



HAL
open science

Point-process-based Markovian dynamics and their applications

Michel Davydov

► **To cite this version:**

Michel Davydov. Point-process-based Markovian dynamics and their applications. Probability [math.PR]. Ecole normale supérieure PSL, 2023. English. NNT : . tel-04268154

HAL Id: tel-04268154

<https://theses.hal.science/tel-04268154>

Submitted on 2 Nov 2023

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

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



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée au Département d'Informatique de l'ENS
Dans le cadre d'une cotutelle avec Inria de Paris

**Dynamiques markoviennes à base de processus ponctuels
et leurs applications**

Soutenue par

Michel Davydov

Le 25/09/2023

École doctorale n°386

Mathématiques Paris Centre

Spécialité

Mathématiques

Composition du jury :

François Baccelli INRIA Paris & DI ENS	<i>Directeur de thèse</i>
Eva Löcherbach Université Panthéon Sorbonne	<i>Rapporteuse et Présidente du Jury</i>
Giovanni Luca Torrisi CNR	<i>Rapporteur</i>
Sergei Foss Heriot Watt University	<i>Examineur</i>
Nicolas Gast INRIA Grenoble	<i>Examineur</i>
Eric Luçon Université Paris-Cité	<i>Examineur</i>
Kavita Ramanan Brown University	<i>Examinatrice</i>
Thibaud Taillefumier University of Texas, Austin	<i>Invité</i>

Acknowledgements

Writing this section feels like the hardest part of the PhD: you have to bear with the pressure of forgetting someone important and making an enemy for life, knowing that this is the only part of your thesis manuscript that most people will bother reading. As such, I will try to be brief (nobody wants to read 10 pages of gushing praise) but as exhaustive as possible. If anyone is looking for their name in this section and doesn't find it, please find it in your heart to forgive me.

The first thanks of this section must go to the two people who have very graciously accepted to read beyond this section of the manuscript and provide invaluable feedback on it: my deepest gratitude to Eva and Gianluca for being the best reviewers one could hope to have, surprisingly quick over the summer holidays and yet incredibly in-depth and precise. When one is writing a thesis, it is impossible to fully escape the terror that sometime, somewhere, a critical mistake will be found and your reviews have helped alleviate that fear as much as it is possible. I would also like to thank all the other members of the jury, Sergei, Nicolas, Eric, Kavita and Thibaud, some of which have to wake up quite early for the defense, and for this I apologize. I would like in addition to specifically thank Thibaud for hosting me in Austin at the very beginning of the PhD and for answering my many questions throughout, especially those pertaining to links with computational neuroscience. It has been a pleasure to work with you! Additional special thanks must go to Nicolas (as well as Bruno Gaujal) for being part of my "Comité de suivi" during my PhD.

I would also like to use this opportunity to thank Antonio Galves. We've only met in person at a workshop at IHP relatively briefly, but I was struck by his enthusiasm, his energy and his kindness. He will be missed.

Now, I must go on to try to properly thank the three Fs - friends, family, and François.

Resisting the temptation to save the best for last, I will start with you, François, for you were truly an inspiration and the best possible thesis supervisor anyone could hope to have. You made me feel quite guilty when some of my friends were sharing "horror stories" about their relationships with their supervisors and I had nothing to volunteer! More seriously, I think that you fostered an incredible

environment for a first "real" experience with the world of research, and I'm truly grateful for the opportunity I had to learn under your tutelage. I could continue gushing on and on, but I'm sure similarly warm words will be found in all the manuscripts of your other students, past, present or future.

Speaking of other students, this seems like the best time to thank all my colleagues, brief or otherwise, during this PhD. Out of all these, my first words must go to the "Baccelli Boys and Girls", as we called our Whatsapp group. It's been truly a privilege working at your side and learning so much from you all, mathematics-related or otherwise. Ke, thank you for all the tennis lessons and teaching me so much about different perspectives. Bharath, thank you for making me feel less guilty about taking two desserts at Agraf (and for leading the way to the Agraf cafeteria in the first place!). Sayeh, thank you for all the chocolate "donations" and for patenting the word "asleepy". Pierre, thank you for all the trivia facts, and I officially forgive you for vetoing the calzone restaurant in Ithaca. Of course, I should also mention all the wonderful interns (some of which are now PhD students in our team!), who helped revitalize the group when it really needed it, so thank you Solène, Alessia, Thibaud, Baptiste and Florian! Alessia, thanks for the best tiramisu I've ever had. Guodong, thank you for the rice cooker! Thanks as well to the other members of ERC Nemo, Sanket, Ahmad, Romain. Natasha, thanks to you I will pay proper respects to foliage watching in the US.

Of course, I would be remiss not to mention the other members of the DYOGENE team: Bartek, our fearless leader after the schism; Christine, always willing to lead the way to Agraf or to go play tennis on clay; Ali, whose motivation in reading groups was peerless; Roman, with whom culinary disagreements were always a source of joy; Pierre Brémaud, whose brief appearances were always a pleasant surprise. And even though they are technically not in DYOGENE anymore, thanks to all the ARGO members with whom I interacted, at Agraf or elsewhere, with a special mention for Ana, whose willingness to laugh even at my worst jokes was quite gratifying, Ilya and Jakob. This also seems like a good place to thank people from the MAMBA team I've had the pleasure to interact with, namely Philippe, Jana and Lucie. A notable mention must also go to all the helpful staff at Inria, namely Julien and Marion.

Since I had the privilege of being part of not one, but two labs, I would like to thank everyone at LINCS, either at the original site near Place d'Italie, or at the new location in Palaiseau (one of these is more accessible than the other, though...). In particular, François (the other François), Ludovic, Fabien and Anna, you were an infinite source of positive energy and fascinating stories. I will also mention here Bruno, even though you weren't *technically* part of Lincs, I very much enjoyed our discussions at Télécom.

One name has come up several times in this section, and it is deserving of its own moment to shine: Agraf, the amazing cafeteria we could access through Inria - I have a feeling it will be dearly missed in the years to come.

Another category deserving a highlight is all the wonderful people I met at conferences and workshops: among others, Apolline, all the way back in Guidel; Sebastian from Grenoble; Mariem and Tara in Dunkerque; Mirabel, Philip, Jaap, Ellen, Eyal, Lhassa in Ithaca; Eva in Nancy; too many people to mention them all at the workshop at Institut Henri Poincaré (shoutout to the amazing staff there!); and more recently Lev in Rennes; Anna, Julien, Maxime, Yoan and Sonia in Marseille. Apolline, thank you for your interest in murder parties! Mirabel, I'm still not over the fact that *Kin-dza-dza* is your favorite movie, and I'm not sure I will ever be. Eva and Vaki, thanks again for hosting us in San Diego! Nhi, thanks for all the gossip and the chicken facts!

A special mention has to go to Zoé and Laetitia, who have endured my presence since we sat next to each other during classes at ENS. The fact that we still cross ways at conferences and more generally keep in touch is truly a delight.

Speaking of ENS, I must thank all the wonderful people I've met in Cachan and kept in touch with over the years, especially during the extracurricular activities. I must start with the members of the "team Tisane", Ariane, Blandine and Célia, it's always a rare treat to talk with you. Another shoutout goes to the members of the "VIP Murder party" group, you know who you are ;)! Mathilde, it's always a pleasure with you, whether it is while writing about italian mafia or baking an apple pie. Karim, thanks for hosting us all at your *amazing* house. Constance, thank you for sharing with us the wisdom of the elders. Sardine, thank you for being an amazing organizer. Constaince, *vivere in sal, mori per sal* :p. Géraud, you are a great friend despite your questionable taste in board games. Lucas, I'm still looking for a counter-example to your banana-carrot theory. Simon, you're one of the funniest people I've ever met. Xavier, how's the CR going? It's over 7 years overdue now... Zéphyr, thank you for always putting a laughing emoji whenever I make a joke in a conversation online. Paul, your passion for cinnamon and mustard never ceased to astound me. Joseph, your gift for poetry is unique. Aurore, Alice, the weekly online meetings that you initiated during the first lockdown have always been a pleasure. Of course, there are many more people deserving of a mention, please forgive me if you're not given a direct shoutout here. Antoine, thanks for bearing with me since middle school!

Juliette, I don't know what to say except thank you. This PhD wouldn't have been the same without you by my side. Please give a nice belly rub to Chtulhu from me next time you see him! Huge thanks as well to Nadine, Vincent, Florian and Simon.

I must confess that in the end, I couldn't resist leaving the best for last. So I

will finish this section thanking my family, especially my parents. It's a bit trite to say that I never would've ended up where I did without you, but that doesn't make it false. Thank you so, so much, for everything you did over the years. I will just add that it's a rare privilege to have a father who can actually understand what I did during my PhD, and even offer advice! Stas, Nastia, you're the best siblings one can hope to have. Elena, si tu lis ceci c'est que ton niveau en français a bien progressé !

The very final thanks of this section goes to you, the reader. I hope you find this manuscript worthy of your attention.

Contents

Acknowledgements	1
Contents	5
Résumé détaillé	7
1 Introduction	21
1.1 Motivation	21
1.2 Primer on point process dynamics	24
1.3 Primer on neural intensity-based models	27
1.4 Primer on mean-field theory	30
1.5 Primer on Poisson approximation	34
1.6 Primer on replica-mean-field models	37
1.7 Primer on epidemiological models and migration-contagion processes	40
2 Replica-mean-field limits for discrete-time processes	45
2.1 Introduction	45
2.2 Discrete-time replica-mean-field limit for FIAPs	46
3 Replica-mean-field limits for continuous-time processes	73
3.1 Introduction	73
3.2 Replica-mean-field limits for intensity-based neural networks	74
3.3 Replica-mean-field limit of continuous-time fragmentation-interaction-aggregation processes	110
4 Replica-mean-field limit of a migration-contagion process on a network	139
4.1 Introduction	139
4.2 Replica-mean-field limit of a migration-contagion process on a network	139
Bibliography	149

Résumé détaillé

Dans cette thèse, nous étudions différents modèles d'agents en interaction sur des réseaux, avec un intérêt particulier donné aux phénomènes biologiques. Nous nous intéressons surtout à un modèle de neurosciences computationnelles, le modèle de Galves-Löcherbach, ainsi qu'à un modèle d'épidémiologie, les processus de migration-contagion, même si nous nous attachons à présenter les résultats obtenus dans un plus grand degré de généralité.

Nous nous intéressons à l'évolution temporelle de phénomènes qui peuvent être modélisés par un réseau dans lequel chaque noeud correspond à un individu caractéristique du phénomène (un individu, une particule, un neurone, par exemple) et chaque arête correspond à une possibilité d'interaction entre deux noeuds. Un coefficient associé à chaque arête indique l'information transmise lors d'une interaction. A chaque noeud est associé un processus ponctuel, c'est-à-dire dans le cadre unidimensionnel une suite de temps aléatoires, représentant les instants d'interaction avec les noeuds voisins. On se place dans le cadre où ce processus ponctuel admet une intensité stochastique : autrement dit, si \mathcal{F}_t est une filtration sur \mathbb{R} et N est un processus ponctuel, il existe un processus \mathcal{F}_t -prévisible tel que pour tous $s < t \in \mathbb{R}$,

$$\mathbf{E} [N(s, t) | \mathcal{F}_s] = \mathbf{E} \left[\int_s^t \lambda(u) du | \mathcal{F}_s \right].$$

Pour décrire une dynamique sur un réseau, on peut alors considérer que l'état du système est régi par les intensités stochastiques des processus ponctuels associés aux noeuds du réseau. Dans les dynamiques considérées, cela revient à établir un système d'équations différentielles stochastiques avec un aléa d'origine poissonnienne vérifiées par ces intensités. Chaque équation décrit à un niveau qu'on peut qualifier de "microscopique" l'état du noeud auquel elle est associée. Concrètement, dans les phénomènes que l'on va étudier dans cette thèse, l'état d'un noeud à un instant t est égal à son état initial auquel on agrège les interactions reçues des noeuds voisins jusqu'au temps t et auquel on retire les effets des départs depuis le noeud.

