N)]}$ . To simplify the notations, we introduce

$$\theta_N(t) := \theta(U_N(t)). \quad (5.5.6)$$

Note that from the decomposition (5.4.29) of  $U_N(t)$  and the definition  $M_N(t)$  in (5.4.30), one can write

$$dU_N(t) = B_N(t)dt + dM_N(t)$$

where  $B_N(t) := -U_N(t) + \cos * f(U_N(t)) + \Upsilon_t$ , with

$$\begin{aligned} \Upsilon_t(x) = \sum_{i=1}^N \left( \frac{2\pi}{N} \sum_{j=1}^N \cos(x_i - x_j) f(U_{N,j}(t-)) \right. \\ \left. - \int_S \cos(x - y) f(U_N(t)(y)) dy \right) \mathbf{1}_{B_{N,i}}(x). \end{aligned} \quad (5.5.7)$$

The starting point is to write the semimartingale decomposition of  $\theta(U_N(t))$  from Itô formula:

$$\begin{aligned} \theta(U_N(t)) &= \theta(U_N(t_0)) + \int_{t_0}^t D\theta(U_N(s-)) [-U_N(s) + \cos * f(U_N(s-))] ds \\ &\quad + \int_{t_0}^t D\theta(U_N(s-)) \Upsilon_s ds + \int_{t_0}^t D\theta(U_N(s-)) [dM_N(s)] \\ &+ \sum_{j=1}^N \int \int [\theta(U_N(s-) + \chi_j(s, z)) - \theta(U_N(s-)) - D\theta(U_N(s-)) [\chi_j(s, z)]] \pi_j(ds, dz) \\ &=: \theta(U_N(t_0)) + I_1^N(t_0, t) + I_2^N(t_0, t) + I_3^N(t_0, t) + I_4^N(t_0, t). \end{aligned} \quad (5.5.8)$$

We are going to focus on each of the terms of (5.5.8), that is  $I_k^N(t_0, t)$  for  $k \in \{1, 2, 3, 4\}$ . We have the following lemmas.

**Lemma 5.5.3.** *We have*

$$\sup_{t_0 \in [T_0(N), T_f(N)]} \sup_{t \in (t_0, T_f(N))} |I_1^N(t_0, t)| \xrightarrow{N \rightarrow \infty} 0 \quad (5.5.9)$$

*in probability.*

**Lemma 5.5.4.** *We have*

$$\sup_{t_0 \in [T_0(N), T_f(N)]} \sup_{t \in (t_0, T_f(N))} |I_2^N(t_0, t)| \xrightarrow{N \rightarrow \infty} 0 \quad (5.5.10)$$

*in probability.*

**Lemma 5.5.5.** For any  $t_0, t \in [T_0(N), T_f(N)]$ ,  $t_0 \leq t$ , we have

$$I_3^N(t_0, t) = \tilde{\Theta}_N(t_0, t) + J_3^N(t_0, t) \quad (5.5.11)$$

where  $\sup_{s \in (t_0, T_f(N))} \mathbf{E}(|J_3^N(t_0, s)|) \xrightarrow{N \rightarrow \infty} 0$  and  $\tilde{\Theta}_N$  is a real martingale with quadratic variation

$$[\tilde{\Theta}_N]_t = \frac{1}{N^2} \sum_{j=1}^N \int_{t_0}^t \Phi(x_j, \theta(U_N(s-))) f(u_{\theta(U_N(s-))}(x_j)) ds$$

with  $\Phi$  defined in (5.5.5).

**Lemma 5.5.6.** We have

$$\sup_{t_0 \in [T_0(N), T_f(N)]} \sup_{t \in (t_0, T_f(N))} \mathbf{E}(|I_4^N(t_0, t)|) \xrightarrow{N \rightarrow \infty} 0.$$

The proofs of these four lemmas are postponed to Section 5.5.3. Combining them, we can define some random variable  $J_N(t_0, t)$  such that

$$\sup_{s \in (t_0, T_f(N))} \mathbf{E}(|J^N(t_0, s)|) \xrightarrow{N \rightarrow \infty} 0$$

and for any  $t_0, t \in [T_0(N), T_f(N)]$ ,  $t_0 \leq t$ ,

$$\theta(U_N(t)) = \theta(U_N(t_0)) + J^N(t_0, t) + \tilde{\Theta}_N(t_0, t).$$

Recall the change of variables used to define  $\hat{\theta}$  in (5.5.2). Define similarly the quantities  $\vartheta_N(\tau_0, \tau) := J^N(N\tau_0, N\tau)$  and  $\Theta_N(\tau_0, \tau) = \tilde{\Theta}_N(N\tau_0, N\tau)$  for  $\tau_0 = t_0/N$  and  $\tau = t/N$ . Then we have exactly shown (5.5.3).  $\square$

### 5.5.3 Control of the terms of the decomposition

For simplicity, we may write  $I_k(t)$  instead of  $I_k^N(t_0, t)$ . In the following, we use the notations  $g_s = O(\alpha_N)$  with  $g : s \in I \mapsto g_s \in L^2(S)$  for some time interval  $I$  and a sequence  $(\alpha_N)$  independent of the time  $s$  when there exists some  $C$  (independent of  $N$ ) such that for all  $x \in S$ ,  $\sup_{s \in I} \sup_{x \in S} |g_s(x)| \leq C\alpha_N$ . Recall the definition of  $\theta_N(t)$  in (5.5.6). In the following proofs, this notation will be essentially used for  $t = s-$ , so that we write for simplicity  $\theta_N = \theta_N(s-)$ .

#### 5.5.3.1 Proof of Lemma 5.5.3

Recall that

$$I_1(t) := \int_{t_0}^t D\theta(U_N(s-)) [-U_N(s) + \cos * f(U_N(s))] ds.$$

Define for  $g \in L^2(S)$

$$\mathcal{V}(g) := -g + \cos * f(g).$$

Recall that for any  $\phi \in S$ ,  $\mathcal{L}(u_\phi) = 0$  and  $D\mathcal{V}(u_\phi)[h] = \mathcal{L}_\phi h$ . Let  $g \in B(\mathcal{U}, \varepsilon_0)$ , and  $t \mapsto g_t := \psi_t(g)$  defined in (5.1.17), that is the flow of (5.1.16) under initial condition  $g$ . Note that by definition of the isochron map  $\theta$  in Proposition 5.2.8 and the fact that  $\mathcal{U}$  consists of stationary solutions to (5.1.16), one has that  $\theta(\psi_t(g)) = \theta(\psi_0(g)) = \theta(g)$ . Differentiating with respect to  $t$  (recall Proposition 5.2.8) gives that  $D\theta(g_t)[\partial_t g_t] = D\theta(g_t)[-g_t + \cos * f(g_t)] = 0$ . Since this is for all  $t \geq 0$ , taking  $t = 0$  gives  $D\theta(g)[-g + \cos * f(g)] = 0$ . Hence for any  $s$ ,  $D\theta(U_N(s))[\mathcal{V}(U_N(s))] = 0$  and as  $\mathcal{V}(u_{\theta(U_N(s))}) = 0$ , we have

$$\begin{aligned} I_1(t) &= \int_{t_0}^t D\theta(U_N(s-))[\mathcal{V}(U_N(s))]ds \\ &= \int_{t_0}^t (D\theta(U_N(s-)) - D\theta(U_N(s)))[\mathcal{V}(U_N(s))]ds \\ &= \int_{t_0}^t (D\theta(U_N(s-)) - D\theta(U_N(s)))[\mathcal{V}(U_N(s)) - \mathcal{V}(u_{\theta(U_N(s))})]ds. \end{aligned}$$

As  $\theta$  and  $\mathcal{V}$  are Lipschitz continuous, as from (5.1.4) a jump of the process gives a.s. at most an increment of  $\frac{2\pi}{N}$  between  $U_N(s-)$  and  $U_N(s)$ , using (5.5.1) there exists some  $C > 0$  (independent of  $N$  and of the time) such that

$$I_1(t) \leq (t - t_0)\|\theta\|_L \frac{2\pi}{N} \|\mathcal{V}\|_L \|U_N(s-) - u_{\theta(U_N(s))}\|_2 \leq \frac{CT_f(N)}{N} N^{\eta-1/2}$$

on the event  $\Omega_N$  (given by Theorem 5.2.9). As  $T_f(N) \propto N$  and from the choice on  $\eta$ , (5.5.9) follows.

### 5.5.3.2 Proof of Lemma 5.5.4

We place ourselves again on the event  $\Omega_N$  (given by Theorem 5.2.9) on which we have (5.5.1). Recall that  $I_2(t) := \int_{t_0}^t D\theta(U_N(s-)) \Upsilon_s ds$ , where the definition of  $\Upsilon$  is given in (5.5.7). We have

$$\begin{aligned} I_2(t) &= \int_{t_0}^t (D\theta(U_N(s-)) - D\theta(u_{\theta_N}))[\Upsilon_s] ds + \int_{t_0}^t D\theta(u_{\theta_N})[\Upsilon_s - \tilde{\Upsilon}_s] ds \\ &\quad + \int_{t_0}^t D\theta(u_{\theta_N})[\tilde{\Upsilon}_s] ds, \end{aligned} \quad (5.5.12)$$

with

$$\tilde{\Upsilon}_s(x) = \sum_{i=1}^N \left( \frac{2\pi}{N} \sum_{j=1}^N \cos(x_i - x_j) f(u_{\theta_N}(x_j)) - \int_S \cos(x - y) f(u_{\theta_N}(y)) dy \right) \mathbf{1}_{B_{N,i}}. \quad (5.5.13)$$

From (5.4.18) and (5.4.19) we have that  $\|\Upsilon_s\|_2 \leq \frac{C}{N}$  for some  $C > 0$  independent of  $N$  and  $s$ , thus, for the first term of (5.5.12), as done before using (5.5.1),

$$\begin{aligned} \int_{t_0}^t (D\theta(U_N(s-)) - D\theta(u_{\theta(U_N(s-))}))[\Upsilon_s] ds &\leq (t - t_0)C\|\theta\|_L N^{\eta-1/2} \frac{C}{N} \\ &\leq \frac{CT_f(N)}{N} N^{\eta-1/2}. \end{aligned}$$

For the third term of (5.5.12), using (5.2.16), we have  $D\theta(u_{\theta(U_N(s-))})[\tilde{\Upsilon}_s] = \frac{\langle v_{\theta(U_N(s-))}, \tilde{\Upsilon}_s \rangle_{\theta(U_N(s-))}}{\|v_{\theta(U_N(s-))}\|_{\theta(U_N(s-))}}$ . As shown in (5.3.2),  $\|v_{\theta_N}\|_{\theta_N} = A$ . From trigonometric formula one has

$$\begin{aligned}
 & \langle v_{\theta_N}, \tilde{\Upsilon}_s \rangle_{\theta_N} \\
 &= \langle v_{\theta_N}, \frac{2\pi}{N} \sum_{i,j=1}^N \cos(x_i - x_j) f(u_{\theta_N}(x_j)) \mathbf{1}_{B_{N,i}} - \int_S \cos(\cdot - y) f(u_{\theta_N}(y)) dy \rangle_{\theta_N} \\
 &= \left( \frac{2\pi}{N} \sum_{j=1}^N \cos(x_j + \theta_N) f(u_{\theta_N}(x_j)) \right) \left( \sum_{i=1}^N \cos(x_i + \theta_N) \langle v_{\theta_N}, \mathbf{1}_{B_{N,i}} \rangle_{\theta_N} \right) \\
 &+ \left( \frac{2\pi}{N} \sum_{j=1}^N \sin(x_j + \theta_N) f(u_{\theta_N}(x_j)) \right) \left( \sum_{i=1}^N \sin(x_i + \theta_N) \langle v_{\theta_N}, \mathbf{1}_{B_{N,i}} \rangle_{\theta_N} \right) \\
 &- \left( \int_S \cos(y + \theta_N) f(u_{\theta_N}(y)) dy \right) \langle v_{\theta_N}, \cos(\cdot + \theta_N) \rangle_{\theta_N} \\
 &- \left( \int_S \sin(y + \theta_N) f(u_{\theta_N}(y)) dy \right) \langle v_{\theta_N}, \sin(\cdot + \theta_N) \rangle_{\theta_N}. \tag{5.5.14}
 \end{aligned}$$

By invariance of rotation and with Lemma 5.3.1 we have  $\langle v_{\theta_N}, \cos(\cdot + \theta_N) \rangle_{\theta_N} = \mathcal{I}(\sin \cos) = 0$ , and similarly  $\int_S \sin(y + \theta_N) f(u_{\theta_N}(y)) dy = 0$ . We can then write (5.5.14) as  $\langle v_{\theta_N}, \tilde{\Upsilon}_s \rangle_{\theta_N} = A_1 A_2 + A_3 A_4$ . From the computations (5.B.11), (5.B.12), (5.B.13) and (5.B.14) of Lemma 5.B.4, we obtain that

$$\langle v_{\theta_N}, \tilde{\Upsilon}_s \rangle_{\theta_N} := \frac{\pi A^2}{N} + o\left(\frac{1}{N}\right). \tag{5.5.15}$$

For the second term of (5.5.12), we have with Lemma 5.3.1 that  $\langle v_{\theta_N}, \sin(\cdot + \theta_N) \rangle_{\theta_N} = -A\mathcal{I}(\sin^2) = -A$  thus

$$\begin{aligned}
 \langle v_{\theta_N}, \Upsilon_s - \tilde{\Upsilon}_s \rangle_{\theta_N} &= A_2 \left( \frac{2\pi}{N} \sum_{j=1}^N \cos(x_j + \theta_N) (f(U_N(s-)(x_j)) - f(u_{\theta_N}(x_j))) \right) \\
 &+ A_4 \left( \frac{2\pi}{N} \sum_{j=1}^N \sin(x_j + \theta_N) (f(U_N(s-)(x_j)) - f(u_{\theta_N}(x_j))) \right) \\
 &+ A \int_S \sin(y + \theta_N) f(U_N(s-)(y)) dy.
 \end{aligned}$$

Let us show that

$$D_N := \frac{2\pi}{N} \sum_{j=1}^N \cos(x_j + \theta_N) (f(U_N(s-)(x_j)) - f(u_{\theta_N}(x_j))) = O(N^{\eta-1/2}). \tag{5.5.16}$$

Setting  $\widehat{u}_{\theta_N}(y) := \sum_{k=1}^N u_{\theta_N}(x_k) \mathbf{1}_{y \in B_{N,k}}$ , we have

$$\begin{aligned} |D_N| &= \left| \sum_{j=1}^N \cos(x_j + \theta_N) \int_S (f(U_N(s-)(x_j)) - f(u_{\theta_N}(x_j))) \mathbf{1}_{y \in B_{N,j}} dy \right| \\ &\leq \|f\|_L \sum_{j=1}^N \int_S |U_N(s-)(y) - \widehat{u}_{\theta_N}(y)| \mathbf{1}_{y \in B_{N,j}} dy. \end{aligned}$$