En général, la description "microscopique" d'un tel système n'est pas tractable, dans le sens où il n'existe en général pas de solution analytique au système d'équations

différentielles stochastiques. Une idée naturelle est donc de le simplifier d'une certaine manière pour pouvoir obtenir des formes closes pour des quantités d'intérêt.

Comme la complexité du problème provient avant tout des dépendances entre les noeuds, il est naturel de choisir une simplification qui les rend indépendants. Une approche classique pour ce faire est de considérer le régime de champ moyen, qui consiste à regarder la limite d'échelle du système dynamique lorsque le nombre de noeuds dans le réseau tend vers l'infini et que les arrivées à un noeud sont considérées comme étant une moyenne empirique de toutes les interactions du réseau. Un argument de type loi des grands nombres permet dans de nombreux modèles de justifier rigoureusement cette convergence et de montrer que les noeuds deviennent asymptotiquement indépendants, raison pour laquelle on appelle cette convergence propagation du chaos.

Dans les modèles classiques de champ moyen, le réseau considéré doit être supposé pleinement connecté et l'effet des interactions doit être faible, typiquement inversement proportionnel au nombre de noeuds dans le réseau. Ces pré-requis représentent des contraintes significatives sur les architectures des réseaux considérés et donc sur les types de phénomènes pour lesquels l'approximation de champ moyen présente un intérêt.

Diverses approches ont été développées ces dernières années pour dépasser ces limitations, dont certaines sont présentées dans la section 1.4. Celle sur laquelle nous nous sommes concentrés est l'hypothèse poissonnienne. Initialement popularisée par Kleinrock et Dobrushin dans le domaine des réseaux de télécommunications, elle consiste à considérer que les noeuds dans un réseau sont indépendants et que les arrivées à chaque noeud sont données par un processus de Poisson.

Cette hypothèse est souvent conjecturée sans être prouvée ou alors seulement vérifiée numériquement. Une question naturelle est alors la construction d'un système physique à partir du modèle initial qui, une fois correctement échelonné, converge vers le système initial sous hypothèse poissonnienne. Autrement dit, on cherche à prouver un résultat de type propagation du chaos pour le "bon" système physique.

Un tel système a été introduit par Baccelli et Taillefumier pour étudier le modèle de Galves-Löcherbach, utilisé en neurosciences computationnelles pour décrire le comportement d'une population de neurones dans le cerveau en décrivant chaque neurone via l'évolution du potentiel électrique de sa membrane.

Cela conduit au système d'équations suivant, pour une population de K neurones, pour $1 \leq i \leq K$ et $t \geq 0$:

$$\lambda_i(t) = \lambda_i(0) + \sum_{j \neq i} \mu_{j \rightarrow i} \int_0^t N_j(ds) + \int_0^t (r_i - \lambda_i(s)) N_i(ds) + \frac{1}{\tau_i} \int_0^t (b_i - \lambda_i(s)) ds, \quad (1)$$

où λ_i est l'intensité stochastique du processus ponctuel N_i , b_i, r_i et τ_i sont des constantes strictement positives, et $\mu_{j \rightarrow i}$ sont des réels positifs.

Bacelli et Taillefumier ont introduit un modèle de champ moyen à répliques, construit à partir du modèle de Galves-Löcherbach. Ils ont considéré M répliques du modèle initial de K neurones avec les mêmes conditions initiales et interagissant comme suit : lorsqu'une interaction devait se produire entre deux noeuds i et j dans une des répliques en raison d'un départ depuis le noeud i , à la place une réplique est choisie aléatoirement et uniformément et l'interaction se produit entre le noeud i de la réplique initiale et le noeud j de la réplique ainsi choisie. Ils ont ensuite étudié ce modèle dans le régime stationnaire à la limite du nombre infini de répliques, en conjecturant que cette limite existait et correspondait précisément au modèle initial sous l'hypothèse poissonnienne, et en particulier que les répliques étaient asymptotiquement indépendantes.

Un des objectifs de ce travail a été de prouver cette conjecture. Plus généralement, nous nous sommes intéressés à ces dynamiques de champ moyen à répliques dans différents cadres et à divers degrés de généralité, avec le modèle de Galves-Löcherbach comme exemple motivant. Nous avons donc introduit une classe de processus, les processus d'interaction-agrégation-fragmentation (abrégé en PIAF par la suite), initialement en temps discret, puis en temps continu, à partir desquels nous avons construit des dynamiques avec des champs moyens à répliques.

Rappelons ici la définition d'un PIAF en temps discret.

Définition 0.0.1. *Un élément de la classe \mathcal{C} des processus d'interaction-agrégation-fragmentation en temps discret est déterminé par :*

- *Un entier K représentant le nombre de noeuds;*
- *La donnée de conditions initiales pour les variables d'état discrètes à l'instant 0, notées $\{X_i\}$, où $i \in \{1, \dots, K\}$;*
- *La donnée de variables aléatoires de fragmentation $\{U_i\}$, qui sont i.i.d., uniformes sur $[0, 1]$ et indépendantes de $\{X_i\}$, pour $i \in \{1, \dots, K\}$;*
- *La donnée de fonctions de fragmentation $\{g_{1,i} : \mathbb{N} \rightarrow \mathbb{N}\}_{i \in \{1, \dots, K\}}$ et $\{g_{2,i} : \mathbb{N} \rightarrow \mathbb{N}\}_{i \in \{1, \dots, K\}}$;*
- *La donnée de fonctions d'interaction bornées $\{h_{j \rightarrow i} : \mathbb{N} \rightarrow \mathbb{N}\}_{i,j \in \{1, \dots, K\}}$;*
- *La donnée de probabilités d'activation $\{\sigma_i(0), \sigma_i(1), \dots\}_{i \in \{1, \dots, K\}}$ vérifiant les conditions $\sigma_i(0) = 0$ et $0 < \sigma_i(1) \leq \sigma_i(2) \leq \dots \leq 1$ pour tout i .*

Les dynamiques associées prennent en entrée les variables d'état initiales $\{X_i\}$ et définissent les variables d'état après un pas de temps comme

$$Y_i = g_{1,i}(X_i) \mathbb{1}_{\{U_i < \sigma_i(X_i)\}} + g_{2,i}(X_i) \mathbb{1}_{\{U_i > \sigma_i(X_i)\}} + A_i, \quad \forall i = 1, \dots, K, \quad (2)$$

avec les processus d'arrivées

$$A_i = \sum_{j \neq i} h_{j \rightarrow i}(X_j) \mathbb{I}_{\{U_j < \sigma_j(X_j)\}}, \quad \forall i = 1, \dots, K. \quad (3)$$

L'interprétation est la suivante : le noeud i s'active avec probabilité $\sigma_i(k)$ si son état X_i est égal à k . L'état de ce noeud se fragmente en $g_{1,i}(k)$ en cas d'activation et en $g_{2,i}(k)$ sinon. Cette activation entraîne une arrivée de $h_{i \rightarrow j}(k)$ unités au noeud j . Ainsi, les fonctions d'interaction encodent la structure du réseau. La variable A_i donne le nombre total d'arrivées au noeud i . Cette variable est agrégée à l'état du noeud comme vu dans (2). Notons que la condition $\sigma_i(0) = 0$ pour tout i assure que les variables d'état dans l'état 0 ne peuvent pas être fragmentées.

La classe \mathcal{C} des PIAFs inclut diverses dynamiques de réseau d'intérêt en théorie des files d'attente ou en biologie mathématique. Par exemple, en prenant $g_{1,i}(k) = k - 1$, $g_{2,i}(k) = k$ et $h_{j \rightarrow i}(k) = \mathbb{I}_{\{i=j+1 \pmod K\}}$, on récupère une description de réseaux de files d'attente dits de Gordon-Newell [50]. En prenant $g_{1,i}(k) = 0$, $g_{2,i}(k) = k$ et $h_{j \rightarrow i}(k) = \mu_{j \rightarrow i} \in \mathbb{N}$, on définit une version discrète de dynamiques de Galves-Löcherbach pour des réseaux de neurones. En prenant $g_{1,i}(k) = \lfloor \frac{k}{2} \rfloor$ et $g_{2,i}(k) = k + 1$, on retrouve des processus d'agrégation-fragmentation modélisant, par exemple, des réseaux de communication TCP [7].

On peut ensuite, comme dit précédemment, introduire le champ moyen à répliques d'un PIAF. L'état d'un modèle à M répliques est alors donné par les variables d'état $X_{m,i}^M$, où m est l'indice de la réplique et i correspond à l'indice du noeud dans le réseau d'origine. Plus rigoureusement:

Définition 0.0.2. *Pour tout processus dans \mathcal{C} , la dynamique à M répliques associée est entièrement déterminée par*

- *La donnée de conditions initiales pour les variables d'état au temps 0, notées $\{X_{m,i}^M\}$, où $m \in \{1, \dots, M\}$ et $i \in \{1, \dots, K\}$, tels que pour tous M, m et i , $X_{m,i}^M = X_i$;*
- *La donnée de variables aléatoires de fragmentation $\{U_{m,i}\}$, qui sont i.i.d., uniformes sur $[0, 1]$ et indépendantes de $\{X_{m,i}^M\}$, où $m \in \{1, \dots, M\}$ et $i \in \{1, \dots, K\}$;*
- *La donnée de variables de routage i.i.d. $\{R_{(n,j) \rightarrow i}^M\}$ indépendantes de $\{X_{m,i}^M\}$ et $\{U_{m,i}\}$, uniformément distribuées sur $\{1, \dots, M\} \setminus \{n\}$ pour tous $i, j \in \{1, \dots, K\}$ et $n \in \{1, \dots, M\}$. En d'autres termes, si $R_{(n,j) \rightarrow i}^M = m$, alors une activation éventuelle du noeud j dans la réplique n au temps 0 induit une arrivée de taille $h_{j \rightarrow i}(X_{n,j}^M)$ au noeud i de la réplique m , et m est choisi uniformément parmi les répliques et indépendamment des variables d'état. On notera que ces variables sont définies que l'activation se produise ou*

non, et que pour $i' \neq i$, l'activation en question entrainera une arrivée au noeud i' dans la réplique m' , avec m' déterminé de la même manière mais indépendamment de n .