With Cauchy–Schwarz inequality and Jensen’s discrete inequality, we have

$$\begin{aligned} |D_N| &\leq C_f \sum_{j=1}^N \left( \int_S |U_N(s-)(y) - \widehat{u}_{\theta_N}(y)|^2 \mathbf{1}_{y \in B_{N,j}} dy \right)^{1/2} \left( \int_S \mathbf{1}_{y \in B_{N,j}} dy \right)^{1/2} \\ &= \frac{\sqrt{2\pi N} C_f}{N} \sum_{j=1}^N \left( \int_S |U_N(s-)(y) - \widehat{u}_{\theta_N}(y)|^2 \mathbf{1}_{y \in B_{N,j}} dy \right)^{1/2} \\ &\leq \sqrt{2\pi N} C_f \left( \frac{1}{N} \sum_{j=1}^N \int_S |U_N(s-)(y) - \widehat{u}_{\theta_N}(y)|^2 \mathbf{1}_{y \in B_{N,j}} dy \right)^{1/2} \\ &= C\sqrt{N} \left( \frac{1}{N} \int_S |U_N(s-)(y) - \widehat{u}_{\theta_N}(y)|^2 dy \right)^{1/2} \\ &= C \|U_N(s-) - \widehat{u}_{\theta_N}\|_2 \leq C \|U_N(s-) - u_{\theta_N}\|_2 + C \|u_{\theta_N} - \widehat{u}_{\theta_N}\|_2, \end{aligned}$$

hence with (5.5.1) and as  $\|u_{\theta_N} - \widehat{u}_{\theta_N}\|_2 = O(1/N)$ , we have indeed shown that  $D_N = O(N^{\eta-1/2})$ . Similarly, one can show that

$$\frac{2\pi}{N} \sum_{j=1}^N \sin(x_j + \theta_N) (f(U_N(s-)(x_j)) - f(u_{\theta_N}(x_j))) = O(N^{\eta-1/2}).$$

Using Lemma 5.B.4 and as  $\int_S \sin(y + \theta_N) f(u_{\theta_N}(y)) dy = 0$ , we have

$$\begin{aligned} \langle v_{\theta_N}, \Upsilon_s - \widetilde{\Upsilon}_s \rangle_{\theta_N} &= O(N^{\eta-3/2}) + A \int_S \sin(y + \theta_N) (f(U_N(s-)(y)) - f(u_{\theta_N}(y))) dy \\ &\quad - A \frac{2\pi}{N} \sum_{j=1}^N \sin(x_j + \theta_N) (f(U_N(s-)(x_j)) - f(u_{\theta_N}(x_j))). \end{aligned}$$

Using Taylor’s expansion, we obtain

$$\langle v_{\theta_N}, \Upsilon_s - \widetilde{\Upsilon}_s \rangle_{\theta_N} = o\left(\frac{1}{N}\right) + A \Delta_N, \quad (5.5.17)$$

where

$$\begin{aligned} \Delta_N &= \int_S \sin(y + \theta_N) f'(u_{\theta_N}(y)) (U_N(s-)(y) - u_{\theta_N}(y)) dy \\ &\quad - \frac{2\pi}{N} \sum_{j=1}^N \sin(x_j + \theta_N) f'(u_{\theta_N}(x_j)) (U_N(s-)(x_j) - u_{\theta_N}(x_j)). \end{aligned}$$

Define  $\widehat{u}_{\theta_N}(y) := \sum_{j=1}^N u_{\theta_N}(x_j) \mathbf{1}_{B_{N,j}}(y)$ , we introduce it in  $\Delta_N$  so that

$$\begin{aligned} \Delta_N = \sum_{j=1}^N \int_{B_{N,j}} & [\sin(y + \theta_N) f'(u_{\theta_N}(y)) (U_N(s-)(y) - u_{\theta_N}(y)) \\ & - \sin(x_j + \theta_N) f'(u_{\theta_N}(x_j)) (U_N(s-)(y) - \widehat{u}_{\theta_N}(y))] dy, \end{aligned}$$

thus

$$\begin{aligned} \Delta_N = \sum_{j=1}^N \int_{B_{N,j}} & (\sin(y + \theta_N) f'(u_{\theta_N}(y)) - \sin(x_j + \theta_N) f'(u_{\theta_N}(x_j))) \cdot \\ & (U_N(s-)(y) - \widehat{u}_{\theta_N}(y)) dy + \sum_{j=1}^N \int_{B_{N,j}} \sin(y + \theta_N) f'(u_{\theta_N}(y)) (\widehat{u}_{\theta_N}(y) - u_{\theta_N}(y)) dy. \end{aligned} \quad (5.5.18)$$

For the first term of  $\Delta_N$ , let  $\alpha_N(y) := \sin(y + \theta_N) f'(u_{\theta_N}(y)) - \sin(x_j + \theta_N) f'(u_{\theta_N}(x_j))$ , one has with Cauchy–Schwarz inequality that

$$\begin{aligned} & \sum_{j=1}^N \int_S (U_N(s-)(y) - \widehat{u}_{\theta_N}(y)) \alpha_N(y) dy \\ & \leq \sum_{j=1}^N \left( \int_S (U_N(s-)(y) - \widehat{u}_{\theta_N}(y))^2 \mathbf{1}_{B_{N,j}} dy \right)^{1/2} \left( \int_{B_{N,j}} \alpha_N(y)^2 dy \right)^{1/2}. \end{aligned}$$

As  $\int_{B_{N,j}} \alpha_N(y)^2 dy \leq \int_{B_{N,j}} (y - x_j)^2 dy = O\left(\frac{1}{N^{3/2}}\right)$ , for some  $C > 0$ , using Jensen's inequality

$$\begin{aligned} & \sum_{j=1}^N \int_S (U_N(s-)(y) - \widehat{u}_{\theta_N}(y)) \alpha_N(y) dy \\ & \leq \frac{C}{\sqrt{N}} \frac{1}{N} \sum_{j=1}^N \left( \int_S (U_N(s-)(y) - \widehat{u}_{\theta_N}(y))^2 \mathbf{1}_{B_{N,j}} dy \right)^{1/2} \\ & \leq \frac{C}{\sqrt{N}} \sqrt{\frac{1}{N} \sum_{j=1}^N \int_S (U_N(s-)(y) - \widehat{u}_{\theta_N}(y))^2 \mathbf{1}_{B_{N,j}} dy} \\ & \leq \frac{C}{\sqrt{N}} \sqrt{\frac{1}{N} \|U_N(s-) - \widehat{u}_{\theta_N}\|_2^2}. \end{aligned}$$

As  $\|u_{\theta} - \widehat{u}_{\theta_N}\|_2^2 = O\left(\frac{1}{N^2}\right)$  and with (5.5.1), we obtain that the first term of



(5.5.18) is in  $O\left(\frac{N^{\eta-1/2}}{N}\right)$ . For the second term of  $\Delta_N$ , we have

$$\begin{aligned}
 & \sum_{j=1}^N \int_{B_{N,j}} \sin(y + \theta_N) f'(u_{\theta_N}(y)) (\widehat{u}_{\theta_N}(y) - u_{\theta_N}(y)) dy \\
 &= \sum_{j=1}^N \int_{B_{N,j}} \sin(y + \theta_N) f'(u_{\theta_N}(y)) (A \cos(x_j + \theta_N) - A \cos(y + \theta_N)) dy \\
 &= A \sum_{j=1}^N \int_{B_{N,j}} \sin(y + \theta_N) f'(u_{\theta_N}(y)) \sin(x_j + \theta_N) (y - x_j) dy + o\left(\frac{1}{N}\right) \\
 &= A \sum_{j=1}^N \sin(x_j + \theta_N)^2 f'(u_{\theta_N}(x_j)) \int_{B_{N,j}} (y - x_j) dy + o\left(\frac{1}{N}\right) \\
 &= A \sum_{j=1}^N \sin(x_j + \theta_N)^2 f'(u_{\theta_N}(x_j)) \left(-\frac{2\pi^2}{N^2}\right) + o\left(\frac{1}{N}\right) \\
 &= -\frac{\pi}{N} \int_S A \sin(y + \theta_N)^2 f'(u_{\theta_N}(y)) dy + o\left(\frac{1}{N}\right) = -\frac{A\pi}{N} + o\left(\frac{1}{N}\right).
 \end{aligned}$$

Coming back to (5.5.17), we have then that

$$\langle v_{\theta_N}, \Upsilon_s - \tilde{\Upsilon}_s \rangle_{\theta_N} = -\frac{\pi A^2}{N} + o\left(\frac{1}{N}\right). \quad (5.5.19)$$

This term (5.5.19) cancels with the previous computation (5.5.15) up to some rest of order  $o\left(\frac{1}{N}\right)$ . We obtain then (5.5.10) after integrating on  $(t_0, t)$  and using  $T_f(N) \propto N$ .

### 5.5.3.3 Proof of Lemma 5.5.5

Recall that  $I_3(t) := \int_{t_0}^t D\theta(U_N(s-)) [dM_N(s)]$ . Recall the definition of  $\chi_j$  in (5.B.2) and introduce the compensated measure  $\tilde{\pi}_j(ds, dz) := \pi_j(ds, dz) - \lambda_{N,j} ds dz$ , we can re-write the term  $I_3^N(t_0, t)$  and introduce  $D\theta(u_{\theta_N})$ :

$$\begin{aligned}
 I_3(t) &= \sum_{j=1}^N \int_{t_0}^t \int_0^\infty (D\theta(U_N(s-)) - D\theta(u_{\theta_N})) [\chi_j(s, z)] \tilde{\pi}_j(ds, dz) \\
 &\quad + \sum_{j=1}^N \int_{t_0}^t \int_0^\infty D\theta(u_{\theta_N}) [\chi_j(s, z)] \tilde{\pi}_j(ds, dz). \quad (5.5.20)
 \end{aligned}$$

Let us focus first on

$$Q_0(t) := \sum_{j=1}^N \int_{t_0}^t \int_0^\infty (D\theta(U_N(s-)) - D\theta(u_{\theta_N})) [\chi_j(s, z)] \tilde{\pi}_j(ds, dz).$$

It is a real martingale. We denote by  $[Q_0]_t = \sum_{s \leq t} |\Delta Q_0(t)|^2$  its quadratic variation. It is computed as follows (as the  $(\pi_j)_{1 \leq j \leq N}$  are independent, there are almost surely no simultaneous jumps so that  $[\tilde{\pi}_j, \tilde{\pi}_{j'}] = 0$  if  $j \neq j'$ ):

$$\begin{aligned} [Q_0]_t &= \sum_{j=1}^N \int_{t_0}^t \int_0^\infty ((D\theta(U_N(s-)) - D\theta(u_{\theta_N})) [\chi_j(s, z)])^2 \pi_j(ds, dz) \\ &\leq C \|\theta\|_L^2 \left( \sup_{t \in [t_0, t]} \|U_N(s-) - u_{\theta_N}\|_2 \right)^2 \sum_{j=1}^N \int_{t_0}^t \int_0^\infty \|\chi_j(s, z)\|_2^2 \pi_j(ds, dz) \\ &\leq CN^{2\eta-1} \frac{1}{N^2} \sum_{j=1}^N \int_{t_0}^t \int_0^\infty \mathbf{1}_{z \leq \lambda_{N,j}(s)} \pi_j(ds, dz), \end{aligned}$$

using (5.5.1) and the computation (5.B.6) for some constants  $C > 0$ . Then, by Burkholder-Davis-Gundy inequality and as  $f$  is bounded

$$\begin{aligned} \mathbf{E} [Q_0(t)^2] &\leq C \mathbf{E} [[Q_0]_t] \leq CN^{2\eta-3} \mathbf{E} \left[ \sum_{j=1}^N \int_{t_0}^t \int_0^\infty \mathbf{1}_{z \leq \lambda_{N,j}(s)} \pi_j(ds, dz) \right] \\ &\leq CN^{2\eta-1} \frac{T_f(N)}{N} \xrightarrow{N \rightarrow \infty} 0, \end{aligned}$$

hence  $Q_0(t)$  converges in  $L^1$  towards 0 as  $N \rightarrow \infty$  uniformly in  $t$ .

The other term  $Q(t) := \sum_{j=1}^N \int_{t_0}^t \int_0^\infty D\theta(u_{\theta_N(s-)}) [\chi_j(s, z)] \tilde{\pi}_j(ds, dz)$  in (5.5.20) is also a real martingale, we denote by  $[Q]_t = \sum_{s \leq t} |\Delta Q(t)|^2$  its quadratic variation and it is computed as follows:

$$\begin{aligned} [Q]_t &= \sum_{j=1}^N \int_{t_0}^t \int_0^\infty (D\theta(u_{\theta_N}) [\chi_j(s, z)])^2 \pi_j(ds, dz) \\ &= \sum_{j=1}^N \int_{t_0}^t \int_0^\infty \left( \frac{\langle v_{\theta_N}, \chi_j(s, z) \rangle_{\theta_N}}{\|v_{\theta_N}\|_{\theta_N}} \right)^2 \pi_j(ds, dz), \end{aligned}$$

where we used (5.2.16). Recall the notation  $w_{ij}^{(N)} = 2\pi \cos(x_i - x_j)$ , from the computation (5.3.2),  $\|v_{\theta_N}\|_{\theta_N} = A$  hence

$$\begin{aligned} [Q]_t &= \frac{1}{A^2} \sum_{j=1}^N \int_{t_0}^t \int_0^\infty \left( \langle v_{\theta_N}, \sum_{i=1}^N \mathbf{1}_{B_{N,i}} \frac{w_{ij}^{(N)}}{N} \mathbf{1}_{z \leq \lambda_{N,j}} \rangle_{\theta_N} \right)^2 \pi_j(ds, dz) \\ &= \frac{1}{A^2} \sum_{j=1}^N \int_{t_0}^t \int_0^\infty \left( \sum_{i=1}^N \frac{w_{ij}^{(N)}}{N} \langle v_{\theta_N}, \mathbf{1}_{B_{N,i}} \rangle_{\theta_N} \right)^2 \mathbf{1}_{z \leq \lambda_{N,j}} \pi_j(ds, dz). \quad (5.5.21) \end{aligned}$$

Let us focus on the term  $E_N := \sum_{i=1}^N \frac{w_{ij}^{(N)}}{N} \langle v_{\theta_N}, \mathbf{1}_{B_{N,i}} \rangle_{\theta_N}$ . We have with trigonometric

formula

$$E_N = \frac{2\pi}{N} \left( \cos(x_j + \theta_N) \left( \sum_{i=1}^N \cos(x_i + \theta_N) \langle v_{\theta_N}, \mathbf{1}_{B_{N,i}} \rangle_{\theta_N} \right) + \sin(x_j + \theta_N) \left( \sum_{i=1}^N \sin(x_i + \theta_N) \langle v_{\theta_N}, \mathbf{1}_{B_{N,i}} \rangle_{\theta_N} \right) \right).$$

As  $\sum_{i=1}^N \cos(x_i + \theta_N) \langle v_{\theta_N}, \mathbf{1}_{B_{N,i}} \rangle_{\theta_N} \xrightarrow{N \rightarrow \infty} \int_S A \cos \sin f'(A \cos) = 0$  (by symmetry)

and  $\sum_{i=1}^N \sin(x_i + \theta_N) \langle v_{\theta_N}, \mathbf{1}_{B_{N,i}} \rangle_{\theta_N} \xrightarrow{N \rightarrow \infty} - \int_S A \sin^2 f'(A \cos) = -A$  with (5.3.2),

we have that

$$\sum_{i=1}^N \frac{w_{ij}^{(N)}}{N} \langle v_{\theta_N}, \mathbf{1}_{B_{N,i}} \rangle_{\theta_N} \sim_{N \rightarrow \infty} -\frac{2\pi}{N} A \sin(x_j + \theta_N).$$

Hence we have  $\left( \sum_{i=1}^N \frac{w_{ij}^{(N)}}{N} \langle v_{\theta_N}, \mathbf{1}_{B_{N,i}} \rangle_{\theta_N} \right)^2 = \frac{A^2}{N^2} \Phi(x_j, \theta_N)$  with  $\Phi(x_j, \theta_N) \sim_{N \rightarrow \infty} (2\pi \sin(x_j + \theta_N))^2$  (bounded independently of  $N, \theta_N$ ). Coming back to (5.5.21), we have

$$[Q]_t = \frac{1}{N^2} \sum_{j=1}^N \int_{t_0}^t \int_0^\infty \Phi(x_j, \theta_N) \mathbf{1}_{z \leq \lambda_{N,j}} \pi_j(ds, dz) + o\left(\frac{1}{N}\right).$$

Let

$$Q_1(t) := \frac{1}{N^2} \sum_{j=1}^N \int_{t_0}^t \int_0^\infty \Phi(x_j, \theta_N) \left( \mathbf{1}_{z \leq f(U_{N,j}(s-))} - \mathbf{1}_{z \leq f(u_{\theta_N}(x_j))} \right) \pi_j(ds, dz)$$

$$Q_2(t) := \frac{1}{N^2} \sum_{j=1}^N \int_{t_0}^t \int_0^\infty \Phi(x_j, \theta_N) \mathbf{1}_{z \leq f(u_{\theta_N}(x_j))} \tilde{\pi}_j(ds, dz)$$

$$Q_3(t) := \frac{1}{N^2} \sum_{j=1}^N \int_{t_0}^t \int_0^\infty \Phi(x_j, \theta_N) \mathbf{1}_{z \leq f(u_{\theta_N}(x_j))} ds dz,$$

so that  $[Q]_t = Q_1(t) + Q_2(t) + Q_3(t) + o\left(\frac{1}{N}\right)$ . We have (recall that  $\Phi$  is bounded)

$$\begin{aligned} \mathbf{E} [|Q_1(t)|] &\leq \frac{1}{N^2} \sum_{j=1}^N \mathbf{E} \left[ \int_{t_0}^t \int_0^\infty \Phi(x_j, \theta_N) \left| \mathbf{1}_{z \leq f(U_{N,j}(s-))} - \mathbf{1}_{z \leq f(u_{\theta_N}(x_j))} \right| \pi_j(ds, dz) \right] \\ &= \frac{\|\Phi\|_\infty}{N^2} \sum_{j=1}^N \int_{t_0}^t \mathbf{E} [|f(U_{N,j}(s-)) - f(u_{\theta_N}(x_j))|] ds \\ &\leq \frac{\|\Phi\|_\infty \|f\|_L}{N} (t - t_0) C N^{\eta-1/2} \leq C \frac{T_f(N)}{N} N^{\eta-1/2} \xrightarrow{N \rightarrow \infty} 0, \end{aligned}$$

using (5.5.1). About  $Q_2$ , we use once again that  $Q_2$  is a real martingale with quadratic variation

$$\begin{aligned} [Q_2]_t &= \sum_{j=1}^N \int_{t_0}^t \int_0^\infty \left( \frac{1}{N^2} \Phi(x_j, \theta_N) \mathbf{1}_{z \leq f(u_{\theta_N}(x_j))} \right)^2 \pi_j(ds, dz) \\ &\leq \frac{C}{N^4} \sum_{j=1}^N \int_{t_0}^t \int_0^\infty \mathbf{1}_{z \leq f(u_{\theta_N}(x_j))} \pi_j(ds, dz), \end{aligned}$$

hence with Burkholder-Davis-Gundy inequality,

$$\begin{aligned} \mathbf{E} [Q_2(t)^2] &\leq C \mathbf{E} [[Q_2]_t] \leq \frac{C}{N^4} \mathbf{E} \left[ \sum_{j=1}^N \int_{t_0}^t \int_0^\infty \mathbf{1}_{z \leq f(u_{\theta_N}(x_j))} \pi_j(ds, dz) \right] \\ &\leq \frac{C}{N^2} \frac{T_f(N)}{N} \xrightarrow{N \rightarrow \infty} 0. \end{aligned}$$

The last term  $Q_3(t)$  gives the term  $\tilde{\Theta}_N(t_0, t)$  that appears in (5.5.11).

#### 5.5.3.4 Proof of Lemma 5.5.6

Recall that  $I_4(t)$  is defined in (5.5.8):

$$I_4(t) = \sum_{j=1}^N \int_{t_0}^t \int_0^\infty \int_0^1 (1-r) D^2 \theta(U_N(s-) + r \chi_j(s, z)) [\chi_j(s, z)]^2 dr \pi_j(ds, dz).$$

A Taylor's expansion gives that  $I_4(t) = L_1(t) + L_2(t) + L_3(t) + L_4(t)$  with

$$L_1(t) = \sum_{j=1}^N \int_{t_0}^t \int_0^\infty \int_0^1 (1-r) D^2 \theta(U_N(s-) + r \chi_j(s, z)) [\chi_j(s, z)]^2 dr \tilde{\pi}_j(ds, dz),$$

$$\begin{aligned} L_2(t) &= \sum_{j=1}^N \int_{t_0}^t \int_0^\infty \int_0^1 (1-r) (D^2 \theta(U_N(s-) + r \chi_j(s, z)) - D^2 \theta(U_N(s-))) \\ &\quad \cdot [\chi_j(s, z)]^2 dr ds dz, \end{aligned}$$

$$L_3(t) = \sum_{j=1}^N \int_{t_0}^t \int_0^\infty \int_0^1 (1-r) (D^2 \theta(U_N(s-)) - D^2 \theta(u_{\theta_N})) [\chi_j(s, z)]^2 dr ds dz,$$

$$L_4(t) = \sum_{j=1}^N \int_{t_0}^t \int_0^\infty \int_0^1 (1-r) D^2 \theta(u_{\theta_N}) [\chi_j(s, z)]^2 dr ds dz.$$

$L_1$  is a real martingale and

$$\begin{aligned} [L_1](t) &= \sum_{j=1}^N \int_{t_0}^t \int_0^\infty \left( \int_0^1 (1-r) D^2 \theta(U_N(s-) + r \chi_j(s, z)) [\chi_j(s, z)]^2 dr \right)^2 \pi_j(ds, dz) \\ &\leq \frac{\|D^2 \theta\|_\infty^2}{2} \sum_{j=1}^N \int_{t_0}^t \int_0^\infty \|\chi_j(s, z)\|_2^4 \pi_j(ds, dz) \leq \frac{C}{N^4} \sum_{j=1}^N \int_{t_0}^t \int_0^\infty \mathbf{1}_{z \leq \lambda_{N,j}(s)} \pi_j(ds, dz). \end{aligned}$$

As done for  $Q_2$  in the proof of Lemma 5.5.5, we obtain that

$$\mathbf{E} [|L_1(t)^2|] \leq \frac{C}{N^2} \frac{T_f(N)}{N} \xrightarrow{N \rightarrow \infty} 0. \quad (5.5.22)$$

We have, using (5.B.6) and the fact that  $f$  is bounded

$$\begin{aligned} L_2(t) &= \sum_{j=1}^N \int_{t_0}^t \int_0^\infty \int_0^1 (1-r) \\ &\quad (D^2\theta(U_N(s-)) + r\chi_j(s, z) - D^2\theta(U_N(s-))) [\chi_j(s, z)]^2 dr ds dz \\ &\leq \sum_{j=1}^N \int_{t_0}^t \int_0^\infty \|D^2\theta\|_L \|\chi_j(s, z)\|_2 \|\chi_j(s, z)\|_2^2 ds dz \\ &\leq C \sum_{j=1}^N \int_{t_0}^t \int_0^\infty \left( \frac{1}{N} \mathbf{1}_{z \leq \lambda_{N,j}(s)} \right)^3 ds dz \leq \frac{C}{N^3} \sum_{j=1}^N \int_{t_0}^t \lambda_{N,j}(s) ds \leq \frac{CT_f(N)}{N^2}. \end{aligned}$$

Similarly, using (5.5.1)

$$\begin{aligned} L_3(t) &= \sum_{j=1}^N \int_{t_0}^t \int_0^\infty \int_0^1 (1-r) (D^2\theta(U_N(s-)) - D^2\theta(u_{\theta_N})) [\chi_j(s, z)]^2 dr ds dz \\ &\leq \frac{1}{2} \|D^2\theta\|_L \sum_{j=1}^N \int_{t_0}^t \int_0^\infty \|U_N(s-) - u_{\theta_N}\|_2 \|\chi_j(s, z)\|_2^2 ds dz \\ &\leq CN^{\eta-1/2} \sum_{j=1}^N \int_{t_0}^t \int_0^\infty \frac{1}{N^2} \mathbf{1}_{z \leq \lambda_{N,j}(s)} ds dz \leq C \frac{T_f(N)}{N} N^{\eta-1/2}. \end{aligned}$$

For  $L_4$ , we use the computation of  $D^2\theta(u_{\theta_N}) [\chi_j(s, z)]^2$  given by Lemma 5.B.3: for some  $C = C_{A,\gamma}$ ,

$$\begin{aligned} L_4(t) &= \frac{1}{2} \sum_{j=1}^N \int_{t_0}^t \int_0^\infty D^2\theta(u_{\theta_N}) [\chi_j(s, z)]^2 ds dz \\ &= \sum_{j=1}^N \int_{t_0}^t \int_0^\infty \mathbf{1}_{z \leq \lambda_{N,j}(s)} \left( \frac{C}{N^2} \cos(x_j + \theta) \sin(x_j + \theta) + O(N^{-3}) \right) ds dz \\ &= \frac{C}{N^2} \sum_{j=1}^N \int_{t_0}^t \lambda_{N,j}(s) \cos(x_j + \theta_N) \sin(x_j + \theta_N) ds + O\left(\frac{T_f(N)}{N^2}\right) \\ &= \frac{C}{N^2} \sum_{j=1}^N \int_{t_0}^t (f(U_N(s-))(x_j) - f(u_{\theta_N}(x_j))) \cos(x_j + \theta_N) \sin(x_j + \theta_N) ds \\ &\quad + O\left(\frac{T_f(N)}{N^2}\right) + \frac{C}{N^2} \sum_{j=1}^N \int_{t_0}^t f(u_{\theta_N}) \cos(x_j + \theta_N) \sin(x_j + \theta_N) ds. \end{aligned}$$

As done before for  $D_N$  in (5.5.16),  $\frac{2\pi}{N} \sum_{j=1}^N (f(U_N(s-))(x_j) - f(u_{\theta_N}(x_j))) \cos(x_j + \theta_N) \sin(x_j + \theta_N) = O(N^{\eta-1/2})$  and

$$\begin{aligned} & \frac{C}{N^2} \sum_{j=1}^N f(u_{\theta_N}(x_j)) \cos(x_j + \theta_N) \sin(x_j + \theta_N) \\ &= \frac{C}{N} \left( \int_S f(u_{\theta_N}(x)) \cos(x + \theta_N) \sin(x + \theta_N) dx + O(N^{-1}) \right) = O\left(\frac{1}{N^2}\right), \end{aligned}$$

hence as  $T_f(N) \propto N$ ,  $L_4(t) = O(N^{\eta-1/2})$ . Combining our results on  $L_2, L_3, L_4$ , we have then shown that  $\sup_{t \in [T_0(N), T_f(N)]} (L_2(t) + L_3(t) + L_4(t)) = O(N^{\eta-1/2}) \xrightarrow{N \rightarrow \infty} 0$ . We conclude with (5.5.22).

## 5.A On the stationary solutions to the Neural Field Equation

### 5.A.1 When $f$ is the Heaviside function

Here we study the NFE equation (5.1.11) and its stationary solutions (5.2.1) when  $f = H_\varrho$ . We recall the results from Kilpatrick & Ermentrout (2013) and Veltz & Faugeras (2010).

**Proposition 5.A.1.** *There exist non-zero stationary solutions to (5.1.11) when  $f = H_\varrho$ ,  $\nu(dy) = \frac{\mathbf{1}_{[-\pi, \pi]}}{2\pi} dy$  and  $w(x, y) = 2\pi \cos(x - y)$  if and only if  $\varrho \in [-1, 1]$ , and in this case, the set of stationary solutions is  $\mathcal{U}_0 \cup \mathcal{U}_{A_+(0)} \cup \mathcal{U}_{A_-(0)}$ , where  $A_+(0)$  and  $A_-(0)$  are defined in (5.2.3).*

*Proof.* (following Kilpatrick & Ermentrout (2013)) First,  $u = 0$  is an evident solution to (5.2.1). We focus now on the other solutions. To solve (5.2.1), we need to find  $A$  solving (5.2.2). As  $A \cos(x) = -A \cos(x + \pi)$ ,  $\mathcal{U}_A = \mathcal{U}_{-A}$  and we can focus on the case  $A > 0$ .

Let  $A > 0$  be a solution to (5.2.2) with  $f = H_\varrho$ . Note that we necessarily need  $A \geq |\varrho|$ , because if  $A < \varrho$ , the threshold  $\varrho$  is never reached in (5.2.2) hence the unique solution is  $A = 0$  which is a contradiction (and similarly for  $\varrho < -A$ ). Then as  $|\varrho| \leq A$ ,  $\text{Arccos}(\varrho/A) \in [0, \pi]$  is well defined and verifies  $A \cos(y) \geq \varrho \Leftrightarrow |y| \leq \text{Arccos}(\varrho/A)$ , hence (5.2.2) becomes

$$A = 2 \int_0^{\text{Arccos}(\varrho/A)} \cos(y) dy = 2 \sin(\text{Arccos}(\varrho/A)) = 2 \sqrt{1 - \left(\frac{\varrho}{A}\right)^2}. \quad (5.A.1)$$

Equation (5.A.1) has two non-negative solutions  $A_+(0)$  and  $A_-(0)$  defined in (5.2.3) if and only if  $\varrho \in [-1, 1]$ , which indeed verify  $\varrho \in [-A, A]$ , hence the result.  $\square$

### 5.A.2 When $f$ is a sigmoid

Here we prove Proposition 5.2.2, following the previous result when  $f = H_\varrho$  and using the fact that  $f_{\kappa,\varrho} \xrightarrow{\kappa \rightarrow 0} H_\varrho$ .