Alors les variables d'état après un pas de temps, notées $\{Y_{m,i}^M\}$, sont données par les équations

$$Y_{m,i}^M = g_{1,i}(X_{m,i}^M) \mathbb{I}_{\{U_{m,i} < \sigma_i(X_{m,i}^M)\}} + g_{2,i}(X_{m,i}^M) \mathbb{I}_{\{U_{m,i} > \sigma_i(X_{m,i}^M)\}} + A_{m,i}^M, \quad (4)$$

où $g_{1,i}$, $g_{2,i}$ sont les fonctions de fragmentation, σ_i sont les probabilités d'activation, et où

$$A_{m,i}^M = \sum_{n \neq m} \sum_{j \neq i} h_{j \rightarrow i}(X_{n,j}^M) \mathbb{I}_{\{U_{n,j} < \sigma_j(X_{n,j}^M)\}} \mathbb{I}_{\{R_{(n,j) \rightarrow i}^M = m\}} \quad (5)$$

est le nombre d'arrivées au noeud i de la réplique m .

Nous avons prouvé une propriété de propagation du chaos pour ces dynamiques au sens suivant: si les variables d'état au temps 0 vérifient une propriété que nous appelons indépendance asymptotique par paires, alors cette propriété est toujours vérifiée au temps 1. Cela permet de propager la propriété en question, définie ci-après, à tout temps fini.

Définition 0.0.3. *Etant donné $M \in \mathbb{N}$, et un tableau de variables aléatoires discrètes $Z = \{Z_{m,i}^M\}_{1 \leq m \leq M, 1 \leq i \leq K}$ tel que pour tout M , les variables aléatoires $Z_{m,i}^M$ sont échangeables en m , on dit que les variables aléatoires $Z_{n,i}^M$ sont asymptotiquement indépendantes par paires, ce qu'on note $\text{PAI}(Z)$, s'il existe des variables aléatoires $(\tilde{Z}_i)_{i \in \{1, \dots, K\}}$ telles que $\forall (n, i) \neq (m, j), \forall u, v \in [0, 1]$,*

$$\lim_{M \rightarrow \infty} \mathbf{E}[u^{Z_{n,i}^M} v^{Z_{m,j}^M}] = \mathbf{E}[u^{\tilde{Z}_i}] \mathbf{E}[v^{\tilde{Z}_j}]. \quad (6)$$

Nous pouvons alors énoncer rigoureusement le théorème suivant, qui est le résultat principal du chapitre 2:

Théorème 0.0.4. *Avec les notations précédentes, $\text{PAI}(X)$ implique $\text{PAI}(Y)$. De plus, les arrivées à un noeud donné sont asymptotiquement distribuées selon une loi de Poisson composée et sont indépendantes des états des noeuds.*

La preuve repose sur une analyse de la fonction génératrice de $Y_{m,i}$ et sur la loi des grands nombres triangulaire suivante, qui est induite par la propriété d'indépendance asymptotique par paires :

Définition 0.0.5. *Etant donné $M \in \mathbb{N}$ et un tableau de variables aléatoires discrètes $Z = \{Z_m^M\}_{m \in \{1, \dots, M\}}$ tel que pour tout M , les variables aléatoires Z_m^M*

sont échangeables en m , on dit que Z vérifie la loi triangulaire des grands nombres TLLN(Z) s'il existe une variable aléatoire discrète \tilde{Z} telle que pour toutes les fonctions $f : \mathbb{N} \rightarrow \mathbb{R}$ à support compact, on a la limite suivante dans L^2 :

$$\lim_{M \rightarrow \infty} \frac{1}{M} \sum_{m=1}^M f(Z_m^M) = \mathbf{E}[f(\tilde{Z})]. \quad (7)$$

Dans le cas particulier du modèle de Galves-Löcherbach en temps discret, on dispose de la forme explicite suivante pour les arrivées dans la limite du nombre infini de répliques: pour $i \in \{1, \dots, K\}$ et $z \in [0, 1]$,

$$\mathbf{E} \left[z^{\tilde{A}_i} \right] = e^{\theta_i \sum_{j \neq i} (z^{\mu_{j \rightarrow i} - 1})} = \prod_{j \neq i} e^{\theta_i (z^{\mu_{j \rightarrow i} - 1})}, \quad (8)$$

où $\theta_i = \mathbf{E} \left[\sigma_i(\tilde{X}_i) \right]$.

Nous avons ensuite cherché à généraliser ce résultat de propagation du chaos au modèle de Galves-Löcherbach en temps continu, étant donné qu'il s'agissait de la motivation initiale du travail. A partir du modèle de Galves-Löcherbach (1), on introduit une version à M répliques, similairement à la situation en temps discret.

Formellement, pour $1 \leq m \leq M, 1 \leq i, j \leq K$, soient $\{V_{(m,i) \rightarrow j}^M(t)\}_{t \in \mathbb{R}}$ des processus stochastiques tels que pour tout point de $N_{m,i}^M$, les variables aléatoires $\{V_{(m,i) \rightarrow j}^M(T)\}_j$ sont mutuellement indépendantes, indépendantes du passé et uniformément distribuées sur $\{1, \dots, M\} \setminus \{m\}$. Ici, $V_{(m,i) \rightarrow j}^M(T)$ donne l'indice de la réplique dans laquelle un départ du neurone i dans la réplique m à l'instant T provoque une arrivée au neurone j .

Les intensités stochastiques associées aux processus ponctuels vont alors être solutions du système d'équations différentielles stochastiques suivant :

$$\begin{aligned} \lambda_{m,i}^M(t) &= \lambda_{m,i}^M(0) + \frac{1}{\tau_i} \int_0^t (b_i - \lambda_{m,i}^M(s)) ds \\ &+ \sum_{j \neq i} \mu_{j \rightarrow i} \sum_{n \neq m} \int_0^t \mathbb{I}_{\{V_{(n,j) \rightarrow i}^M(s) = m\}} N_{n,j}^M(ds) + \int_0^t (r_i - \lambda_{m,i}^M(s)) N_{m,i}^M(ds). \end{aligned} \quad (9)$$

L'intuition motivant l'introduction de ce système physique reste identique à celle en temps discret : le routage induit un mélange uniforme entre les interactions qui devrait donner une limite poissonnienne lorsque le nombre de répliques augmente. Par ailleurs, le caractère uniforme du routage suggère une indépendance asymptotique entre les répliques, puisque la probabilité que deux noeuds donnés dans deux répliques données interagissent est d'ordre $\frac{1}{M}$.

Pour obtenir un résultat de convergence, nous faisons l'hypothèse suivante sur les conditions initiales:

Hypothèse 0.0.6. *Il existe $\xi_0 > 0$ tel que pour tous $p \geq 1$, $1 \leq m \leq M$, $1 \leq i \leq K$, et tout $0 < \xi \leq \xi_0$, $\mathbf{E}[e^{\xi \lambda_{m,i}^p(0)}] < \infty$.*

Nous introduisons le système d'équations différentielles stochastiques suivant qui est le candidat naturel pour la dynamique limite lorsque le nombre de répliques tend vers l'infini : les instants d'interaction sont remplacées par des processus de Poisson, ce qui correspond précisément à l'hypothèse poissonnienne. Nous notons avec des tildes tout ce qui correspond à ce régime.

Nous considérons des processus ponctuels $\tilde{N}_1, \dots, \tilde{N}_K$ sur \mathbb{R}^+ avec (\mathcal{F}_t) intensités stochastiques respectives $\tilde{\lambda}_1, \dots, \tilde{\lambda}_K$, où \mathcal{F}_t est l'histoire du réseau, vérifiant les équations différentielles stochastiques suivantes pour $t > 0$, et pour $1 \leq i \leq K$,

$$\tilde{\lambda}_i(t) = \tilde{\lambda}_i(0) + \sum_{j \neq i} \mu_{j \rightarrow i} \tilde{A}_{j \rightarrow i}(t) + \int_0^t (r_i - \tilde{\lambda}_i(s)) \tilde{N}_i(ds), \quad (10)$$

où $\tilde{A}_{j \rightarrow i}$ sont des processus de Poisson inhomogènes avec intensités respectives $a_j(t) = \int_0^t \mathbf{E}[\tilde{\lambda}_j(s)] ds = \mathbf{E}[\tilde{N}_j([0, t])]$ et $(\tilde{\lambda}_1(0), \dots, \tilde{\lambda}_K(0))$ satisfont l'Hypothèse 0.0.6.

Nous avons alors le résultat de convergence suivant :

Théorème 0.0.7. *Soit $t \in [0, T]$. Soit*

$$A_{m,i}^M(t) = \sum_{j \neq i} \mu_{j \rightarrow i} \sum_{n \neq m} \int_0^t \mathbb{1}_{\{V_{(n,j) \rightarrow i}^M(s) = m\}} N_{n,j}^M(ds),$$

où $N_{m,i}^M$ est défini par (9), et soit

$$\tilde{A}_i(t) = \sum_{j \neq i} \mu_{j \rightarrow i} \tilde{A}_{j \rightarrow i}(t),$$

avec $\tilde{A}_{j \rightarrow i}(t)$ défini par (10).

Alors,

1. les processus $(\tilde{A}_1, \dots, \tilde{A}_K)$ sont indépendants, tout comme les processus $(\tilde{\lambda}_1, \dots, \tilde{\lambda}_K)$;
2. pour tout $(m, i) \in \{1, \dots, M\} \times \{1, \dots, K\}$, la variable aléatoire $A_{m,i}^M(t)$ converge en variation totale vers $\tilde{A}_i(t)$ quand $M \rightarrow \infty$;

3. pour tout $(m, i) \in \{1, \dots, M\} \times \{1, \dots, K\}$, la variable aléatoire $\lambda_{m,i}^M(t)$ définie par (3.4) converge en variation totale vers $\tilde{\lambda}_i(t)$ défini par (3.5) quand $M \rightarrow \infty$;
4. soit \mathcal{N} un sous-ensemble fini de \mathbb{N}^* . Pour tout $i \in \{1, \dots, K\}$, les processus $(A_{m,i}^M(\cdot))_{m \in \mathcal{N}}$ et $(\lambda_{m,i}^M(\cdot))_{m \in \mathcal{N}}$ convergent faiblement dans l'espace de Skorokhod $D([0, T]^{\text{card } \mathcal{N}})$ muni de la topologie de Skorokhod produit vers $\text{card}(\mathcal{N})$ copies indépendantes des processus limites correspondants $(\tilde{A}_i(\cdot))$ et $(\tilde{\lambda}_i(\cdot))$ quand $M \rightarrow \infty$.

La preuve repose sur l'observation suivante : à la différence des modèles de champ moyen classiques, la simplification de type champ moyen dans les modèles à répliques vient des routages aléatoires entre répliques. Les arrivées à un noeud dans le modèle à M répliques consistent en une superposition de M processus ponctuels rares, ce qui explique informellement les limites poissonniennes. Cette observation nous conduit à fixer un instant $t \in \mathbb{R}^+$ et à considérer la variable aléatoire des arrivées jusqu'à l'instant t comme une somme d'un nombre aléatoire de variables de Bernoulli de paramètre $\frac{1}{M-1}$. La méthode de Chen-Stein est un candidat naturel pour obtenir une borne explicite en variation totale entre cette somme aléatoire et une variable aléatoire suivant une loi de Poisson. Nous généralisons la méthode classique pour obtenir cette borne en utilisant une propriété d'indépendance des routages conditionnellement au nombre de départs dans les autres noeuds. Il n'est pas évident de conclure que la borne obtenue par la méthode de Chen-Stein converge vers 0 lorsque M tend vers l'infini en raison de la présence d'un terme similaire à une norme L^1 d'une moyenne empirique de variables aléatoires centrées qui sont échangeables mais non indépendantes. Pour contourner l'analyse directe d'un tel terme, nous découplons les arrivées et les départs d'un noeud fixé en considérant la dynamique (3.4) comme le point fixe d'une certaine fonction sur l'espace des lois de probabilité sur l'espace des fonctions càdlàg muni d'une métrique le rendant complet. Cette procédure est inspirée d'une méthode courante pour prouver l'existence et l'unicité de solutions à des équations différentielles stochastiques, voir par exemple [75] ou [18]. Ici, nous l'utilisons pour prouver qu'une certaine propriété, en l'occurrence la convergence d'une moyenne empirique, est vraie au point fixe en prouvant que la propriété est préservée par la fonction et que la suite des itérées de la fonction converge vers son unique point fixe.