*Proof of Proposition 5.2.2.* Define the function  $g : \mathbb{R} \times (|\varrho|, +\infty) \rightarrow \mathbb{R}$  such that

$$\begin{cases} g(\kappa, a) := a - \int_{-\pi}^{\pi} \cos(y) f_{\kappa,\varrho}(a \cos(y)) dy, & (\kappa, a) \in \mathbb{R}_+^* \times (|\varrho|, +\infty), \\ g(\kappa, a) := a - \int_{-\pi}^{\pi} \cos(y) H_\varrho(a \cos(y)) dy \\ \qquad = a - 2\sqrt{1 - \left(\frac{\varrho}{a}\right)^2}, & (\kappa, a) \in \mathbb{R}_- \times (|\varrho|, +\infty). \end{cases} \quad (5.A.2)$$

As  $f_{\kappa,\varrho} \xrightarrow{\kappa \rightarrow 0} H_\varrho$ , by dominated convergence,  $g$  is continuous on  $\mathbb{R} \times (\varrho, +\infty)$ . It is differentiable on  $\mathbb{R}_+^* \times (\varrho, +\infty)$  and on  $\mathbb{R}_-^* \times (\varrho, +\infty)$ , we now focus on its differentiability in  $(0, a)$  for any  $a \in (\varrho, +\infty)$ . We first show the continuity of  $\frac{dg}{da}$ , that is showing

$$\lim_{\kappa \rightarrow 0} \frac{dg}{da}(\kappa, a) = \frac{dg}{da}(0, a) = 1 - \frac{2\varrho^2}{a^3 \sqrt{1 - \left(\frac{\varrho}{a}\right)^2}}. \quad (5.A.3)$$

For any  $\kappa > 0$ , recalling the definition of  $f_{\kappa,\varrho}$  in (5.1.3),

$$\frac{dg}{da}(\kappa, a) = 1 - \int_{-\pi}^{\pi} \frac{\cos(y)^2 e^{-(a \cos(y) - \varrho)/\kappa}}{\kappa (1 + e^{-(a \cos(y) - \varrho)/\kappa})^2} dy = 1 - 2 \int_0^\pi \frac{\cos(y)^2 e^{-(a \cos(y) - \varrho)/\kappa}}{\kappa (1 + e^{-(a \cos(y) - \varrho)/\kappa})^2} dy,$$

and by the change of variables  $a \cos(y) - \varrho = u$ , we get

$$\begin{aligned} \int_0^\pi \frac{\cos(y)^2 e^{-(a \cos(y) - \varrho)/\kappa}}{\kappa (1 + e^{-(a \cos(y) - \varrho)/\kappa})^2} dy &= \int_{-a-\varrho}^{a-\varrho} \frac{(u + \varrho)^2}{a^3 \sqrt{1 - \left(\frac{u+\varrho}{a}\right)^2}} \frac{e^{-u/\kappa}}{\kappa (1 + e^{-u/\kappa})^2} du \\ &= \frac{1}{a^3} \int_{\mathbb{R}} h(-u) \varphi_\kappa(u) du = \frac{1}{a^3} (h * \varphi_\kappa)(0) \end{aligned}$$

with  $h(u) := \mathbf{1}_{(\varrho-a, a+\varrho)}(u) \frac{(-u + \varrho)^2}{\sqrt{1 - \left(\frac{-u+\varrho}{a}\right)^2}}$  and  $\varphi_\kappa(u) := \frac{e^{-u/\kappa}}{\kappa (1 + e^{-u/\kappa})^2}$ . By Lemma

5.B.1,  $(h * \varphi_\kappa)(0) \xrightarrow{\kappa \rightarrow 0} h(0) = \frac{\varrho^2}{\sqrt{1 - \left(\frac{\varrho}{a}\right)^2}}$  and (5.A.3) follows. We show now the

continuity of  $\frac{dg}{d\kappa}$ , that is

$$\lim_{\kappa \rightarrow 0} \frac{dg}{d\kappa}(\kappa, a) = 0. \quad (5.A.4)$$

For any  $\kappa > 0$ , we obtain similarly

$$\begin{aligned} \frac{dg}{d\kappa}(\kappa, a) &= 2 \int_0^\pi \frac{\cos(y) (a \cos(y) - \varrho) e^{-(a \cos(y) - \varrho)/\kappa}}{\kappa^2 (1 + e^{-(a \cos(y) - \varrho)/\kappa})^2} dy \\ &= \frac{2}{a^2 \kappa} \int_{(-a-\varrho)/\kappa}^{(a-\varrho)/\kappa} \tilde{h}(\kappa v) \frac{e^{-v}}{(1 + e^{-v})^2} dv \end{aligned}$$

with  $\tilde{h}(u) := \frac{u(u + \varrho)}{\sqrt{1 - (\frac{u+\varrho}{a})^2}}$ . Let  $F(\kappa) := \int_0^{(a-\varrho)/\kappa} \tilde{h}(\kappa v) \frac{e^{-v}}{(1 + e^{-v})^2} dv$ , by dominated convergence  $F(\kappa) \xrightarrow{\kappa \rightarrow 0} \tilde{h}(0) \int_0^\infty \frac{e^{-v}}{(1 + e^{-v})^2} dv = \frac{\tilde{h}(0)}{2} = 0$ . Setting  $F(0) := 0$ ,  $F$  is continuous on  $[0, \infty)$  and differentiable on  $(0, \infty)$  with

$$F'(\kappa) = -\frac{(a - \varrho)}{\kappa^2} \tilde{h}(a - \varrho) \frac{e^{-(a-\varrho)/\kappa}}{(1 + e^{-(a-\varrho)/\kappa})^2} + \int_0^{(a-\varrho)/\kappa} v \tilde{h}'(\kappa v) \frac{e^{-v}}{(1 + e^{-v})^2} dv.$$

By dominated convergence,  $F'(\kappa) \xrightarrow{\kappa \rightarrow 0} 0 + \tilde{h}'(0) \int_0^\infty \frac{v e^{-v}}{(1 + e^{-v})^2} dv = \tilde{h}'(0) \ln(2) = \frac{\varrho \ln(2)}{\sqrt{1 + (\frac{\varrho}{a})^2}}$ . Hence by Taylor's theorem,  $F(\kappa) = \kappa \frac{\varrho \ln(2)}{\sqrt{1 + (\frac{\varrho}{a})^2}} + o(\kappa)$  as  $\kappa \rightarrow 0$ . Similarly, let

$$G(\kappa) := \int_{(-a-\varrho)/\kappa}^0 \tilde{h}(\kappa v) \frac{e^{-v}}{(1 + e^{-v})^2} dv = \int_0^{(a+\varrho)/\kappa} \tilde{h}(-v\kappa) \frac{e^v}{(1 + e^v)^2} dv,$$

we also have  $G(\kappa) \rightarrow 0$ . Setting  $G(0) := 0$ ,  $G$  is differentiable on  $(0, \infty)$  with

$$G'(\kappa) = -\frac{a + \varrho}{\kappa^2} \tilde{h}(-a - \varrho) \frac{e^{(a+\varrho)/\kappa}}{(1 + e^{(a+\varrho)/\kappa})^2} - \int_0^{(a+\varrho)/\kappa} v \tilde{h}'(\kappa v) \frac{e^v}{(1 + e^v)^2} dv,$$

thus  $G'(\kappa) \xrightarrow{\kappa \rightarrow 0} -\tilde{h}'(0) \ln(2)$ . Hence by Taylor's theorem  $G(\kappa) = -\kappa \frac{\varrho \ln(2)}{\sqrt{1 + (\frac{\varrho}{a})^2}} + o(\kappa)$  as  $\kappa \rightarrow 0$ . We obtain then

$$\frac{dg}{d\kappa}(\kappa, a) = \frac{2}{a^2 \kappa} (F(\kappa) + G(\kappa)) = \frac{2}{a^2 \kappa} \left( \kappa \tilde{h}'(0) \ln(2) - \kappa \tilde{h}'(0) \ln(2) + o(\kappa) \right) = o(1),$$

hence (5.A.4) is true. We have shown that  $g$  is indeed  $\mathcal{C}^1$  on  $\mathbb{R} \times (|\varrho|, +\infty)$ .

Our aim is to apply the implicit function theorem. With Proposition 5.A.1, we have that  $g(0, A_+(0)) = 0$ . Let us show that  $\frac{dg}{da}(0, A_+(0)) \neq 0$ . Using (5.2.3), we obtain

$$\frac{dg}{da}(0, A_+(0)) = 1 - \frac{2\varrho^2}{2 \left(1 + \sqrt{1 - \varrho^2}\right) \sqrt{2 + 2\sqrt{1 - \varrho^2} - \varrho^2}},$$

we then need  $\varrho^2 \neq \left(1 + \sqrt{1 - \varrho^2}\right) \sqrt{2 + 2\sqrt{1 - \varrho^2} - \varrho^2}$ , which is true if and only if  $\varrho \neq 1$ . We conclude by implicit function theorem.



It remains now to prove that there exists  $\kappa_1 > 0$  such that for any  $\kappa \in (0, \kappa_1)$ ,  $I(1, \kappa) = \int_S f'_{\kappa, \varrho}(A(\kappa) \cos(x)) dx \in (1, 2)$ . We have

$$\begin{aligned} \mathcal{I}(1, \kappa) &= 2 \int_0^\pi \frac{e^{-(A(\kappa) \cos(y) - \varrho)/\kappa}}{\kappa (1 + e^{-(A(\kappa) \cos(y) - \varrho)/\kappa})^2} dy \\ &= 2 \int_{-A(\kappa) - \varrho}^{A(\kappa) - \varrho} \frac{1}{A(\kappa) \sqrt{1 - \left(\frac{u + \varrho}{A(\kappa)}\right)^2}} \frac{e^{-u/\kappa}}{\kappa (1 + e^{-u/\kappa})^2} \\ &= h * \phi_\kappa(0) \xrightarrow{\kappa \rightarrow 0} \frac{2}{A_+(0) \sqrt{1 - \frac{\varrho^2}{A_+(0)^2}}}, \end{aligned}$$

with  $h(u) = \mathbf{1}_{(\varrho - A(\kappa), A(\kappa) + \varrho)}(u) \frac{2}{A(\kappa) \sqrt{1 - \left(\frac{-u + \varrho}{A(\kappa)}\right)^2}}$  and using Lemma 5.B.1 and as  $A(\kappa) \xrightarrow{\kappa \rightarrow 0} A_+(0)$  defined in (5.2.3). As

$$\frac{1}{A_+(0) \sqrt{1 - \frac{\varrho^2}{A_+(0)^2}}} = \frac{1}{\sqrt{2 + 2\sqrt{1 - \varrho^2} - \varrho^2}} < 1$$

when  $\varrho \in (-1, 1)$ , by continuity of  $\kappa \mapsto A(\kappa)$  there exists  $\kappa_1 > 0$  such that for  $\kappa < \kappa_1$ , we have indeed  $I(1, \kappa) < 2$ . Let us show now that for small  $\kappa$  we have also  $I(1, \kappa) > 1$ . We have

$$I(1, 0) - 1 = \frac{2}{\sqrt{2 + 2\sqrt{1 - \varrho^2} - \varrho^2}} - 1 = \frac{2 - \sqrt{2 + 2\sqrt{1 - \varrho^2} - \varrho^2}}{\sqrt{2 + 2\sqrt{1 - \varrho^2} - \varrho^2}},$$

and as  $2\sqrt{1 - \varrho^2} - \varrho^2 < 2$  we have indeed  $I(1, 0) - 1 > 0$ . Similarly by continuity it implies that  $I(1, \kappa) > 1$  for  $\kappa$  small enough.  $\square$

## 5.B Some computations

### 5.B.1 Control of the noise perturbation

We prove here Proposition 5.4.3, which is a part of the **Step 2** of the proof of Theorem 5.2.9 in Section 5.4. The proof relies on an adaptation of an argument given in (Zhu *et al.*, 2017, Theorem 4.3), where a similar quantity to the following (5.B.1) is considered for  $N = 1$ , and used in the proof of Proposition 4.2 of Agathe-Nerine (2023a).

*Proof of Proposition 5.4.3.* Recall the expression of  $(Z_{N,j})_{1 \leq j \leq N}$  in (5.1.14) and the compensated measure  $\tilde{\pi}_j(ds, dz) := \pi_j(ds, dz) - \lambda_{N,j} ds dz$ , so that with the linearity of  $(e^{t\mathcal{L}_{\phi_{n-1}}})_{t \geq 0}$ , we obtain that  $\zeta_n$  can be written as

$$\zeta_n(t) = \sum_{j=1}^N \int_0^t \int_0^\infty e^{(t-s)\mathcal{L}_{\phi_{n-1}}} \chi_j(s, z) \tilde{\pi}_j(ds, dz), \quad (5.B.1)$$

with

$$\chi_j(s, z) := \left( \sum_{i=1}^N \mathbf{1}_{B_{N,i}} \frac{w_{ij}^{(N)}}{N} \mathbf{1}_{z \leq \lambda_{N,j}(s)} \right). \quad (5.B.2)$$

Fix  $m \geq 1$ . The functional  $\phi : L^2(I) \rightarrow \mathbb{R}$  given by  $\phi(v) = \|v\|_2^{2m}$  is of class  $\mathcal{C}^2$  (recall that  $\zeta_n(t) \in L^2(I)$ ) so that by Itô formula on the expression (5.B.1) we obtain

$$\begin{aligned} \phi(\zeta_n(t)) &= \int_0^t \phi'(\zeta_n(s)) \mathcal{L}_{\phi_{n-1}}(\zeta_n(s)) ds \\ &\quad + \sum_{j=1}^N \int_0^t \int_0^\infty \phi'(\zeta_n(s-)) \chi_j(s, z) \tilde{\pi}_j(ds, dz) \\ &\quad + \sum_{j=1}^N \int_0^t \int_0^\infty [\phi(\zeta_n(s-) + \chi_j(s, z)) - \phi(\zeta_n(s-)) - \phi'(\zeta_n(s-)) \chi_j(s, z)] \pi_j(ds, dz) \\ &:= I_0(t) + I_1(t) + I_2(t). \end{aligned} \quad (5.B.3)$$

We also have that for any  $v, h, k \in L^2(I)$ ,  $\phi'(v)h = 2m\|v\|_2^{2m-2} \langle v, h \rangle \in \mathbb{R}$  and  $\phi''(v)(h, k) = 2m(2m-1)\|v\|_2^{2m-4} \langle v, k \rangle \langle v, h \rangle + 2m\|v\|_2^{2m-2} \langle h, k \rangle$ .

We have  $I_0(t) = \int_0^t 2m\|\zeta_N(s)\|_2^{2m-2} \langle \zeta_N(s), \mathcal{L}(\zeta_N(s)) \rangle ds$ . From Proposition 5.2.5, the operator  $\mathcal{L}_{\phi_{n-1}}$  has only three non-positive eigenvalues hence by Lumer-Philipp's Theorem (see Section 1.4 of Pazy (1974)),  $\langle \zeta_n(s), \mathcal{L}_{\phi_{n-1}}(\zeta_n(s)) \rangle \leq 0$ . Then for any  $t \geq 0$ ,  $I_0(t) \leq 0$ .

Let some  $\varepsilon > 0$  to be chosen later. About  $I_1$ , using Burkholder-Davis-Gundy inequality, the independence of the family  $(\pi_j)$  and Hölder inequality with some well chosen parameter, one can show

$$\begin{aligned} \mathbf{E} \left[ \sup_{s \leq T} |I_1(s)| \right] &\leq C(2m-1)\varepsilon \mathbf{E} \left[ \sup_{0 \leq s \leq T} (\|\zeta_n(s)\|_2^{2m}) \right] \\ &\quad + C\varepsilon^{-(2m-1)} \mathbf{E} \left[ \left( \sum_{j=1}^N \int_0^T \int_0^\infty \|\chi_j(s, z)\|_2^2 \pi_j(ds, dz) \right)^m \right], \end{aligned} \quad (5.B.4)$$

as done for Proposition 4.2 of Agathe-Nerine (2023a), with  $C$  some deterministic constant. About  $I_2$ , using Taylor's Lagrange formula and both Hölder and Young's inequalities, one can show

$$\begin{aligned} \mathbf{E} \left[ \sup_{s \leq T} |I_2(s)| \right] &\leq m(2m-2)\varepsilon \mathbf{E} \left[ \sup_{0 \leq s \leq t} (\|\zeta_n(s)\|_2^{2m}) \right] \\ &\quad + 2m\varepsilon^{-(2m-2)} \mathbf{E} \left[ \left( \sum_{j=1}^N \int_0^t \int_0^\infty \|\chi_j(s, z)\|_2^2 \pi_j(ds, dz) \right)^m \right]. \end{aligned} \quad (5.B.5)$$

Fixing  $\varepsilon$  such that  $\varepsilon(C(2m-1) + m(2m-2)) \leq \frac{1}{2}$  and taking the expectation in (5.B.3), we get

$$\mathbf{E} \left[ \sup_{s \leq T} \|\zeta_n(s)\|_2^{2m} \right] \leq 2C \mathbf{E} \left[ \left( \sum_{j=1}^N \int_0^T \int_0^\infty \|\chi_j(s, z)\|_2^2 \pi_j(ds, dz) \right)^m \right],$$

where  $C > 0$  depends only on  $m$ . As  $\sup_{i,j} w_{ij}^{(N)} \leq 2\pi$ ,

$$\|\chi_j(s, z)\|_2^2 = \frac{1}{N} \sum_{i=1}^N \left( \frac{w_{ij}^{(N)}}{N} \right)^2 \mathbf{1}_{z \leq \lambda_{N,j}(s)} \leq \frac{4\pi^2}{N^2} \mathbf{1}_{z \leq \lambda_{N,j}(s)}. \quad (5.B.6)$$

As  $f$  is bounded by 1, we have that  $\mathbf{E} \left[ \sup_{s \leq T} \|\zeta_n(s)\|_2^{2m} \right] \leq \frac{C}{N^m} \mathbf{E} \left[ \frac{1}{N} \sum_{j=1}^N \tilde{Z}_j(T)^m \right]$ ,

where  $(\tilde{Z}_j(t))$  are i.i.d copies of a Poisson process on intensity 1. Hence for some constant  $C = C(T, m, \kappa, \rho) > 0$ , for any  $1 \leq n \leq n_f$ ,  $\mathbf{E} \left[ \sup_{0 \leq t \leq T} \|\zeta_n(t)\|_2^{2m} \right] \leq \frac{C}{N^m}$ .

It implies

$$\mathbf{P} \left( \sup_{t \in [0, T]} \|\zeta_n(t)\|_2 \geq \frac{N^\eta}{\sqrt{N}} \right) \leq \frac{\mathbf{E} \left[ \sup_{0 \leq t \leq T} \|\zeta_n(t)\|_2^{2m} \right]}{N^{2\eta m}} N^m \leq CN^{-2m\eta},$$

hence by a union bound  $\mathbf{P}(A_N^C) \leq Cn_f N^{-2m\eta} = CN^{\alpha-2m\eta}$ . We can then choose  $m$  large enough to obtain the result of Proposition 5.4.3.  $\square$

## 5.B.2 Analysis complements

**Lemma 5.B.1.** Define  $\varphi(u) = \frac{e^{-u}}{(1 + e^{-u})^2}$ . For any  $\kappa > 0$ , let  $\varphi_\kappa(u) := \frac{1}{\kappa} \varphi\left(\frac{u}{\kappa}\right)$ . Then  $(\varphi_\kappa)_{\kappa > 0}$  is an approximate identity and  $\varphi_\kappa * h \xrightarrow{\kappa \rightarrow 0} h$  for any  $h \in L^p$ , with  $1 \leq p < \infty$ .

*Proof.* It suffices to check that

$$\int_{\mathbb{R}} \varphi(u) du = \int_{\mathbb{R}} \frac{e^{-u}}{(1 + e^{-u})^2} du = \left[ \frac{1}{1 + e^{-u}} \right]_{-\infty}^{+\infty} = 1.$$

$\square$

**Lemma 5.B.2.** Let  $N \geq 1$ , recall that  $S = [-\pi, \pi)$  and its regular subdivision  $x_i = \frac{2i\pi}{N} - \pi$  for  $0 \leq i \leq N$ . For any function  $g \in \mathcal{C}^2(I, \mathbb{R})$ , we have

$$\frac{2\pi}{N} \sum_{j=1}^N g(x_j) = \int_S g(y) dy - \frac{1}{2} \left( \frac{2\pi}{N} \right)^2 \sum_{j=1}^N g'(x_j) + o\left(\frac{1}{N}\right). \quad (5.B.7)$$

Moreover, for any function  $h \in \mathcal{C}^1(I, \mathbb{R})$ , we have

$$\sum_{i=1}^N h(x_i) \int_{x_{i-1}}^{x_i} g(y) dy = \int_S h(x) g(x) dx - \sum_{i=1}^N h'(x_i) \int_{x_{i-1}}^{x_i} (y - x_i) g(y) dy + o\left(\frac{1}{N}\right). \quad (5.B.8)$$

*Proof.* Let  $C_j = (x_{j-1}, x_j)$  for  $1 \leq j \leq N$ . From Taylor's expansion,  $g(y) = g(x_j) + g'(x_j)(y - x_j) + \int_{x_j}^y g''(t)(y - t)dt$  hence the result (5.B.7) as  $\int_S g(y)dy = \sum_{j=1}^N \int_{C_j} g(y)dy$ . About (5.B.8), we proceed similarly as

$$\int_S hg = \sum_j \int_{C_j} g(y) \left( h(x_j) + h'(x_j)(y - x_j) + \int_{x_j}^y h''(t)(y - t)dt \right) dy.$$

□

### 5.B.3 Auxilliary lemmas

#### 5.B.3.1 About the derivatives of the isochron

**Lemma 5.B.3.** *Let  $\phi \in S$ . There exists  $C = C_{A,\gamma}$  such that*

$$D^2\theta(u_\phi)[\chi_j(s, z)]^2 = \mathbf{1}_{z \leq \lambda_{N,j}(s)} \left( \frac{C}{N^2} \cos(x_j + \phi) \sin(x_j + \phi) + O(N^{-3}) \right), \quad (5.B.9)$$

where the notation  $O(N^{-3})$  is uniform in  $(s, z, \phi)$ .

*Proof.* Recall (5.2.17), we have

$$\begin{aligned} D^2\theta(u_\phi)[\chi_j(s, z)]^2 &= \frac{1}{2A^2} (2\alpha_\phi^\circ(\chi_j(s, z))\beta_\phi(v_\phi, \chi_j(s, z)) + \beta_\phi(\chi_j(s, z), \chi_j(s, z))) \\ &\quad + \frac{1+\gamma}{A^2(1-\gamma)} \alpha_\phi^\gamma(\chi_j(s, z))\beta_\phi(u_\phi, \chi_j(s, z)) \\ &\quad - \frac{(2-\gamma)(1+\gamma)}{2(1-\gamma)} (\alpha_\phi^\circ(\chi_j(s, z))^2 + \alpha_\phi^\gamma(\chi_j(s, z))^2), \end{aligned} \quad (5.B.10)$$

Let us compute each term. About  $\alpha$ , using some trigonometric formula and Lemma 5.B.2 we have

$$\begin{aligned} \alpha_\phi^\circ(\chi_j(s, z)) &= \frac{\langle \chi_j, v_\phi \rangle_\phi}{A} = \frac{1}{A} \int_S \chi_j v_\phi f'(u_\phi) \\ &= \frac{2\pi}{A} \mathbf{1}_{z \leq \lambda_{N,j}(s)} \sum_{i=1}^N \frac{\cos(x_i - x_j)}{N} \int_S v_\phi f'(u_\phi) \mathbf{1}_{B_{N,i}} \\ &= \frac{2\pi}{AN} \mathbf{1}_{z \leq \lambda_{N,j}(s)} \left( \cos(x_j + \phi) \sum_{i=1}^N \cos(x_i + \phi) \int_S v_\phi f'(u_\phi) \mathbf{1}_{B_{N,i}} \right. \\ &\quad \left. + \sin(x_j + \phi) \sum_{i=1}^N \sin(x_i + \phi) \int_S v_\phi f'(u_\phi) \mathbf{1}_{B_{N,i}} \right) \\ &= \frac{2\pi}{AN} \mathbf{1}_{z \leq \lambda_{N,j}(s)} \left( \cos(x_j + \phi) \int_S \cos(x + \phi) v_\phi(x) f'(u_\phi(x)) dx \right. \\ &\quad \left. + \sin(x_j + \phi) \int_S \sin(x + \phi) v_\phi(x) f'(u_\phi(x)) dx + O(N^{-1}) \right), \end{aligned}$$

using Lemma 5.3.1, hence

$$\begin{aligned}\alpha_\phi^\circ(\chi_j(s, z)) &= \mathbf{1}_{z \leq \lambda_{N,j}(s)} \left( -\frac{2\pi}{N} \sin(x_j + \phi) \mathcal{I}(\sin^2) + O(N^{-2}) \right) \\ &= \mathbf{1}_{z \leq \lambda_{N,j}(s)} \left( -\frac{2\pi}{N} \sin(x_j + \phi) + O(N^{-2}) \right).\end{aligned}$$

We prove in a same way that

$$\alpha_\phi^\gamma(\chi_j(s, z)) = \mathbf{1}_{z \leq \lambda_{N,j}(s)} \left( \frac{2\pi}{\|u_\phi\|_\phi N} \cos(x_j + \phi) (\mathcal{I}(1) - 1) + O(N^{-2}) \right).$$

About  $\beta$ , we have similarly using Lemma 5.B.2 that

$$\begin{aligned}\beta_\phi(v_\phi, \chi_j(s, z)) &= \int_S f''(u_\phi(y)) v_\phi(y)^2 \chi_j(s, z)(y) dy \\ &= \sum_{i=1}^N \frac{w_{ij}^{(N)}}{N} \mathbf{1}_{z \leq \lambda_{N,j}(s)} \int_S f''(u_\phi(y)) v_\phi(y)^2 \mathbf{1}_{B_{N,i}}(y) dy \\ &= \mathbf{1}_{z \leq \lambda_{N,j}(s)} \frac{2\pi}{N} \left( \cos(x_j + \phi) \sum_{i=1}^N \cos(x_i + \phi) \int_S f''(u_\phi(y)) v_\phi(y)^2 \mathbf{1}_{B_{N,i}}(y) dy \right. \\ &\quad \left. + \sin(x_j + \phi) \sum_{i=1}^N \sin(x_i + \phi) \int_S f''(u_\phi(y)) v_\phi(y)^2 \mathbf{1}_{B_{N,i}}(y) dy \right) \\ &= \mathbf{1}_{z \leq \lambda_{N,j}(s)} \frac{2\pi}{N} \left( \cos(x_j + \phi) \int_S \cos(y + \phi) f''(u_\phi(y)) v_\phi(y)^2 dy \right. \\ &\quad \left. + \sin(x_j + \phi) \int_S \sin(y + \phi) f''(u_\phi(y)) v_\phi(y)^2 dy + O\left(\frac{1}{N}\right) \right) \\ &= \mathbf{1}_{z \leq \lambda_{N,j}(s)} \frac{2\pi}{N} \left( \cos(x_j + \phi) \int_S \cos(y + \phi) f''(u_\phi(y)) v_\phi(y)^2 dy + O\left(\frac{1}{N}\right) \right).\end{aligned}$$

With Lemma 5.3.1 and an integration by parts, we obtain

$$\begin{aligned}\int_S \cos(y + \phi) f''(u_\phi(y)) v_\phi(y)^2 dy &= A^2 \int_S \cos(y + \phi) f''(A \cos(y + \phi)) \sin^2(y + \phi) dy \\ &= \int_S (-A \sin(y) f''(A \cos(y)) (-A \sin(y) \cos(y)) dy \\ &= - \int_S f'(A \cos(y)) (-A + 2A \sin^2) dy \\ &= A (\mathcal{I}(1) - 2\mathcal{I}(\sin^2)) = A\gamma\end{aligned}$$

recalling (5.3.7), hence  $\beta_\phi(v_\phi, \chi_j(s, z)) = \mathbf{1}_{z \leq \lambda_{N,j}(s)} \left( \frac{2\pi}{N} A\gamma \cos(x_j + \phi) + O(N^{-2}) \right)$ .