Étant donné le degré de généralité obtenu en temps discret, une question naturelle qui s'est posée a été de chercher à construire une classe de processus en temps continu analogue aux PIAFs pour laquelle le résultat de convergence du modèle de champ moyen à répliques vers une dynamique sous hypothèse poissonnienne reste valable.

Nous les introduisons en les définissant à travers leurs intensités stochastiques :

Définition 0.0.8. Soit $K \geq 2$. Un processus d'interaction-agrégation-fragmentation à temps continu, noté PIAFc par la suite, est défini comme une collection de processus ponctuels $(N_i)_{1 \leq i \leq K}$ admettant des intensités stochastiques (λ_i) par rapport à l'histoire du réseau telles que pour tout $t \in \mathbb{R}^+$, il existe

- des processus ponctuels $(\widehat{N}_{j \rightarrow i})_{1 \leq i \leq K, j \neq i}$ admettant des intensités stochastiques $(\widehat{\lambda}_{j \rightarrow i})$ dits processus d'interaction;
- des fonctions $(h_{j \rightarrow i})_{1 \leq i \leq K, j \neq i} : \mathbb{R} \mapsto \mathbb{R}$ dites fonctions d'interaction telles qu'il existe $H > 0$ satisfaisant pour tous i, j et tout $t \in \mathbb{R}$, $|h_{j \rightarrow i}(t)| \leq H$;
- des fonctions $(g_i)_{1 \leq i \leq K} : \mathbb{R} \mapsto \mathbb{R}^+$ et $(\sigma_i)_{1 \leq i \leq K} : \mathbb{R} \mapsto \mathbb{R}^+$ dites fonctions d'évolution autonome;
- une fonction lipschitzienne $f : \mathbb{R} \mapsto \mathbb{R}^+$ telle que $f(0) = 0$;

tels que pour tout $1 \leq i \leq K$ et tout $t \in \mathbb{R}^+$,

$$\begin{aligned} \lambda_i(t) &= \lambda_i(0) + f \left(\sum_{j \neq i} \int_0^t h_{j \rightarrow i}(s) \widehat{N}_{j \rightarrow i}(ds) \right) + \int_0^t (g_i(s, \lambda_i(s)) - \lambda_i(s)) N_i(ds) \\ &\quad + \int_0^t (\sigma_i(s, \lambda_i(s)) - \lambda_i(s)) ds. \end{aligned} \tag{11}$$

Nous pouvons alors formaliser ce que nous entendons dans ce cadre par hypothèse poissonnienne:

Définition 0.0.9. Nous disons qu'un PIAFc vérifie l'hypothèse poissonnienne si tous les temps d'interaction sont donnés par des processus de Poisson indépendants. Nous notons avec des tildes tous les processus dans ce régime. Plus précisément, pour $i, j \in 1 \leq i \leq K$ avec $i \neq j$, $\tilde{N}_{j \rightarrow i}$ sont des processus de Poisson indépendants avec intensités $s \rightarrow \mathbf{E}[\tilde{\lambda}_j(s)]$ et pour tout $t \in \mathbb{R}^+$,

$$\begin{aligned} \tilde{\lambda}_i(t) &= \tilde{\lambda}_i(0) + f \left(\sum_{j \neq i} \int_0^t h_{j \rightarrow i}(s) \tilde{N}_{j \rightarrow i}(ds) \right) + \int_0^t (g_i(s, \tilde{\lambda}_i(s)) - \tilde{\lambda}_i(s)) \tilde{N}_i(ds) \\ &\quad + \int_0^t (\sigma_i(s, \tilde{\lambda}_i(s)) - \tilde{\lambda}_i(s)) ds. \end{aligned} \tag{12}$$

Étant donné un PIAFc, nous pouvons alors construire une dynamique à répliques correspondante.

Définition 0.0.10. Soient $K, M \geq 2$. Un PIAFc en champ moyen à M répliques est défini par la donnée de processus ponctuels $(N_{m,i}^M)_{1 \leq i \leq K, 1 \leq m \leq M}$ admettant les intensités stochastiques $(\lambda_{m,i}^M)$ telles que pour tout $t \in \mathbb{R}^+$, il existe

- des processus ponctuels $(\hat{N}_{n,j \rightarrow i})_{1 \leq i \leq K, j \neq i, 1 \leq n \leq M}$ admettant des intensités stochastiques $(\hat{\lambda}_{n,j,i})$ appelés processus d'agrégation;
- des fonctions $(h_{j \rightarrow i})_{1 \leq i \leq K, j \neq i} : \mathbb{R} \mapsto \mathbb{R}$ appelées fonctions d'interaction telles qu'il existe $H > 0$ vérifiant pour tous i, j et tout $t \in \mathbb{R}$, $|h_{j \rightarrow i}(t)| \leq H$;
- des fonctions $(g_i)_{1 \leq i \leq K} : \mathbb{R} \mapsto \mathbb{R}^+$ et $(\sigma_i)_{1 \leq i \leq K} : \mathbb{R} \mapsto \mathbb{R}^+$ appelées fonctions d'évolution autonome;
- une fonction lipschitzienne $f : \mathbb{R} \mapsto \mathbb{R}^+$ telle que $f(0) = 0$;
- des processus de routage $\{V_{(m,i) \rightarrow j}^M(t)\}_{t \in \mathbb{R}}$ prévisibles vis-à-vis de la filtration engendrée par les processus $N_{m,i}^M$ pour $1 \leq m \leq M, 1 \leq i, j \leq K$, tels que pour tout instant d'interaction T , i.e., chaque point de $N_{m,i}^M$, les variables aléatoires $\{V_{(m,i) \rightarrow j}^M(T)\}_j$ sont mutuellement indépendantes, indépendantes du passé, et uniformément distribuées sur $\{1, \dots, M\} \setminus \{m\}$, à valeurs dans

$$\mathcal{V}_{m,i}^M = \{v \in [1, \dots, M]^K \mid v_i = m \text{ and } v_j \neq m, j \neq i\};$$

tels que pour tous $1 \leq m \leq M, 1 \leq i \leq K$ et tout $t \in \mathbb{R}^+$,

$$\begin{aligned} \lambda_{m,i}(t) &= \lambda_{m,i}(0) + f\left(\sum_{j \neq i} \sum_{n \neq m} \int_0^t h_{j \rightarrow i}(s) \mathbb{1}_{\{V_{(n,j) \rightarrow i}^M(s) = m\}} \hat{N}_{n,j \rightarrow i}(ds)\right) \\ &\quad + \int_0^t (g_i(s, \lambda_{m,i}(s)) - \lambda_{m,i}(s)) N_{m,i}(ds) + \int_0^t (\sigma_i(s, \lambda_{m,i}(s)) - \lambda_{m,i}(s)) ds. \end{aligned} \tag{13}$$

Nous supposons toujours que $\hat{N}_{n,j \rightarrow i} = N_{n,j}$ pour tous $1 \leq n \leq M, 1 \leq j \leq K$ par la suite.

Par ailleurs, nous avons besoin des conditions suivantes sur g_i et σ_i :

Hypothèse 0.0.11. Pour tous $s, t \in \mathbb{R}$, pour tout $i \in \{1, \dots, K\}$,

$$g_i(s, t) \leq t \text{ et } \sigma_i(s, t) \leq t.$$

En particulier,

$$g_i(s, \lambda_i(s)) \leq \lambda_i(s) \text{ et } \sigma_i(s, \lambda_i(s)) \leq \lambda_i(s).$$

Cela implique que les processus d'état λ_i sont toujours décroissants entre deux agrégations.

Nous avons enfin besoin de la même contrainte sur les conditions initiales 0.0.6 que précédemment pour établir le théorème suivant :

Théorème 0.0.12. *Soit $T \in \mathbb{R}^+$. Soit $t \in [0, T]$. Soit*

$$A_{m,i}^M(t) = \sum_{j \neq i} \sum_{n \neq m} \int_0^t h_{j \rightarrow i}(s) \mathbb{I}_{\{V_{(n,j) \rightarrow i}^M(s)=m\}} \widehat{N}_{n,j \rightarrow i}^M(ds),$$

avec \widehat{N}^M défini par (3.32), et soit

$$\tilde{A}_i(t) = \sum_{j \neq i} \int_0^t h_{j \rightarrow i}(s) \widehat{N}_{j \rightarrow i}(ds),$$

avec $(\widehat{N}_{j \rightarrow i})_j$ des processus de Poisson indépendants avec intensités $s \mapsto \mathbf{E}[\widehat{\lambda}(s)]$.

Alors,

1. les processus $(\tilde{A}_1, \dots, \tilde{A}_K)$ sont indépendants, tout comme les processus $(\tilde{\lambda}_1, \dots, \tilde{\lambda}_K)$;
2. pour tout $(m, i) \in \{1, \dots, M\} \times \{1, \dots, K\}$, la variable aléatoire $A_{m,i}^M(t)$ converge en variation totale vers $\tilde{A}_i(t)$ quand $M \rightarrow \infty$;
3. pour tout $(m, i) \in \{1, \dots, M\} \times \{1, \dots, K\}$, la variable aléatoire $\lambda_{m,i}^M(t)$ définie par (3.32) converge en variation totale vers $\tilde{\lambda}_i(t)$ défini par (3.31) quand $M \rightarrow \infty$;
4. soit \mathcal{N} un sous-ensemble fini de \mathbb{N}^* , pour tout $i \in \{1, \dots, K\}$, les processus $(A_{m,i}^M(\cdot))_{m \in \mathcal{N}}$ et $(\lambda_{m,i}^M(\cdot))_{m \in \mathcal{N}}$ convergent faiblement dans l'espace de Skorokhod $D([0, T])^{\text{card}(\mathcal{N})}$ muni de la topologie de Skorokhod produit vers $\text{card}(\mathcal{N})$ copies indépendantes des processus limites respectifs $(\tilde{A}_i(\cdot))$ et $(\tilde{\lambda}_i(\cdot))$ when $M \rightarrow \infty$.

La structure de la preuve est similaire à celle du cas particulier du modèle de Galves-Löcherbach en temps continu.

Quelques exemples de processus qui sont dans la classe des PIAFc sont les suivants :

- en prenant pour tous $1 \leq i, j \leq K$ et tout $t \in \mathbb{R}$,
 $h_{j \rightarrow i}(t) = \mu_{j \rightarrow i} \geq 0$, $f(t) = |t|$, $g_i(t, \lambda_i(t)) = r_i > 0$, $\sigma_i(t, \lambda_i(t)) = b_i > 0$, nous retrouvons le modèle de Galves-Löcherbach en temps continu introduit plus haut [35] [8].