We prove in a same way that

$$\beta_\phi(u_\phi, \chi_j(s, z)) = -\mathbf{1}_{z \leq \lambda_{N,j}(s)} \left( \frac{2\pi}{N} A\gamma \sin(x_j + \phi) + O(N^{-2}) \right).$$

Finally we have

$$\begin{aligned}
 \beta_\phi(\chi_j(s, z), \chi_j(s, z)) &= \int_S f''(u_\phi(y))v_\phi(y) \left( \sum_{i=1}^N \mathbf{1}_{B_{N,i}}(y) \frac{w_{ij}^{(N)}}{N} \mathbf{1}_{z \leq \lambda_{N,j}(s)} \right)^2 dy \\
 &= \mathbf{1}_{z \leq \lambda_{N,j}(s)} \left( \frac{2\pi}{N} \right)^2 \sum_{i=1}^N \int_{B_{N,i}} f''(u_\phi(y))v_\phi(y) dy \\
 &\quad (\cos(x_i + \phi) \cos(x_j + \phi) + \sin(x_i + \phi) \sin(x_j + \phi))^2 \\
 &= \mathbf{1}_{z \leq \lambda_{N,j}(s)} \left[ \left( \frac{2\pi}{N} \right)^2 2 \cos(x_j + \phi) \sin(x_j + \phi) \right. \\
 &\quad \left. \int_S \cos(y + \phi) \sin(y + \phi) f''(u_\phi(y))v_\phi(y) dy + O(N^{-3}) \right].
 \end{aligned}$$

With an integration by parts and recognising (5.3.7),

$$\begin{aligned}
 &\int_S \cos(y + \phi) \sin(y + \phi) f''(u_\phi(y))v_\phi(y) dy \\
 &= -A \int_S \cos(y) \sin(y) f''(A \cos(y)) \sin(y) dy \\
 &= \int_S (-A \sin(y) f''(A \cos(y))) (\cos(y) \sin(y)) dy = -\gamma,
 \end{aligned}$$

hence we obtain that

$$\beta_\phi(\chi_j(s, z), \chi_j(s, z)) = \mathbf{1}_{z \leq \lambda_{N,j}(s)} \left( -\frac{8\gamma}{\pi^2} \cos(x_j + \phi) \sin(x_j + \phi) + O(N^{-3}) \right).$$

Putting all the previous estimates together in (5.B.10), we obtain (5.B.9) for some constant  $C = C_{A,\gamma}$ .  $\square$

### 5.B.3.2 About the fluctuations

**Lemma 5.B.4** (Some computations for the proof of Proposition 5.5.2). *Let  $\phi \in S$ . Recall the definitions of  $u_\phi$  and  $v_\phi$  in (5.2.4) and (5.2.7). We have*

$$A_1 := \frac{2\pi}{N} \sum_{j=1}^N \cos(x_j + \phi) f(u_\phi(x_j)) = A + o\left(\frac{1}{N}\right) \quad (5.B.11)$$

$$A_2 := \sum_{i=1}^N \cos(x_i + \phi) \langle v_\phi, \mathbf{1}_{B_{N,i}} \rangle_\phi = \frac{A\pi}{N} + o\left(\frac{1}{N}\right) \quad (5.B.12)$$

$$A_3 := \frac{2\pi}{N} \sum_{j=1}^N \sin(x_j + \phi) f(u_\phi(x_j)) = o\left(\frac{1}{N}\right) \quad (5.B.13)$$

$$A_4 := \sum_{i=1}^N \sin(x_i + \phi) \langle v_\phi, \mathbf{1}_{B_{N,i}} \rangle_\phi = -A + o\left(\frac{1}{N}\right), \quad (5.B.14)$$

where the notation  $o\left(\frac{1}{N}\right)$  is uniform in the choice of  $\phi$ .

*Proof.* From Lemma 5.B.2, more especially (5.B.7) applied to  $g(y) = \cos(y + \phi)f(u_\phi(y))$ , we have using (5.3.2) that

$$\begin{aligned} A_1 &= \frac{2\pi^2}{N^2} \sum_{j=1}^N (\sin(x_j + \phi)f(u_\phi(x_j)) - \cos(x_j + \phi)f'(u_\phi(x_j))v_\phi(x_j)) \\ &\quad + \int_S \cos(x + \phi)f(u_\phi(x))dx + o\left(\frac{1}{N}\right) \\ &= A + \frac{2\pi^2}{N^2} \sum_{j=1}^N (\sin(x_j + \phi)f(u_\phi(x_j)) - \cos(x_j + \phi)f'(u_\phi(x_j))v_\phi(x_j)) + o\left(\frac{1}{N}\right) \\ &= A + o\left(\frac{1}{N}\right), \end{aligned}$$

as

$$\begin{aligned} &\frac{2\pi}{N} \sum_{j=1}^N (\sin(x_j + \phi)f(u_\phi(x_j)) - \cos(x_j + \phi)f'(u_\phi(x_j))v_\phi(x_j)) \\ &= \int_S \sin(y + \phi)f(A \cos(y + \phi))dy + A \int_S \cos(y + \phi)f'(A \cos(y + \phi)) \sin(y + \phi)dy + O\left(\frac{1}{N}\right) \\ &= O\left(\frac{1}{N}\right). \end{aligned}$$

Similarly we can prove (5.B.13) as

$$\begin{aligned} A_3 &= -\frac{2\pi^2}{N^2} \sum_{j=1}^N (\cos(x_j + \phi)f(u_\phi(x_j)) + \sin(x_j + \phi)f'(u_\phi(x_j))v_\phi(x_j)) \\ &\quad + \int_S \sin(x + \phi)f(u_\phi(x))dx + o\left(\frac{1}{N}\right) \\ &= -\frac{2\pi^2}{N^2} \sum_{j=1}^N (\cos(x_j + \phi)f(u_\phi(x_j)) + \sin(x_j + \phi)f'(u_\phi(x_j))v_\phi(x_j)) + o\left(\frac{1}{N}\right) \\ &= o\left(\frac{1}{N}\right), \end{aligned}$$

using that  $\int_S \sin(x + \phi)f(u_\phi(x))dx = 0$  by symmetry and

$$\begin{aligned} &\frac{2\pi}{N} \sum_{j=1}^N (\cos(x_j + \phi)f(u_\phi(x_j)) + \sin(x_j + \phi)f'(u_\phi(x_j))v_\phi(x_j)) \\ &= \int_S (\cos f(A \cos) - \sin f'(A \cos)A \sin) + O\left(\frac{1}{N}\right) = A - A + O\left(\frac{1}{N}\right) = O\left(\frac{1}{N}\right). \end{aligned}$$

From Lemma 5.B.2, more especially (5.B.8) applied to  $g(y) = v_\phi(y)f'(u_\phi(y))$

and  $h(x) = \cos(x + \phi)$ , we have that

$$\begin{aligned}
 A_2 &= \sum_{i=1}^N \cos(x_i + \phi) \langle v_\phi, \mathbf{1}_{B_{N,i}} \rangle_\phi = \sum_{i=1}^N \cos(x_i + \phi) \int_{B_{N,i}} v_\phi(y) f'(u_\phi(y)) dy \\
 &= \sum_{i=1}^N \sin(x_i + \phi) \int_{B_{N,i}} (y - x_i) v_\phi(y) f'(u_\phi(y)) dy \\
 &\quad + \int_S \cos(x + \phi) v_\phi(x) f'(u_\phi(x)) dx + o\left(\frac{1}{N}\right) \\
 &= \sum_{i=1}^N \sin(x_i + \phi) \int_{B_{N,i}} (y - x_i) v_\phi(y) f'(u_\phi(y)) dy + o\left(\frac{1}{N}\right) \\
 &= \sum_{i=1}^N \sin(x_i + \phi) v_\phi(x_i) f'(u_\phi(x_i)) \int_{B_{N,i}} (y - x_i) dy + o\left(\frac{1}{N}\right) \\
 &= - \sum_{i=1}^N \sin(x_i + \phi) v_\phi(x_i) f'(u_\phi(x_i)) \frac{1}{2} \left(\frac{2\pi}{N}\right)^2 + o\left(\frac{1}{N}\right) \\
 &= \frac{\pi}{N} \left( A \int \sin(x + \phi)^2 f'(A \cos(x + \phi)) dx + O\left(\frac{1}{N}\right) \right) + o\left(\frac{1}{N}\right) \\
 &= \frac{A\pi}{N} + o\left(\frac{1}{N}\right),
 \end{aligned}$$

and similarly, for the choice  $h(x) = \sin(x + \phi)$  and using (5.3.2)

$$\begin{aligned}
 A_4 &= \sum_{i=1}^N \sin(x_i + \phi) \langle v_\phi, \mathbf{1}_{B_{N,i}} \rangle_\phi = \sum_{i=1}^N \sin(x_i + \phi) \int_{B_{N,i}} v_\phi(y) f'(u_\phi(y)) dy \\
 &= \int_S \sin(x + \phi) v_\phi(x) f'(u_\phi(x)) dx \\
 &\quad - \sum_{i=1}^N \cos(x_i + \phi) \int_{B_{N,i}} (y - x_i) v_\phi(y) f'(u_\phi(y)) dy + o\left(\frac{1}{N}\right) \\
 &= -A - \sum_{i=1}^N \cos(x_i + \phi) \int_{B_{N,i}} (y - x_i) v_\phi(y) f'(u_\phi(y)) dy + o\left(\frac{1}{N}\right) \\
 &= -A + A \sum_{i=1}^N \cos(x_i + \phi) \int_{B_{N,i}} (y - x_i) \sin(y + \phi) f'(u_\phi(y)) dy + o\left(\frac{1}{N}\right).
 \end{aligned}$$



As

$$\begin{aligned} & \sum_{i=1}^N \cos(x_i + \phi) \int_{B_{N,i}} (y - x_i) \sin(y + \phi) f'(u_\phi(y)) dy \\ &= \sum_{i=1}^N \cos(x_i + \phi) \sin(x_i + \phi) f'(u_\phi(x_i)) \int_{B_{N,i}} (y - x_i) dy + O\left(\frac{1}{N^2}\right) \\ &= -\frac{\pi}{N} \frac{2\pi}{N} \sum_{i=1}^N \cos(x_i + \phi) \sin(x_i + \phi) f'(u_\phi(x_i)) + o\left(\frac{1}{N}\right) \\ &= -\frac{\pi}{N} \int_S \cos(x + \phi) \sin(x + \phi) f'(u_\phi(x)) dx + O\left(\frac{1}{N^2}\right) + o\left(\frac{1}{N}\right) = o\left(\frac{1}{N}\right), \end{aligned}$$

we obtain (5.B.14). □

# Bibliographie

- Z. P. ADAMS : Existence, Regularity, and a Strong Itô Formula for the Isochronal Phase of SPDE, 2023. URL <https://arxiv.org/abs/2109.04515>. arXiv :2109.04515.
- Z. P. ADAMS & J. MACLAURIN : The Isochronal Phase of Stochastic PDE and Integral Equations : Metastability and Other Properties, 2022. URL <https://arxiv.org/abs/2210.10681>. arXiv :2210.10681.
- Z. AGATHE-NERINE : Multivariate Hawkes processes on inhomogeneous random graphs. *Stochastic Process. Appl.*, 152:86–148, 2022. ISSN 0304-4149. URL <https://doi.org/10.1016/j.spa.2022.06.019>.
- Z. AGATHE-NERINE : Long-term stability of interacting Hawkes processes on random graphs. *Electronic Journal of Probability*, 28:1 – 42, 2023a. URL <https://doi.org/10.1214/23-EJP1006>.
- Z. AGATHE-NERINE : Stability of wandering bumps for Hawkes processes interacting on the circle, 2023b. URL <https://arxiv.org/abs/2307.05982>. arXiv :2307.05982.
- K. AIHARA, G. MATSUMOTO & Y. IKEGAYA : Periodic and nonperiodic responses of a periodically forced Hodgkin-Huxley oscillator. *J. Theoret. Biol.*, 109(2): 249–269, 1984. ISSN 0022-5193,1095-8541. URL [https://doi.org/10.1016/S0022-5193\(84\)80005-3](https://doi.org/10.1016/S0022-5193(84)80005-3).
- N. ALON & A. NAOR : Approximating the cut-norm via Grothendieck’s inequality. *SIAM Journal on Computing*, 35(4):787–803, jan. 2006. URL <https://doi.org/10.1137/s0097539704441629>.
- S.-I. AMARI : Dynamics of pattern formation in lateral-inhibition type neural fields. *Biological Cybernetics*, 27(2):77–87, 1977. URL <https://doi.org/10.1007/bf00337259>.
- K. ATHREYA & K. R. MURTHY : Feller’s renewal theorem for systems of renewal equations. *Journal of the Indian Institute of Science*, 58(10):437, 1976.
- J. BALADRON, D. FASOLI, O. FAUGERAS & J. 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. URL <https://doi.org/10.1186/2190-8567-2-10>.

- E. BAYRAKTAR, S. CHAKRABORTY & R. WU : Graphon mean field systems. *Ann. Appl. Probab.*, 33(5):3587–3619, 2023. ISSN 1050-5164,2168-8737. URL <https://doi.org/10.1214/22-aap1901>.
- L. BERTINI, G. GIACOMIN & C. POQUET : Synchronization and random long time dynamics for mean-field plane rotators. *Probab. Theory Related Fields*, 160(3-4):593–653, 2014. ISSN 0178-8051. URL <https://doi.org/10.1007/s00440-013-0536-6>.
- G. BET, F. COPPINI & F. R. NARDI : Weakly interacting oscillators on dense random graphs, 2020. arXiv :2006.07670.
- S. BHAMIDI, A. BUDHIRAJA & R. WU : Weakly interacting particle systems on inhomogeneous random graphs. *Stochastic Process. Appl.*, 129(6):2174–2206, 2019. ISSN 0304-4149,1879-209X. URL <https://doi.org/10.1016/j.spa.2018.06.014>.
- P. BILLINGSLEY : *Convergence of Probability Measures*. Wiley, New York, 1968.
- F. BOLLEY, I. GENTIL & A. GUILLIN : Uniform convergence to equilibrium for granular media. *Arch. Ration. Mech. Anal.*, 208(2):429–445, 2013. ISSN 0003-9527. URL <https://doi.org/10.1007/s00205-012-0599-z>.
- A. BONNET, M. MARTINEZ HERRERA & M. SANGNIER : Maximum likelihood estimation for Hawkes processes with self-excitation or inhibition. *Statist. Probab. Lett.*, 179:Paper No. 109214, 7, 2021. ISSN 0167-7152,1879-2103. URL <https://doi.org/10.1016/j.spl.2021.109214>.
- C. BORGS, J. T. CHAYES, L. LOVÁSZ, V. T. SÓS & K. VESZTERGOMBI : Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing. *Adv. Math.*, 219(6):1801–1851, 2008. ISSN 0001-8708. URL <https://doi.org/10.1016/j.aim.2008.07.008>.
- C. BORGS, J. T. CHAYES, L. LOVÁSZ, V. T. SÓS & K. VESZTERGOMBI : Convergent sequences of dense graphs II. Multiway cuts and statistical physics. *Ann. of Math. (2)*, 176(1):151–219, 2012. ISSN 0003-486X. URL <https://doi.org/10.4007/annals.2012.176.1.2>.
- C. BORGS, J. CHAYES, L. LOVÁSZ, V. SÓS & K. VESZTERGOMBI : Limits of randomly grown graph sequences. *European Journal of Combinatorics*, 32(7): 985–999, oct. 2011. URL <https://doi.org/10.1016/j.ejc.2011.03.015>.
- C. BORGS, J. T. CHAYES, H. COHN & Y. ZHAO : An  $L^p$  theory of sparse graph convergence II : LD convergence, quotients and right convergence. *Ann. Probab.*, 46(1):337–396, 2018. ISSN 0091-1798. URL <https://doi.org/10.1214/17-AOP1187>.
- C. BORGS, J. T. CHAYES, H. COHN & Y. ZHAO : An  $L^p$  theory of sparse graph convergence I : Limits, sparse random graph models, and power law distributions. *Trans. Amer. Math. Soc.*, 372(5):3019–3062, 2019. ISSN 0002-9947. URL <https://doi.org/10.1090/tran/7543>.

- W. H. BOSKING, Y. ZHANG, B. SCHOFIELD & D. FITZPATRICK : Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *The Journal of Neuroscience*, 17(6):2112–2127, mars 1997. URL <https://doi.org/10.1523/jneurosci.17-06-02112.1997>.
- M. BOSSY, O. FAUGERAS & D. 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. ISSN 2190-8567. URL <https://doi.org/10.1186/s13408-015-0031-8>.
- P. BRÉMAUD & L. MASSOULIÉ : Stability of nonlinear Hawkes processes. *The Annals of Probability*, p. 1563–1588, 1996. URL [https://projecteuclid.org/download/pdf\\_1/euclid.aop/1065725193](https://projecteuclid.org/download/pdf_1/euclid.aop/1065725193).
- P. C. BRESSLOFF : Spatiotemporal dynamics of continuum neural fields. *J. Phys. A*, 45(3):033001, 109, 2012. ISSN 1751-8113,1751-8121. URL <https://doi.org/10.1088/1751-8113/45/3/033001>.
- P. C. BRESSLOFF : *Waves in neural media*. Lecture Notes on Mathematical Modelling in the Life Sciences. Springer, New York, 2014. ISBN 978-1-4614-8865-1; 978-1-4614-8866-8. URL <https://doi.org/10.1007/978-1-4614-8866-8>. From single neurons to neural fields.
- P. C. BRESSLOFF & M. A. WEBBER : Front propagation in stochastic neural fields. *SIAM J. Appl. Dyn. Syst.*, 11(2):708–740, 2012. ISSN 1536-0040. URL <https://doi.org/10.1137/110851031>.
- A. BUDHIRAJA & R. WU : Some fluctuation results for weakly interacting multi-type particle systems. *Stochastic Processes and their Applications*, 126(8):2253–2296, août 2016. URL <https://doi.org/10.1016/j.spa.2016.01.010>.
- P. CATTIAUX, L. COLOMBANI & M. COSTA : Limit theorems for Hawkes processes including inhibition. *Stochastic Process. Appl.*, 149:404–426, 2022. ISSN 0304-4149. URL <https://doi.org/10.1016/j.spa.2022.04.002>.
- J. CHEVALLIER, A. DUARTE, E. LÖCHERBACH & 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. ISSN 0304-4149. URL <http://www.sciencedirect.com/science/article/pii/S030441491830022X>.
- J. CHEVALLIER : Mean-field limit of generalized Hawkes processes. *Stochastic Process. Appl.*, 127(12):3870–3912, 2017. ISSN 0304-4149. URL <https://doi.org/10.1016/j.spa.2017.02.012>.
- J. CHEVALLIER, A. MELNYKOVA & I. TUBIKANEC : Diffusion approximation of multi-class Hawkes processes : theoretical and numerical analysis. *Adv. in Appl. Probab.*, 53(3):716–756, 2021. ISSN 0001-8678. URL <https://doi.org/10.1017/apr.2020.73>.

- J. CHEVALLIER & G. OST : Fluctuations for spatially extended Hawkes processes. *Stochastic Process. Appl.*, 130(9):5510–5542, 2020. ISSN 0304-4149. URL <https://doi.org/10.1016/j.spa.2020.03.015>.
- F. CHUNG & L. LU : Connected components in random graphs with given expected degree sequences. *Annals of Combinatorics*, 6(2):125–145, nov. 2002. URL <https://doi.org/10.1007/p100012580>.
- H. L. CIHAK, T. L. EISSA & Z. P. KILPATRICK : Distinct excitatory and inhibitory bump wandering in a stochastic neural field, 2022. URL <https://arxiv.org/abs/2203.02438>. arXiv :2203.02438.
- L. COLOMBANI & P. L. BRIS : Chaos propagation in mean field networks of FitzHugh-Nagumo neurons, 2022. URL <https://arxiv.org/abs/2206.13291>. arXiv :2206.13291.
- F. COPPINI : Long time dynamics for interacting oscillators on graphs. *Ann. Appl. Probab.*, 32(1):360–391, 2022. ISSN 1050-5164. URL <https://doi.org/10.1214/21-aap1680>.
- F. COPPINI, H. DIETERT & G. GIACOMIN : A Law of Large numbers and Large Deviations for interacting diffusions on Erdős–Rényi graphs. *Stochastics and Dynamics*, 20(02):2050010, juil. 2019. URL <https://doi.org/10.1142/s0219493720500100>.
- F. COPPINI, E. LUÇON & C. POQUET : Central limit theorems for global and local empirical measures of diffusions on Erdős–Rényi graphs, 2022. URL <https://arxiv.org/abs/2206.06655>. arXiv :2206.06655.
- Q. CORMIER, E. TANRÉ & R. VELTZ : Long time behavior of a mean-field model of interacting neurons. *Stochastic Processes and their Applications*, 130(5):2553–2595, mai 2020. URL <https://doi.org/10.1016/j.spa.2019.07.010>.
- M. COSTA, C. GRAHAM, L. MARSALLE & V. C. TRAN : Renewal in Hawkes processes with self-excitation and inhibition. *Advances in Applied Probability*, 52(3):879–915, sept. 2020. URL <https://doi.org/10.1017/apr.2020.19>.
- M. CSÖRGŐ : *Quantile processes with statistical applications*, vol. 42 de *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1983. ISBN 0-89871-185-1. URL <https://doi.org/10.1137/1.9781611970289>.
- D. DALEY & D. VERE-JONES : *An Introduction to the Theory of Point Processes : Volume I : Elementary Theory and Methods*. Probability and Its Applications. Springer New York, 2006. ISBN 9780387215648. URL <https://books.google.fr/books?id=6Sv4BwAAQBAJ>.
- M. DAVYDOV : Propagation of chaos and poisson hypothesis for replica mean-field models of intensity-based neural networks, 2022. URL <https://arxiv.org/abs/2211.11490>. arXiv :2211.11490.

- F. DELARUE, J. INGLIS, S. RUBENTHALER & E. TANRÉ : Global solvability of a networked integrate-and-fire model of McKean–Vlasov type. *The Annals of Applied Probability*, 25(4):2096–2133, août 2015. URL <https://doi.org/10.1214/14-aap1044>.
- F. DELARUE & A. TSE : Uniform in time weak propagation of chaos on the torus, 2021. URL <https://arxiv.org/abs/2104.14973>. arXiv :2104.14973.
- S. DELATTRE, N. FOURNIER & M. HOFFMANN : Hawkes processes on large networks. *Ann. Appl. Probab.*, 26(1):216–261, 02 2016a. URL <https://doi.org/10.1214/14-AAP1089>.
- S. DELATTRE, G. GIACOMIN & E. LUÇON : A Note on Dynamical Models on Random Graphs and Fokker–Planck equations. *Journal of Statistical Physics*, 165(4):785–798, nov. 2016b. URL <https://doi.org/10.1007/s10955-016-1652-3>.
- A. DEMBO & O. ZEITOUNI : *Large deviations techniques and applications*, vol. 38 de *Applications of Mathematics (New York)*. Springer-Verlag, New York, second éd., 1998. ISBN 0-387-98406-2. URL <https://doi.org/10.1007/978-1-4612-5320-4>.
- P. DIACONIS & S. JANSON : Graph limits and exchangeable random graphs. *Rend. Mat. Appl. (7)*, 28(1):33–61, 2008. ISSN 1120-7183,2532-3350.
- C. DION, S. LEMLER & E. LÖCHERBACH : Exponential ergodicity for diffusions with jumps driven by a Hawkes process. *Theory of Probability and Mathematical Statistics*, 102:97–115, mars 2021. URL <https://doi.org/10.1090/tpms/1129>.
- S. DITLEVSEN & E. LÖCHERBACH : Multi-class oscillating systems of interacting neurons. *Stochastic Process. Appl.*, 127(6):1840–1869, 2017. ISSN 0304-4149. URL <https://doi.org/10.1016/j.spa.2016.09.013>.
- A. DUARTE, E. LÖCHERBACH & G. OST : Stability, convergence to equilibrium and simulation of non-linear Hawkes processes with memory kernels given by the sum of Erlang kernels. *ESAIM Probab. Stat.*, 23:770–796, 2019. ISSN 1292-8100,1262-3318. URL <https://doi.org/10.1051/ps/2019005>.
- R. M. DUDLEY : *Real Analysis and Probability*. Cambridge University Press, oct. 2002. URL <https://doi.org/10.1017/cbo9780511755347>.
- R. DURRETT : *Random Graph Dynamics*. Cambridge University Press, oct. 2006. URL <https://doi.org/10.1017/cbo9780511546594>.
- C. DUVAL, E. LUÇON & C. POUZAT : Interacting Hawkes processes with multiplicative inhibition. *Stochastic Process. Appl.*, 148:180–226, 2022. ISSN 0304-4149. URL <https://doi.org/10.1016/j.spa.2022.02.008>.
- G. B. ERMENTROUT & J. B. MCLEOD : Existence and uniqueness of travelling waves for a neural network. *Proc. Roy. Soc. Edinburgh Sect. A*, 123(3):461–478, 1993. ISSN 0308-2105. URL <https://doi.org/10.1017/S030821050002583X>.

- X. ERNY : Annealed limit for a diffusive disordered mean-field model with random jumps, 2023. URL <https://arxiv.org/abs/2210.13128>. arXiv :2210.13128.
- X. ERNY, E. LÖCHERBACH & D. LOUKIANOVA : Mean field limits for interacting Hawkes processes in a diffusive regime. *Bernoulli*, 28(1):125–149, 2022. ISSN 1350-7265. URL <https://doi.org/10.3150/21-bej1335>.
- S. N. ETHIER & T. G. KURTZ, éd. *Markov Processes*. John Wiley & Sons, Inc., mars 1986. URL <https://doi.org/10.1002/9780470316658>.
- P. FATT & B. KATZ : Spontaneous subthreshold activity at motor nerve endings. *The Journal of physiology*, 117(1):109–128, 1952.
- O. FAUGERAS & J. INGLIS : Stochastic neural field equations : a rigorous footing. *J. Math. Biol.*, 71(2):259–300, 2015. ISSN 0303-6812,1432-1416. URL <https://doi.org/10.1007/s00285-014-0807-6>.
- W. FELLER : On the integral equation of renewal theory. *The Annals of Mathematical Statistics*, 12(3):243–267, sept. 1941. URL <https://doi.org/10.1214/aoms/1177731708>.
- R. FITZHUGH : Impulses and physiological states in theoretical models of nerve membrane. *Biophysical Journal*, 1(6):445–466, juil. 1961. URL [https://doi.org/10.1016/s0006-3495\(61\)86902-6](https://doi.org/10.1016/s0006-3495(61)86902-6).
- A. GALVES & E. LÖCHERBACH : Infinite systems of interacting chains with memory of variable length—a stochastic model for biological neural nets. *J. Stat. Phys.*, 151(5):896–921, 2013. ISSN 0022-4715,1572-9613. URL <https://doi.org/10.1007/s10955-013-0733-9>.
- A. GEORGOPOULOS, J. KALASKA, R. CAMINITI & J. MASSEY : On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex. *The Journal of Neuroscience*, 2(11):1527–1537, nov. 1982. URL <https://doi.org/10.1523/jneurosci.02-11-01527.1982>.
- G. GIACOMIN, E. LUÇON & C. POQUET : Coherence stability and effect of random natural frequencies in populations of coupled oscillators. *J. Dynam. Differential Equations*, 26(2):333–367, 2014. ISSN 1040-7294,1572-9222. URL <https://doi.org/10.1007/s10884-014-9370-5>.
- G. GIACOMIN, K. PAKDAMAN, X. PELLEGRIN & C. POQUET : Transitions in Active Rotator Systems : Invariant Hyperbolic Manifold Approach. *SIAM Journal on Mathematical Analysis*, 44(6):4165–4194, 2012. URL <https://doi.org/10.1137/110846452>.
- G. GIACOMIN & C. POQUET : Noise, interaction, nonlinear dynamics and the origin of rhythmic behaviors. *Brazilian Journal of Probability and Statistics*, 29(2):460 – 493, 2015. URL <https://doi.org/10.1214/14-BJPS258>.

- G. GIACOMIN, C. POQUET & A. SHAPIRA : Small noise and long time phase diffusion in stochastic limit cycle oscillators. *J. Differential Equations*, 264(2): 1019–1049, 2018. ISSN 0022-0396. URL <https://doi.org/10.1016/j.jde.2017.09.029>.
- J. GUCKENHEIMER : Isochrons and phaseless sets. *J. Math. Biol.*, 1(3):259–273, 1975. ISSN 0303-6812. URL <https://doi.org/10.1007/BF01273747>.
- A. G. HAWKES : Point spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1):83–90, 1971. URL <https://doi.org/10.1093/biomet/58.1.83>.
- A. G. HAWKES & D. OAKES : A cluster process representation of a self-exciting process. *Journal of Applied Probability*, 11(3):493–503, sept. 1974. URL <https://doi.org/10.2307/3212693>.
- S. HEESSEN & W. STANNAT : Fluctuation limits for mean-field interacting nonlinear Hawkes processes. *Stochastic Process. Appl.*, 139:280–297, 2021. ISSN 0304-4149. URL <https://doi.org/10.1016/j.spa.2021.05.007>.
- P. HODARA & E. LÖCHERBACH : Hawkes processes with variable length memory and an infinite number of components. *Advances in Applied Probability*, 49(1):84–107, mars 2017. URL <https://doi.org/10.1017/apr.2016.80>.
- A. L. HODGKIN & A. F. HUXLEY : Currents carried by sodium and potassium ions through the membrane of the giant axon of loligo. *The Journal of Physiology*, 116(4):449–472, avr. 1952. URL <https://doi.org/10.1113/jphysiol.1952.sp004717>.
- W. HOEFFDING : The strong law of large numbers for u-statistics. Rap. tech., North Carolina State University. Dept. of Statistics, 1961.
- P. W. HOLLAND, K. B. LASKEY & S. LEINHARDT : Stochastic blockmodels : first steps. *Social Networks*, 5(2):109–137, 1983. ISSN 0378-8733. URL [https://doi.org/10.1016/0378-8733\(83\)90021-7](https://doi.org/10.1016/0378-8733(83)90021-7).
- R. HÖPFNER, E. LÖCHERBACH & M. THIEULLEN : Strongly degenerate time inhomogeneous SDEs : densities and support properties. Application to Hodgkin-Huxley type systems. *Bernoulli*, 23(4A):2587–2616, 2017. ISSN 1350-7265,1573-9759. URL <https://doi.org/10.3150/16-BEJ820>.
- J. INGLIS & J. MACLAURIN : A general framework for stochastic traveling waves and patterns, with application to neural field equations. *SIAM J. Appl. Dyn. Syst.*, 15(1):195–234, 2016. ISSN 1536-0040. URL <https://doi.org/10.1137/15M102856X>.
- J. JACOD & A. SHIRYAEV : *Limit theorems for stochastic processes*, vol. 288. Springer Science & Business Media, 2013.
- S. JANSON : Graphons, cut norm and distance, couplings and rearrangements. 4:76, 2013.