- en prenant pour tous $1 \leq i, j \leq K$ et tout $t \geq 0$,
 $h_{j \rightarrow i}(t) = \mu_{j \rightarrow i} \in \mathbb{R}$, $f(t) = \max(0, |t|)$, $g_i(t, \lambda_i(t)) = r_i > 0$, $\sigma_i(t, \lambda_i(t)) = b_i > 0$, on obtient un modèle de Galves-Löcherbach généralisé incorporant de l'inhibition.
- en prenant pour tous $1 \leq i, j \leq K$ et tout $t \geq 0$,
 $h_{j \rightarrow i}(t) = \mathbb{1}_{\{j=i+1 \bmod K\}}$, $f(t) = |t|$, $g_i(t, \lambda_i(t)) = \lambda_i(t) - 1$, $\sigma_i(t) = \lambda_i(t)$, on obtient un modèle de réseau de files d'attentes de concaténation en temps continu. Notons qu'un tel réseau sous hypothèse poissonnienne est un réseau de Gordon-Newell [50].

Il est possible d'élargir un peu ce cadre afin de considérer des processus admettant plusieurs types d'arrivées, tant qu'elles sont indépendantes entre elles. Cela permet de considérer des versions de processus de migration-contagion sur des réseaux et d'en étudier la limite de champ moyen à répliques. Les processus de migration-contagion ont été introduits par Baccelli, Foss et Shneer dans [6] pour modéliser la propagation d'épidémies en combinant des modèles de file d'attente et un modèle classique d'épidémiologie, le modèle SIS. Plus précisément, dans un processus de migration-contagion, les individus se déplacent entre des stations appelées réacteurs quand on les considère isolées. Chaque individu est marqué de son état, soit I (infecté), soit S (susceptible). A l'intérieur de chaque réacteur, chaque individu change d'état suivant des mécanismes qui diffèrent en fonction du processus de migration-contagion considéré. Par exemple, dans le réacteur SIS, nous considérons une file d'attente $M/M/\infty$ avec taux d'arrivée λ et taux de service μ . En notant $X(t)$ (resp. $Y(t)$) le nombre d'individus susceptibles (resp. infectés) à l'instant t . A l'intérieur de la file d'attente, un individu susceptible devient infecté à taux $\alpha Y(t)$ et un individu infecté devient susceptible à taux β .

Nous considérons une variante du réacteur SIS : le réacteur SIS-DOCS, pour "Departure On Change of State". Dans celui-ci, lorsqu'un individu change d'état au sein d'un réacteur, il quitte immédiatement celui-ci. Dans l'objectif d'appliquer la méthodologie des champs moyens à répliques à ce type de dynamiques, nous considérons un réseau de tels réacteurs : chaque noeud du réseau représente un réacteur SIS-DOCS, et lorsqu'il y a un départ d'un noeud, l'individu choisit un noeud voisin et se rend à celui-ci. En première approximation, on peut modéliser chaque réacteur comme une ville et le réseau entier comme un pays. Pour étudier l'impact de la géométrie du réseau sur la propagation de l'épidémie dans celui-ci, nous introduisons un modèle à M répliques de la dynamique et nous considérons la limite de celui-ci lorsque M tend vers l'infini.

Les processus de contagion-migration ne rentrent pas dans la classe des PIAFc en raison du fait que les individus choisissent aléatoirement le réacteur où ils se rendent au moment du départ, et la technique de preuve du champ moyen

à répliques ne s'applique pas directement à ce modèle. Aussi, la convergence du modèle à répliques vers une dynamique sous hypothèse poissonnienne demeure à l'état de conjecture. Dans le chapitre 4, nous étudions l'état stationnaire du modèle SIS-DOCS en réseau sous hypothèse poissonnienne, en admettant la conjecture sus-mentionnée. Plus précisément, dans l'objectif d'une étude numérique pour mettre en évidence des transitions de phase, nous montrons que les nombres moyens d'individus susceptibles et infectés dans chaque ville vérifient le système d'équations non-linéaires suivant:

$$\left\{ \begin{array}{l} \mathbf{E}[\tilde{X}_i] = \frac{(\sum_{j \neq i} \mu_j p_{j \rightarrow i} \mathbf{E}[\tilde{X}_j] + \sum_{j \neq i} \beta_j p_{j \rightarrow i} \mathbf{E}[\tilde{Y}_j])}{\mu_i + \beta_i + \alpha_i} \\ \int_0^1 e^{-\frac{(\mu_i + \beta_i) \mathbf{E}[\tilde{Y}_i] \alpha_i^2 (1-t)}{(\mu_i + \beta_i)(\mu_i + \beta_i + \alpha_i)^2} t} \left(\frac{\mu_i}{\mu_i + \beta_i + \alpha_i} + \frac{(\mu_i + \beta_i) \mathbf{E}[\tilde{Y}_i] \alpha_i}{(\mu_i + \beta_i + \alpha_i)^2} - 1 \right) dt \\ \mathbf{E}[\tilde{Y}_i] = \frac{\sum_{j \neq i} \mu_j p_{j \rightarrow i} \mathbf{E}[\tilde{Y}_j] + \sum_{j \neq i} p_{j \rightarrow i} (\sum_{l \neq j} \mu_l p_{l \rightarrow j} \mathbf{E}[\tilde{X}_l] + \sum_{l \neq j} \beta_l p_{l \rightarrow j} \mathbf{E}[\tilde{Y}_l] - \mu_j \mathbf{E}[\tilde{X}_j])}{\mu_i + \beta_i} \\ \sum_{i=1}^K \mathbf{E}[\tilde{X}_i] + \mathbf{E}[\tilde{Y}_i] = N. \end{array} \right. \quad (14)$$

Chapter 1

Introduction

The goal of this chapter is twofold. First, we present the general motivation behind this work and attempt to convince the reader (if need be) of the interest of this study and the questions behind it, before discussing the organisation of the manuscript. The second part of the chapter is devoted to brief presentations of the theoretical concepts that either are used to answer these questions or help to re-contextualize the answers that we bring. Most of the latter can be skipped by a reader already familiar with the general subject. We bring however attention to Section 1.6, which gives some general ideas behind the replica-mean-field approach, a subject at the heart of this work that is less widely known than the other topics presented in this section.

1.1 Motivation

There is very often a gap between trying to provide a description in the mathematical language of physical phenomena and being able to prove meaningful theorems for these mathematical models. This is due to the complexity of most phenomena, making it so that mathematical descriptions aiming to capture their features quickly become too complex for mathematical analysis. This work falls into a long tradition in applied mathematics of trying to fill that gap, by trying to propose and study models that capture key features of the phenomena while also being computationally tractable to some extent.

We are interested here in studying phenomena that can be described equivalently as interacting particle systems with fixed positions, agent-based models or network dynamics, with inherent variability in between observations. The key characteristic here is to have a microscopic description of the phenomenon with an underlying geometry (the particles or agents interact only with their neighbors, nodes in the network communicate only if there is an edge between them) and

punctuate interactions. This restriction enables us to describe these phenomena through the times at which interactions in between particles/agents/nodes happen and to see these times as realizations of random times, which are an instance of point processes, a class of stochastic processes for which a rich and powerful theory has been developed.

Given this description, it is then possible to describe the behavior of the system through stochastic differential equations, with the unknowns being stochastic processes associated with the point processes attached to each node/particle/agent giving the interaction times. This system of equations, as might be expected of a model trying to provide a microscopic description of what "really" happens at each time in the system, usually isn't tractable. As previously mentioned, a rich area of research is concerned with ways to derive simpler models from these "exact" descriptions while preserving certain characteristics. Among the characteristics that are almost always put on the chopping block to obtain tractability are correlations between particles/agents/nodes (that is, they have to be assumed to be independent) and heterogeneity in the underlying network (that is, it has to be assumed to be fully connected with identical weights on the edges). Moreover, it is usually necessary to consider some scaling of the system (letting the number of nodes go to infinity, for example) in order to reduce dimensions.

In this work, we focus mostly on phenomena from life sciences, which is an arbitrary limitation due to the interests of the author. In particular, the original motivation stems from so-called intensity-based models from computational neuroscience. The goal is to describe the evolution in time of the electrical potentials of a population of neurons seen as nodes in a network. In [8], Baccelli and Taillefumier have devised a new physical system that, when properly scaled, allowed to obtain closed forms for physical quantities such as the mean spiking rates of the neurons at stationarity. This physical system, called the replica-mean-field, consists of randomly interacting copies of the original network, and is at the core of the focus of this work.

Baccelli and Taillefumier studied the scaling limit of an infinite number of replicas, and have shown that equations associated with it become tractable. However, this study was done at a cost, namely the Poisson Hypothesis. A popular simplification coming from the field of telecommunications, a physical system satisfying the Poisson Hypothesis means that arrivals to a given node inside the network of interest are Poisson-distributed and that nodes are independent. This often allows for tractability while, notably, keeping intact the geometry of the network. As was the case in their work, this hypothesis is often only conjectured or numerically validated. As such, one motivation was to prove that the hypothesized convergence does indeed take place.

This was first done in a simpler, discrete-time setting where all interactions

are synchronous, and is presented in Chapter 2. The proof relies on the analysis of the one-step transition of Markov dynamics for the replica-mean-field version of the network of interest. Namely, we show that certain asymptotic properties are preserved by the dynamics, allowing for the Poisson Hypothesis to hold at any finite time under certain assumptions on the initial conditions. Moreover, while the proof is first done for a discrete-time model from computation neuroscience, we introduce a class of processes we call fragmentation-interaction-aggregation processes, for which the proof can be generalized in the goal of widening the scope of phenomena for which it is possible to rigorously prove the Poisson Hypothesis at the infinite replica replica-mean-field limit.

We then proceed in Chapter 3 to perform an analogous study in continuous time. The asynchronous nature of interactions in this setting makes the generalization of the discrete time results nontrivial. We prove in the first half of the chapter the original conjecture of Baccelli and Taillefumier for the replica-mean-field limit of intensity-based neural dynamics, both in transient and stationary regimes (their original study focused on the stationary case). The proof relies on the generalization of a powerful Poisson approximation method, called the Chen-Stein method, to the case of a random amount of Bernoulli random variables to obtain explicit bounds between the conjectured limit dynamics and the finite-replica-mean-field dynamics, when coupling them on their source of randomness. The bounds are then shown to be asymptotically arbitrarily small using fixed-point techniques. The second half of Chapter 3 is aimed at generalizing as much as possible the class of dynamics to which the proof of the replica-mean-field limit applies. By analogy with discrete time, we introduce a class of so-called continuous fragmentation-interaction-aggregation processes for which the same result applies. We justify the name by showing the existence of links between discrete-time and continuous-time versions of fragmentation-interaction-aggregation processes.

Finally, in Chapter 4, we apply the convergence result to study the scaling limit of a replica-mean-field version of migration-contagion processes incorporating a spatial component. These processes have recently been introduced by Baccelli, Foss and Shneer in [6] to study the propagation of epidemics, combining compartmental epidemiological models and queueing theory. Here, we show that the above-mentioned scaling limit does indeed exist, leveraging the results from Chapter 3. Under the Poisson Hypothesis, we derive a system of equations for the mean numbers of infected and susceptible individuals at the nodes of an underlying network, and study this system numerically.