- Z. P. KILPATRICK & B. ERMENTROUT : Wandering bumps in stochastic neural fields. *SIAM J. Appl. Dyn. Syst.*, 12(1):61–94, 2013. URL <https://doi.org/10.1137/120877106>.
- J. KRÜGER & W. STANNAT : Front propagation in stochastic neural fields : a rigorous mathematical framework. *SIAM J. Appl. Dyn. Syst.*, 13(3):1293–1310, 2014. ISSN 1536-0040. URL <https://doi.org/10.1137/13095094X>.
- Y. KURAMOTO : Self-entrainment of a population of coupled non-linear oscillators. *In International Symposium on Mathematical Problems in Theoretical Physics (Kyoto Univ., Kyoto, 1975)*, vol. 39 de *Lecture Notes in Phys.*, p. 420–422. Springer, Berlin-New York, 1975. ISBN 3-540-07174-1.
- D. LACKER, K. RAMANAN & R. WU : Local weak convergence for sparse networks of interacting processes. *Ann. Appl. Probab.*, 33(2):643–688, 2023. ISSN 1050-5164,2168-8737. URL <https://doi.org/10.1214/22-aap1830>.
- E. LANG & W. STANNAT :  $L^2$ -stability of traveling wave solutions to nonlocal evolution equations. *J. Differential Equations*, 261(8):4275–4297, 2016. ISSN 0022-0396. URL <https://doi.org/10.1016/j.jde.2016.06.021>.
- E. LANG & W. STANNAT : Finite-size effects on traveling wave solutions to neural field equations. *J. Math. Neurosci.*, 7:Paper No. 5, 35, 2017. URL <https://doi.org/10.1186/s13408-017-0048-2>.
- L. LAPICQUE : Recherches quantitatives sur l’excitation électrique des nerfs traitée comme une polarisation. *Journal of Physiology and Pathology*, 9:620–635, 1907.
- L. LAPICQUE : Quantitative investigations of electrical nerve excitation treated as polarization. *Biol. Cybernet.*, 97(5-6):341–349, 2007. ISSN 0340-1200,1432-0770. URL <https://doi.org/10.1007/s00422-007-0189-6>. Translated reprint of the 1907 French original [J. Phys. Path. Gén. 9, 620–635], Translated by Nicolas Brunel and Mark C. W. van Rossum.
- L. LOVÁSZ : *Large networks and graph limits*, vol. 60. American Mathematical Soc., 2012.
- L. LOVÁSZ & B. SZEGEDY : Limits of dense graph sequences. *Journal of Combinatorial Theory, Series B*, 96(6):933–957, nov. 2006. URL <https://doi.org/10.1016/j.jctb.2006.05.002>.
- E. LUÇON : Quenched asymptotics for interacting diffusions on inhomogeneous random graphs. *Stochastic Process. Appl.*, 130(11):6783–6842, 2020. ISSN 0304-4149. URL <https://doi.org/10.1016/j.spa.2020.06.010>.
- E. LUÇON & C. POQUET : Long time dynamics and disorder-induced traveling waves in the stochastic Kuramoto model. *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques*, 53(3):1196 – 1240, 2017. URL <https://doi.org/10.1214/16-AIHP753>.

- E. LUÇON & C. POQUET : Emergence of oscillatory behaviors for excitable systems with noise and mean-field interaction : a slow-fast dynamics approach. *Comm. Math. Phys.*, 373(3):907–969, 2020. ISSN 0010-3616. URL <https://doi.org/10.1007/s00220-019-03641-y>.
- E. LUÇON & C. POQUET : Periodicity and longtime diffusion for mean field systems in  $\mathbb{R}^d$ , 2021a. URL <https://arxiv.org/abs/2107.02473>. arXiv :2107.02473.
- E. LUÇON & C. POQUET : Periodicity induced by noise and interaction in the kinetic mean-field FitzHugh-Nagumo model. *Ann. Appl. Probab.*, 31(2):561–593, 2021b. ISSN 1050-5164,2168-8737. URL <https://doi.org/10.1214/20-aap1598>.
- E. LUÇON & W. STANNAT : Transition from gaussian to non-gaussian fluctuations for mean-field diffusions in spatial interaction. *The Annals of Applied Probability*, 26(6):3840–3909, déc. 2016. URL <https://doi.org/10.1214/16-aap1194>.
- L. LUO : *Principles of Neurobiology*. Garland Science, sept. 2020. URL <https://doi.org/10.1201/9781003053972>.
- J. MACLAURIN : Phase Reduction of Waves, Patterns, and Oscillations Subject to Spatially Extended Noise. *SIAM J. Appl. Math.*, 83(3):1215–1244, 2023. ISSN 0036-1399. URL <https://doi.org/10.1137/21M1451221>.
- J. MACLAURIN & P. C. BRESSLOFF : Wandering bumps in a stochastic neural field : a variational approach. *Phys. D*, 406:132403, 9, 2020. ISSN 0167-2789. URL <https://doi.org/10.1016/j.physd.2020.132403>.
- A. D. MASI, A. GALVES, E. LÖCHERBACH & E. PRESUTTI : Hydrodynamic limit for interacting neurons. *Journal of Statistical Physics*, 158(4):866–902, nov. 2014. URL <https://doi.org/10.1007/s10955-014-1145-1>.
- 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)*, vol. Session 7 de *Lecture Series in Differential Equations*, p. 41–57. Air Force Office of Scientific Research, Office of Aerospace Research, United States Air Force, Arlington, VA, 1967.
- G. S. MEDVEDEV : The nonlinear heat equation on W-random graphs. *Archive for Rational Mechanics and Analysis*, 212(3):781–803, déc. 2013. URL <https://doi.org/10.1007/s00205-013-0706-9>.
- V. MOUNTCASTLE : The columnar organization of the neocortex. *Brain*, 120(4):701–722, avr. 1997. URL <https://doi.org/10.1093/brain/120.4.701>.
- J. NAGUMO, S. ARIMOTO & S. YOSHIKAWA : An active pulse transmission line simulating nerve axon. *Proceedings of the IRE*, 50(10):2061–2070, oct. 1962. URL <https://doi.org/10.1109/jrproc.1962.288235>.
- R. B. NELSEN : *An Introduction to Copulas*. Springer New York, 1999. URL <https://doi.org/10.1007/978-1-4757-3076-0>.

- S. L. NGUYEN, G. YIN & T. A. HOANG : On laws of large numbers for systems with mean-field interactions and Markovian switching. *Stochastic Process. Appl.*, 130(1):262–296, 2020. ISSN 0304-4149,1879-209X. URL <https://doi.org/10.1016/j.spa.2019.02.014>.
- Y. OGATA : Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical Association*, 83(401):9–27, mars 1988. URL <https://doi.org/10.1080/01621459.1988.10478560>.
- R. I. OLIVEIRA & G. H. REIS : Interacting diffusions on random graphs with diverging average degrees : Hydrodynamics and large deviations. *Journal of Statistical Physics*, 176(5):1057–1087, juil. 2019. URL <https://doi.org/10.1007/s10955-019-02332-1>.
- R. I. OLIVEIRA, G. H. REIS & L. M. STOLERMAN : Interacting diffusions on sparse graphs : hydrodynamics from local weak limits. *Electron. J. Probab.*, 25:Paper No. 110, 35, 2020. URL <https://doi.org/10.1214/20-ejp505>.
- I. OMELCHENKO, B. RIEMENSCHNEIDER, P. HÖVEL, Y. MAISTRENKO & E. SCHÖLL : Transition from spatial coherence to incoherence in coupled chaotic systems. *Physical Review E*, 85(2), fév. 2012. URL <https://doi.org/10.1103/physreve.85.026212>.
- S. OUADAH, S. ROBIN & P. LATOUCHE : Degree-based goodness-of-fit tests for heterogeneous random graph models : Independent and exchangeable cases. *Scandinavian Journal of Statistics*, 47(1):156–181, oct. 2019. URL <https://doi.org/10.1111/sjos.12410>.
- A. PAZY : *Semi-groups of linear operators and applications to partial differential equations*. University of Maryland, Department of Mathematics, College Park, Md., 1974. Department of Mathematics, University of Maryland, Lecture Note, No. 10.
- P. PFAFFELHUBER, S. ROTTER & J. STIEFEL : Mean-field limits for non-linear Hawkes processes with excitation and inhibition. *Stochastic Process. Appl.*, 153: 57–78, 2022. ISSN 0304-4149. URL <https://doi.org/10.1016/j.spa.2022.07.006>.
- C. POUZAT : Spike train analysis and modeling. *In Lascon VIII summer school*, 2020. URL : [https://plmlab.math.cnrs.fr/xtof/lascon\\_2020](https://plmlab.math.cnrs.fr/xtof/lascon_2020).
- A. PRODHOMME : Strong Gaussian approximation of metastable density-dependent Markov chains on large time scales. *Stochastic Process. Appl.*, 160:218–264, 2023. ISSN 0304-4149. URL <https://doi.org/10.1016/j.spa.2023.01.018>.
- M. B. RAAD, S. DITLEVSEN & E. LÖCHERBACH : Stability and mean-field limits of age dependent Hawkes processes. *Ann. Inst. Henri Poincaré Probab. Stat.*, 56(3):1958–1990, 2020. ISSN 0246-0203,1778-7017. URL <https://doi.org/10.1214/19-AIHP1023>.

- M. B. RAAD & E. LÖCHERBACH : Stability for Hawkes processes with inhibition. *Electron. Commun. Probab.*, 25:Paper No. 33, 9, 2020. URL <https://doi.org/10.1214/20-ecp312>.
- J. RINZEL & R. N. MILLER : Numerical calculation of stable and unstable periodic solutions to the Hodgkin-Huxley equations. *Math. Biosci.*, 49(1-2):27–59, 1980. ISSN 0025-5564,1879-3134. URL [https://doi.org/10.1016/0025-5564\(80\)90109-1](https://doi.org/10.1016/0025-5564(80)90109-1).
- C. ROSSANT, D. F. GOODMAN, B. FONTAINE, J. PLATKIEWICZ, A. MAGNUSSON & R. BRETTE : Fitting neuron models to spike trains. *Frontiers in Neuroscience*, 5, 2011. ISSN 1662-453X. URL <https://www.frontiersin.org/articles/10.3389/fnins.2011.00009>.
- L. SACERDOTE & M. T. GIRAUDO : Stochastic integrate and fire models : a review on mathematical methods and their applications. In *Stochastic biomathematical models*, vol. 2058 de *Lecture Notes in Math.*, p. 99–148. Springer, Heidelberg, 2013. ISBN 978-3-642-32156-6 ; 978-3-642-32157-3. URL [https://doi.org/10.1007/978-3-642-32157-3\\_5](https://doi.org/10.1007/978-3-642-32157-3_5).
- H. H. SCHAEFER : *Banach Lattices and Positive Operators*. Springer Berlin Heidelberg, 1974. URL <https://doi.org/10.1007/978-3-642-65970-6>.
- O. SHRIKI, D. HANSEL & H. SOMPOLINSKY : Rate models for conductance-based cortical neuronal networks. *Neural Computation*, 15(8):1809–1841, août 2003. URL <https://doi.org/10.1162/08997660360675053>.
- D. SMITH : Mechanisms of action potential propagation failure at sites of axon branching in the crayfish. *The Journal of physiology*, 301(1):243–259, 1980.
- A.-S. SZNITMAN : Topics in propagation of chaos. In P.-L. HENNEQUIN, éd. : *Ecole d'Été de Probabilités de Saint-Flour XIX — 1989*, p. 165–251, Berlin, Heidelberg, 1991. Springer Berlin Heidelberg. ISBN 978-3-540-46319-1.
- T. TAO : *Topics in random matrix theory*, vol. 132 de *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012. ISBN 978-0-8218-7430-1. URL <https://doi.org/10.1090/gsm/132>.
- J. TOUBOUL : Propagation of chaos in neural fields. *The Annals of Applied Probability*, 24(3):1298–1328, juin 2014. URL <https://doi.org/10.1214/13-aap950>.
- V. S. VARADARAJAN : On the convergence of sample probability distributions. *Sankhyā*, 19:23–26, 1958. ISSN 0036-4452.
- R. VELTZ & O. FAUGERAS : Local/global analysis of the stationary solutions of some neural field equations. *SIAM J. Appl. Dyn. Syst.*, 9(3):954–998, 2010. URL <https://doi.org/10.1137/090773611>.
- A. A. VERVEEN & H. E. DERKSEN : Fluctuation phenomena in nerve membrane. *Proceedings of the IEEE*, 56(6):906–916, 1968.

- H. R. WILSON & J. D. COWAN : Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical Journal*, 12(1):1–24, jan. 1972. URL [https://doi.org/10.1016/s0006-3495\(72\)86068-5](https://doi.org/10.1016/s0006-3495(72)86068-5).
- A. T. WINFREE : Patterns of phase compromise in biological cycles. *Journal of Mathematical Biology*, 1(1):73–93, mai 1974. ISSN 0303-6812,1432-1416. URL <https://doi.org/10.1007/BF02339491>.
- M. ZERNER : Quelques propriétés spectrales des opérateurs positifs. *Journal of Functional Analysis*, 72(2):381–417, juin 1987. URL [https://doi.org/10.1016/0022-1236\(87\)90094-2](https://doi.org/10.1016/0022-1236(87)90094-2).
- J. ZHU, Z. BRZEŹNIAK & E. HAUSENBLAS : Maximal inequalities for stochastic convolutions driven by compensated Poisson random measures in Banach spaces. *Ann. Inst. Henri Poincaré Probab. Stat.*, 53(2):937–956, 2017. ISSN 0246-0203. URL <https://doi.org/10.1214/16-AIHP743>.
- L. ZHU : Large deviations for Markovian nonlinear Hawkes processes. *Ann. Appl. Probab.*, 25(2):548–581, 2015. ISSN 1050-5164,2168-8737. URL <https://doi.org/10.1214/14-AAP1003>.