1.2 Primer on point process dynamics

In this section, we will give an extremely brief and incomplete overview of some elements of point process theory. As previously mentioned, our motivation comes from phenomena with punctuate events (spikings of neurons in the brain, first and foremost, but also individual-based epidemiology models). Due to the variability in between observations of the phenomena, a natural approach is to model these phenomena as realizations of *point processes*, that is, of random sets of points. In all our applications, these points represent the times at which the events of interest happen and can be seen as random times.

General point process theory

We will quickly give a few general notions about point processes.

Definition 1.2.1. *Given a probability space $(\Omega, \mathcal{F}, \mathbf{P})$, a Polish space (E, d) and \mathbb{K} a so-called configuration space, let $\mathcal{K} = \sigma(\pi_A, A \in \mathcal{B}(E))$, where if $x \in \mathbb{K}$, $\pi_A(x) = \text{card}(A \cap x)$. A simple point process is a measurable map $(\Omega, \mathcal{F}, \mathbf{P}) \mapsto (\mathbb{K}, \mathcal{K})$.*

It can be shown that an equivalent definition would be to consider a point process to be a counting measure Φ such that

$$\Phi = \sum_{x \in \mathcal{X}} \delta_x,$$

where $\mathcal{X} \in \mathbb{K}$. When represented in this fashion, $\{\Phi(A), A \in \mathcal{B}(E)\}$ is a random process on $\mathcal{B}(E)$. Denote $m(A) = \mathbf{E}[\Phi(A)]$. Then it can be shown that $A \rightarrow m(A)$ is a measure, called the *intensity measure* of the point process Φ . Another object associated to a point process is its *void probability*: namely for $A \in \mathcal{B}(E)$, the void probability of A is $\mathbf{P}(N(A) = 0)$. If N is simple, Renyi's theorem states that the distribution of N is characterized by the family of its void probabilities $\{\mathbf{P}(N(A) = 0)\}$ for all closed $A \in \mathcal{B}(E)$.

We will use the following notation throughout the text:

Notation 1.2.2. *If N is a point process on E and f a function, we write for any $A \in \mathcal{B}(E)$*

$$\sum_{x \in N \cap A} f(x) = \int_A f(t) N(dt).$$

The fundamental simple point process is called the *Poisson point process*, and it is characterized by its intensity measure. We define it here in \mathbb{R}^d , $d \geq 1$ but the construction can be extended to any Polish space.

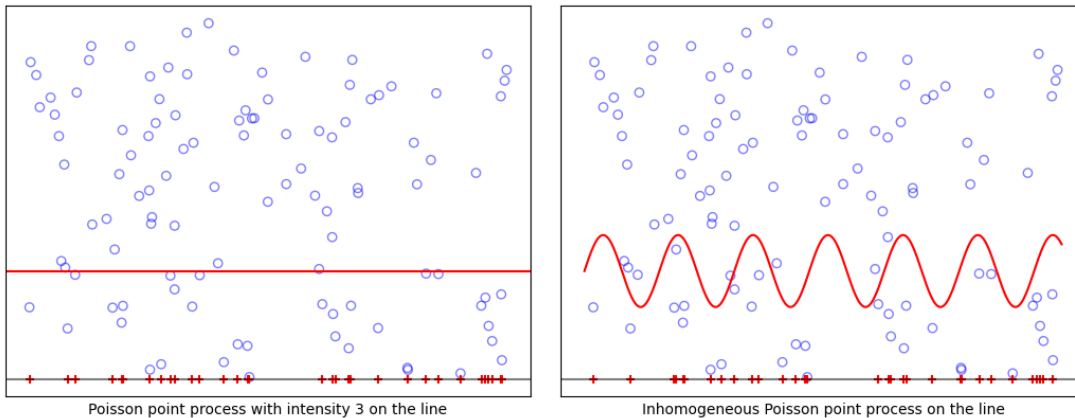
Definition 1.2.3. Let $d \geq 1$. A point process N is said to be a Poisson point process with intensity measure m on \mathbb{R}^d if

1. For all $A \in \mathcal{B}(\mathbb{R}^d)$ such that $m(A) < \infty$, $N(A)$ follows a Poisson distribution with parameter $m(A)$;
2. If A_1, \dots, A_p are disjoint elements of $\mathcal{B}(\mathbb{R}^d)$, then $N(A_1) \dots N(A_p)$ are independent.

It is not obvious that such a process exists, but it is possible to prove that if m is a locally finite measure (i.e. non-atomic), then a Poisson point process with intensity measure m exists. If m is proportional to the Lebesgue measure on \mathbb{R}^d , the Poisson point process is said to be *homogeneous*. The more general case we have presented above is often referred to as the *inhomogeneous* Poisson point process. We now further restrain ourselves to the one-dimensional case. This allows us to state the following embedding theorem:

Proposition 1.2.1. Let \bar{N} be a unit-intensity homogeneous Poisson point process on \mathbb{R}^2 . For a given non-negative function λ on \mathbb{R} such that $\int_{\mathbb{R}} \lambda(x) dx < \infty$, consider the point process N_λ on \mathbb{R} consisting of the projections of the atoms of \bar{N} which are below the curve of λ on the x -axis. Then N_λ is a inhomogeneous Poisson point process on \mathbb{R} with intensity measure $\Lambda(dx) = \lambda(x) dx$.

This stems from the fact that being a Poisson point process is preserved by random transformations (i.e. random independent displacements) of points.



One last theoretical result presented in this section concerns convergence of point processes. We abusively say point processes weakly converge when their distributions do. We recall the following result from Kallenberg [45]:

Proposition 1.2.2. Let N, N_1, N_2, \dots be point processes on \mathbb{R}^d , where N is simple. Then N_n weakly converges to N when $n \rightarrow \infty$ iff for all bounded $A \in \mathcal{B}(\mathbb{R}^d)$, $\mathbf{P}(N_n(A) = 0) \rightarrow \mathbf{P}(N(A) = 0)$ when $n \rightarrow \infty$.

In other words, weak convergence of point processes is equivalent to convergence of their void probabilities.

Stochastic intensities

The seemingly innocuous Proposition 1.2.1 gives insight into a natural extension of Poisson point processes: what happens if the function λ in the theorem above is allowed to be a predictable process? The answer consists in the class of point processes admitting a *stochastic intensity*, that we will now define before linking the definition to the above theorem.

Definition 1.2.4. Let \mathcal{F}_t be a filtration on \mathbb{R} . Let N be a point process on \mathbb{R} . We say that N admits the \mathcal{F}_t -stochastic intensity λ if λ is a \mathcal{F}_t -predictable random process satisfying for all $s < t \in \mathbb{R}$:

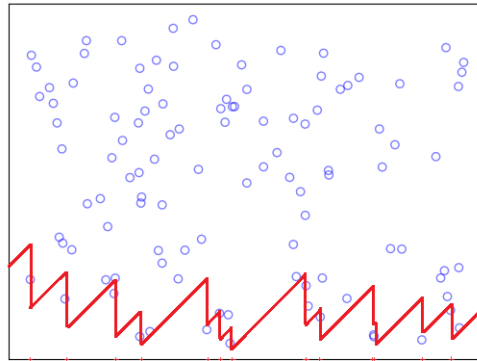
$$\mathbf{E} [N(s, t) | \mathcal{F}_s] = \mathbf{E} \left[\int_s^t \lambda(u) \, du | \mathcal{F}_s \right]. \quad (1.1)$$

Typically, \mathcal{F}_t is taken to be the history of the process N up to time t . We will hereafter refer to (1.1) as the stochastic intensity property.

Point processes with stochastic intensities can be thought of as generalized Poisson point processes, as the following embedding result makes clear:

Proposition 1.2.3. Let N be a point process on \mathbb{R} . Let (\mathcal{F}_t) be an internal history of N . Suppose N admits a (\mathcal{F}_t) -stochastic intensity $\{\lambda(t)\}_{t \in \mathbb{R}}$. Then there exists a Poisson point process \bar{N} with intensity 1 on \mathbb{R}^2 such that, for all $C \in \mathcal{B}(\mathbb{R})$,

$$N(C) = \int_{C \times \mathbb{R}} \mathbb{I}_{[0, \lambda(s)]}(u) \bar{N}(ds \times du).$$



Note that the reverse is also true: given a process constructed by projection from a Poisson point process \bar{N} with intensity 1 on \mathbb{R}^2 in the manner described above, the resulting point process on the real line will admit a stochastic intensity.

Hawkes processes

An important example of point processes admitting stochastic intensities are Hawkes processes. Also referred to as self-exciting processes, they model the situation where the presence of a point in the point process depends on its points in the past. Formally, we have the following definition:

Definition 1.2.5. *We say that a point process N admitting a stochastic intensity is a Hawkes process if there exists a function $\Phi : \mathbb{R} \rightarrow \mathbb{R}^+$, $\lambda > 0$ and a signed measurable function $g : \mathbb{R}^{+*} \rightarrow \mathbb{R}$ such that its intensity Λ_g verifies for all $t \in \mathbb{R}$*

$$\Lambda_g(t) = \Phi\left(\lambda + \int_{(-\infty, t)} g(t-u)N(du)\right).$$

Originally introduced to study earthquakes, Hawkes processes have been extensively used in many applications, including but not limited to neural computations, finance and insurance.

It is also possible to construct Hawkes point processes in the multidimensional setting, to describe a system of particles for example.

Definition 1.2.6. *We say that a finite family of point processes N_1, \dots, N_K admitting stochastic intensities form a multivariate Hawkes process if there exist functions $\Phi_1, \dots, \Phi_K : \mathbb{R} \rightarrow \mathbb{R}^+$ and signed measurable functions $(g_{j \rightarrow i})_{1 \leq i \leq K} : \mathbb{R}^{+*} \rightarrow \mathbb{R}$ such that their respective intensities Λ_i verify for all $t \in \mathbb{R}$ and all $1 \leq i \leq K$,*

$$\Lambda_i(t) = \Phi_i\left(\sum_{j=1}^K \int_{(-\infty, t)} g_{j \rightarrow i}(t-u)N_j(du)\right).$$

The asymptotics of Hawkes processes have been extensively studied, see [47] or [20] for laws of large numbers and large deviation bounds for extended Hawkes processes.

1.3 Primer on neural intensity-based models

Modeling neural computations

As a considerable part of our work concerned applications to neural dynamics, the goal of this section is to provide a partial introduction to a few relevant models and concepts from computational neuroscience.

It is generally understood that computations in the brain are the result of electrical signals being transmitted by neurons to one another. A common modeling choice is to focus on the electrical activity of a neuron, neglecting the mechanical

aspects. The focus is then put on the membrane potential of the neuron and on the differences in concentration of different ions present inside and outside the cell. Various deterministic models to describe the behavior of the neuron through the change in concentration of these ions have been proposed, such as the Morris-Lecar model [62] or the Hodgkin and Huxley model [41]. The basic idea behind these models is to consider an electrical circuit "equivalent" to the neuron and study the current balance equation. The basic behavior of a neuron is usually described as follows: its membrane potential starts at some resting potential, it sharply increases when aggregating signals from other neurons and *spikes*, that is, resets to a value close to its resting potential, when it gets high enough. After a spike, there is usually a short refraction period during which the potential stays below the resting potential and the neuron does not spike.

These models suffer however from several problems. First, they depend on several unknowns resulting in complex systems of ODEs that are hard to study. The Hodgkin-Huxley model, for instance, is a four-dimensional one. The excitability phenomenon that gives rise to the spiking mechanism can be reduced to two dimensions: this is the FitzHugh-Nagumo model [33] [63]. However, this mechanism cannot be reduced to one dimension [42].

Second, experimental data shows variability in the sequences of spiking times of a given neuron, something that these models, being deterministic, cannot incorporate.

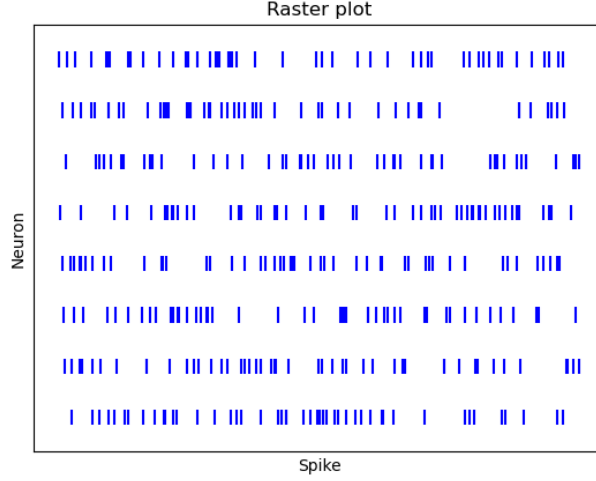
To solve these two issues, a popular class of models is the Leaky Integrate and Fire (LIF) models. The basic idea behind these goes as follows: there exists some fixed threshold and the neuron aggregates arrivals from other neurons until reaching the threshold, at which point it spikes and resets back to its resting value. The aggregations from other neurons are approximated by a diffusion, such as a Brownian motion or a Ornstein-Uhlenbeck process, for which the law of the hitting time of the threshold is known. We refer to [37] for more details on LIF models.

However, as it is a one-dimensional model, the previous observation means that considering the threshold to be fixed is unrealistic. Other models have thus been developed with a stochastic threshold, with the Galves-Löcherbach model central among them. This, however, is done at the expense of the "real" state of the system. We will devote the rest of this section to the presentation of this class of models.

The Galves-Löcherbach model

Consider the case of a population of neurons. The spiking times of the neurons composing it can be represented on a so-called raster plot, seen below, with each line corresponding to a neuron's spiking times. Now, consider these spiking times as realizations of a point process. If we consider these point processes as admitting

stochastic intensities, these can represent the membrane potentials of the neurons. The behavior of the system is then described by a system of stochastic differential equations verified by the stochastic intensities of the neurons.



In the Galves-Löcherbach model, the point processes of spiking times are taken to be generalizations of Hawkes processes, in the sense that instead of keeping track of all the history of the process, only the dynamics after the last spiking time are integrated.

Historically, GL models have been first introduced in discrete time as Markov chains with variable memory length, based on ideas of Rissanen [69]. We give here a formal definition in discrete time:

Definition 1.3.1. *Consider a finite or countable set of neurons I . Consider time to be discrete, with a given time length step. For any time t , let $X_t = (X_t(i), i \in I)$, where $X_t(i) = 1$ if neuron i spikes at time t , and 0 otherwise. Let $(\mu_{j \rightarrow i})_{i \neq j \in I}$ be real numbers representing the synaptic weights of the system, i.e. the amount the potential of neuron i is incremented by when neuron j spikes. Let $L_t(i)$ be the last spiking time of neuron i before time t . Then the membrane potential of neuron i at time t is given by*

$$V_t(i) = \sum_{j \in I, j \neq i} \mu_{j \rightarrow i} \sum_{s=L_t(i)}^{t-1} g_j(t-s)X_s(j).$$

In this setting, the probability of having a spike at neuron i at time t is given by

$$\mathbf{P}(X_t(i) = 1 | \mathcal{F}_{t-1}) = \Phi_i \left(\sum_j \mu_{j \rightarrow i} \sum_{s=L_t(i)}^{t-1} g_j(t-s)X_s(j), t - L_t(i) \right),$$

where $\Phi_i : \mathbb{R} \times \mathbb{N} \rightarrow [0, 1]$ and $g_j : \mathbb{N} \rightarrow \mathbb{R}^+$ are measurable functions and Φ is assumed to be uniformly Lipschitz continuous.

From this representation, it is clear that this model is reminiscent of a discrete version of a Hawkes process, and it is naturally generalized in the continuous time setting in the following fashion, which is heavily linked to the multivariate Hawkes process, except that we do not integrate on the entire past:

Definition 1.3.2. *Let $K \geq 2$. The point processes N_1, \dots, N_K be point processes are said to follow the continuous-time Galves-Löcherbach model if they admit stochastic intensities $\lambda_1, \dots, \lambda_K$ such that for all $t \in \mathbb{R}$, there exist functions Φ_i and g_i such that*

$$\lambda_i(t) = \Phi_i\left(\sum_j \mu_{j \rightarrow i} \int_{L_t(i)}^t g_i(t-s) N_j(ds), t - L_t(i)\right),$$

where $L_t(i)$ is once again the last spiking time of neuron i before t .

One common choice for the functions Φ_i and g_i are given by $\Phi_i(x, s) = x + b_i + (r_i - b_i)e^{-\frac{s}{\tau_i}}$ and $g_i(t-s) = e^{-\frac{t-s}{\tau_i}}$. This is sometimes referred to as the linear Galves-Löcherbach model [8] due to the linearity of Φ_i in x . Note also that the coefficients $\mu_{j \rightarrow i}$ are also often denoted $\mu_{i,j}$, unlike many other similar fields where the destination index is almost always located after the origin index, which, as the author has witnessed, has led to frustration among researchers both inside and outside the field. In an effort to be part of the solution, rather than part of the problem, we will exclusively use the notation $\mu_{j \rightarrow i}$ hereafter.

The Galves-Löcherbach model have been primarily studied in the infinite population limit [35], notably to construct perfect simulation algorithms [67]. In the next section, we describe common techniques to derive these scaling limits.

1.4 Primer on mean-field theory

The mean-field framework

As we have previously seen, intensity-based models such as Galves-Löcherbach models aim to describe the spiking activity of a population of neurons. As such, they are a particular instance of what are often called agent-based models, in which a given phenomenon is described through a particle system, with differential equations governing the behavior of each particle. We will interchangeably see agents as particles in a system with fixed positions or nodes on an underlying graph of interactions, with edges signifying the possibility of interactions and weights on the edges denoting the physical quantities transmitted in the event of interaction.

As in the case in the Galves-Löcherbach model, neither the system of equations nor associated functionals such as moment generating functions allow for closed forms except in some very particular cases. As such, one very general approach is to simplify the model, usually by scaling it in a certain way, to obtain tractability. Both replica-mean-fields, which we will extensively present later on, and fluid limits [53] [64], in which both time and particle quantity are scaled together, fall into this category of approaches. In this section, we present the most common and simple approach of this type, namely the mean-field regime, also referred to as the thermodynamic mean-field in certain applications.

The basic idea behind the mean-field approach goes as follows: to simplify the dynamics, consider the limit behavior of the particle system of interest when the number of particles goes to infinity. Of course, to prevent explosion in finite time of the system, the updates to the state of a given particle due to the interactions with its neighbors must be scaled accordingly. The most common scaling involves approximating these interactions as the empirical mean of all the interactions between particles in the network. Heuristically, as the network size grows, the particles become independent and a law of large numbers-type result shows that the interactions received by a single particle converges to some deterministic quantity, the theoretical mean. This asymptotic independence between particles is referred to as *propagation of chaos*, a terminology originating from Kac [44]. It is closely linked to the convergence of the empirical measure of the particles, see [75].

In this framework, all the particles satisfy the same nonlinear PDE at the limit, where the nonlinearity comes from the fact that for a given particle, the aggregation of interactions is replaced by a deterministic term that depends on the law of the particle. These nonlinear PDEs are known as *McKean-Vlasov* equations [60].

One of the main questions in mean-field theory concerns proving propagation of chaos in various models. This involves proving some kind of convergence of the particle system to infinite copies of the limit dynamics. There are two main methods for proving such results. The first involves working in the weak convergence framework: to prove weak convergence, first prove tightness for the sequence of empirical measures. Then, identify a limit equation and prove uniqueness of solutions to it. The second method, popularized by Sznitman [75], is based on a coupling technique. The idea is to couple together i.i.d. copies of the supposed limit equation (established heuristically) and the state equations of the particle system in a clever way that allows to compare the trajectories in some suitable metric, such as the Wasserstein distance. Of course, this requires rigorously establishing first that the limit equation is well-defined. The second approach often has the benefit of providing explicit convergence rates. We will use elements of both techniques in our work.

A toy model with diffusions

In what follows, we will present an example of application of Sznitman's coupling method to a toy model with diffusions, inspired by both his original work [75] and notes by Chevallier [25]. Our goal here is not to go into much detail, but rather highlight a few key points in the approach that we will later adapt to a novel setting.

Consider the following toy particle system:

Let $K \geq 1$. We consider the processes $(X_i^K)_{1 \leq i \leq K}$ satisfying the mean field equations

$$X_i^K(t) = X_i(0) + \omega_i(t) + \int_0^t \frac{1}{K} \sum_{j=1}^K b(X_i^K(s), X_j^K(s)) ds$$

for all $1 \leq i \leq K$, where (ω_i) are i.i.d. Brownian motions and b is a globally Lipschitz function.

Here, the interactions are clearly of mean-field type but, unlike previously encountered models, the stochasticity in the system is given by Brownian motions.

Heuristically, the limit equation that should be verified for all $i \in \mathbb{N}$ is as follows:

$$\bar{X}_i(t) = X_i(0) + \omega_i(t) + \int_0^t \int_{\mathbb{R}^d} b(\bar{X}_i(s), y) \mu_s^i(dy) ds,$$

where ω_i are i.i.d. Brownian motions and for $s \geq 0$, μ_s^i is the law of $\bar{X}^i(s)$.

To compare the two trajectories, a natural idea, which is the basis of Sznitman's approach, is to couple the two dynamics through their sources of randomness, in this case, the initial conditions and their Brownian motions. After thus constructing the two sets of processes on the same probability space, this allows the use of Grönwall-lemma-type arguments to obtain the following bound :

for $i \geq 1$ and $T > 0$,

$$\sup_{K \geq 1} \sqrt{K} \mathbf{E} \left[\sup_{s \in [0, T]} |X_i^K(s) - \bar{X}_i(s)| \right] < \infty.$$

However, up to now, nothing has been said about the existence and uniqueness of a solution to the limit equation, due to the presence of the nonlinear term. To account for it, Sznitman's approach involves looking at the hypothetical solution as the unique fixed point of a map on the space of probability distributions on the space of trajectories, which is in this case the space C of continuous functions from $[0, T]$ to \mathbb{R}^d . Namely, given the following map:

$$\begin{aligned} \Phi: \mathcal{P}(C) &\rightarrow \mathcal{P}(C) \\ m &\mapsto \Phi(m), \end{aligned}$$

where $\Phi(m)$ is the law of the solution of

$$Z(t) = X_1(0) + \omega_1(t) + \int_0^t \int_{\mathbb{R}^d} b(Z(s), y) m(dy) ds,$$

it is clear that the solution of the limit equation is a fixed point of Φ . Thus, to prove that it admits a unique solution, it is sufficient to show that Φ admits a unique fixed point.

To do so, it is sufficient to show that the sequence of iterates of Φ is a Cauchy sequence for a metric that makes $\mathcal{P}(C)$ complete and separable and apply a Banach theorem. The Kantorovitch-1 metric, more commonly known due to an application of Stigler's law of eponymy [74] as the Wasserstein-1 metric [46], is a well-adapted metric to answer this question positively. We recall here its definition:

Definition 1.4.1. For $\mu, \nu \in \mathcal{P}(C)$, where C is the space of continuous functions from $[0, T]$ to \mathbb{R}^d , let

$$K_T(\mu, \nu) = \inf_{\Pi \in C \times C} \mathbf{E}[d_C(x, y)],$$

where Π is a coupling s.t. $x \stackrel{\mathcal{L}}{=} \mu$ and $y \stackrel{\mathcal{L}}{=} \nu$ and d_C is the uniform distance on C .

We will hereafter omit the "1" and simply refer to it as the Kantorovitch metric as we do not consider power distances in this work.

In this metric, one can show the following contraction-type result, from which it easily follows that the sequence of iterates is a Cauchy sequence:

Proposition 1.4.1. Let $\mu, \nu \in \mathcal{P}(C)$. For all $t \in [0, T]$, there exists a positive constant C_T such that

$$K_T(\Phi(\mu), \Phi(\nu)) \leq C_T \int_0^t K_s(\mu, \nu) ds.$$

Thus, the fixed-point approach allows to establish the existence and uniqueness of the solution to the limit equation. We will later adapt these techniques to càdlàg trajectories.

Generalizations, refinements and beyond

In the rest of this section, we discuss some limitations of the mean-field approach and present a small selection of works from the literature that attempt to provide answers to these limitations.

In classical mean-field models such as [34] or [56], the network considered must be assumed fully connected, the effect of interactions on the state of a given agent must be small, typically inversely proportional to the number of agents, in order

to prevent explosion in finite time in the system. These assumptions represent significant constraints on the architectures and sizes of the networks and thus on the types of phenomena for which a mean-field approximation is relevant. The term replacing interactions at the limit is also deterministic, which isn't realistic in certain applications, e.g. neuroscience.

To circumvent these limitations, different approaches have been explored in recent years. To apply mean-field approximation to small-sized networks (with less than 100 agents, for example), the refined-mean-field approach [36][3] adds a correction term to the macroscopic ODE. Mean-field models have also been studied in other scalings, for example diffusive, where the effect of interactions on a particle is inversely proportional to the square root of the number of agents in the system. Conditional propagation of chaos properties have been proven in that setting [32]. This allows to retain stochasticity at the limit, as the noise from the particle system is coalesced into a common Brownian motion.

To incorporate heterogeneity, the properties of graphons (large dense graphs) have been used to derive new limit equations [1]. In this setting, the limit object is an infinite system of ODEs. However, this approach is only valid for dense networks; when the average degree of a node is of order lesser than the amount of nodes in the network, as is the case for example in the human brain, graphon theory does not apply. To account for that, some authors have introduced extended graphons [43] and graphops [52] and proven propagation of chaos in these extended settings. Another approach that retains heterogeneity involves dividing the nodes into multiple classes, considering all the nodes within one class to be fully connected with one another and taking the mean-field limit inside each class [30]. Finally, it is possible to consider the weights of interactions between nodes as realizations of certain i.i.d. random variables, such as Bernoulli random variables [38], and prove propagation of chaos in this setting.

Another approach to incorporate heterogeneity circumvents mean-field models altogether, relying instead on conditional independence properties and local weak limits to obtain local convergence, thus preserving the local topology of the network [55].

1.5 Primer on Poisson approximation

Although the content of this section might seem at first glance unrelated to the topic presented in the previous section, that is, incorporating heterogeneity in mean-field-type models, we shall see that one approach to tackle the question involves replacing interaction times by Poisson point processes. As such, the question of propagation of chaos in that setting becomes related to matters of Poisson approximation, thus justifying the presence of this section. We will mostly focus

on the case of Poisson random variables, but we will mention a result for Poisson point processes as well.

Base results and coupling approaches

Even heuristically, the Poisson distribution is best understood as the limit of the sum of independent rare events. Thus, the question of Poisson approximation is a very natural one, and one that has been extensively studied in the twentieth century. We refer to the very comprehensive book by Barbour [10], and aim here to only present a few fundamental results. Our goal is not to provide optimal rates of convergence, but rather present a few different approaches to Poisson approximation.

First, we cannot present Poisson approximation without stating the following fundamental undergraduate-level result, which makes rigorous the heuristic mentioned above:

Proposition 1.5.1. *For $n > 0$, let $(A_{n,j})_{\{1 \leq j \leq M_n\}}$ be a finite family of independent events. Let $p_{n,j} = \mathbf{P}(A_{n,j})$ and $S_n = \sum_{j=1}^{M_n} \mathbb{1}_{A_{n,j}}$. Suppose $\max_{1 \leq j \leq M_n} p_{n,j} \rightarrow 0$ when $n \rightarrow \infty$ and $\sum_{j=1}^{M_n} p_{n,j} \rightarrow \lambda$ when $n \rightarrow \infty$. Then $(S_n)_{n>0}$ converges in distribution to the law of the Poisson distribution with parameter λ when $n \rightarrow \infty$.*

The proof of this result relies on Levy's theorem. In fact, a more powerful result can be obtained using coupling techniques in a different convergence metric, namely total variation convergence, which we briefly recall here.

Definition 1.5.1. *For P, Q two probability measures on a σ -algebra \mathcal{F} of subsets of the sample space Ω , we define the total variation distance by*

$$d_{TV}(P, Q) = \sum_{A \in \mathcal{F}} |P(A) - Q(A)|.$$

This is a natural metric to measure proximity between probability distributions. When Ω is countable, an equivalent definition is

$$d_{TV}(P, Q) = \frac{1}{2} \sum_{\omega \in \Omega} |P(\omega) - Q(\omega)|.$$

Note that some authors define this distance with a multiplicative constant 2. We will also abusively say that random variables converge in total variation when their distributions do. Finally, note that convergence in total variation implies weak convergence.

In this setting, we have the following result by Le Cam [57]:

Theorem 1.5.2 (Le Cam). *Let $(B_k)_{k \in \mathbb{N}}$ be a sequence of independent Bernoulli r.v.s with parameters p_k . For $N \in \mathbb{N}$, let $\lambda_N = \sum_{k=1}^N p_k$. Let $S_N = \sum_{k=1}^N B_k$. Let L_N be a Poisson-distributed random variable with mean λ_N . Then*

$$d_{TV}(S_N, L_N) \leq \sum_{k=1}^N p_k^2.$$

In particular, if $\lambda_N \rightarrow \lambda > 0$ when $N \rightarrow \infty$ and $\sum_{k=1}^N p_k^2 \rightarrow 0$ when $N \rightarrow \infty$, S_N converges to L in total variation when $N \rightarrow \infty$, where L is a Poisson random variable with parameter λ .

The proof relies on the following fundamental result, called the *coupling lemma*:

Lemma 1.5.3. *If X, Y are two random variables such that X admits the probability distribution P and Y admits the probability distribution Q ,*

$$d_{TV}(P, Q) \leq \mathbf{P}(X \neq Y).$$

Note that one can show that in fact, $d_{TV}(P, Q) = \inf_{(X, Y), X \sim P, Y \sim Q} \mathbf{P}(X \neq Y)$.

The idea of the proof of Le Cam's theorem is to find a coupling of S_N and L_N such that

$$\mathbf{P}(S_N \neq L_N) \leq \sum_{k=1}^N p_k^2.$$

Heuristically, this is done by considering the sum of N independent Poisson random variables with parameters p_k . We refer to [58] for extensive coverage of coupling techniques.

A functional approach: the Chen-Stein method

The previous method had the advantage of being very constructive, with the bound being given by an explicit coupling. In this section, we present a more modern point of view which uses tools derived from functional analysis. Even though it doesn't provide the satisfaction of an explicit construction, it has several benefits over the latter. First, it gives tighter bounds. Second, and perhaps most importantly, it allows us to relax the independence assumption for the Bernoulli random variables. In fact, it can be applied to a much more general class of distributions. Here, our focus will however remain on the Bernoulli case.

The fundamental idea behind the Chen-Stein method, developed by Chen [23] following the works of Stein in the Gaussian setting, is to characterize the target distribution, in this case the Poisson distribution, as the solution to some functional equation. Then, if a distribution is close to the solution to this functional equation, it should be close to the target distribution.

In the Poisson case, the functional equation in question is based on the following observation: if Z is a random variable taking values in \mathbb{N} with $\mathbf{E}[Z] < \infty$, Z is a Poisson random variable iff the distribution of $Z + 1$ is equal to the distribution of the size-biased version of Z , in other words, iff for all bounded functions f on \mathbb{N} ,

$$\mathbf{E}[Z] \mathbf{E}[f(Z + 1)] = \mathbf{E}[f(Z)Z]. \quad (1.2)$$

Morally, a distribution is close to Poisson if “adding a point” doesn’t heavily affect the distribution. One can see a link with the Slivnyak-Mecke theorem from Palm calculus, which we will not present here. See for example [13] for more details.

In the case of a sum of Bernoulli random variables that are not necessarily independent, we have the following result:

Lemma 1.5.4. *Let $l \in \mathbb{N}$. Consider $W = \sum_{i=1}^l Y_i$, where Y_i are Bernoulli random variables with respective means p_i , without any independence assumptions. Let Z be a Poisson distributed random variable with mean $\mathbf{E}[W] = \sum_i p_i$. For $1 \leq k \leq l$, let U_k and V_k be random variables on the same probability space such that U_k has the same distribution as W and $1 + V_k$ has the same distribution as W conditioned on the event $Y_k = 1$ (with the convention $V_k = 0$ if $\mathbf{P}(Y_k = 1) = 0$). Then*

$$d_{TV}(W, Z) \leq \left(1 \wedge \frac{1}{\mathbf{E}[W]}\right) \sum_{i=1}^l p_i \mathbf{E}[|U_i - V_i|].$$

Then, it suffices to exhibit a coupling of U_i and V_i such that $\mathbf{E}[|U_i - V_i|]$ is small. We refer to [13], [58] or [10] for a comprehensive overview of the Chen-Stein method.

A few references for Poisson process approximation

The base heuristic mentioned at the beginning of the section, namely that a sum of rare events is asymptotically Poisson, can be generalized to point processes in the following fashion: Grigelionis [39] has shown that on \mathbb{R}^+ , the superposition of independent sparse point processes weakly converges to a Poisson point process. Chen [24] has given error bounds for this approximation combining the Chen-Stein method and Palm theory.

1.6 Primer on replica-mean-field models

The goal of this section is to introduce the replica-mean-field models that are extensively used in the following chapters to prove the main results of this work, they are presented therein as well. As such, we focus here on the base principle of

the replica-mean-field approach, as well as the heuristics of certain properties of the replica-mean-field structure.

To contextualize the core of the replica-mean-field approach, it is helpful to contrast it with other existing mean-field-type models. By "mean-field-type", we refer here to models designed to converge to a regime of interest when scaled along a certain parameter.

In Section 1.4, we introduced the classical mean-field scaling, also referred to as the thermodynamical mean-field, in which the particle system consists of K particles with interactions averaged over all the system and the mean-field regime is obtained by letting K go to infinity. The key characteristic of this type of model is having the scaling of interactions be inversely proportional to K^α , with α usually equal to 1, sometimes 2.

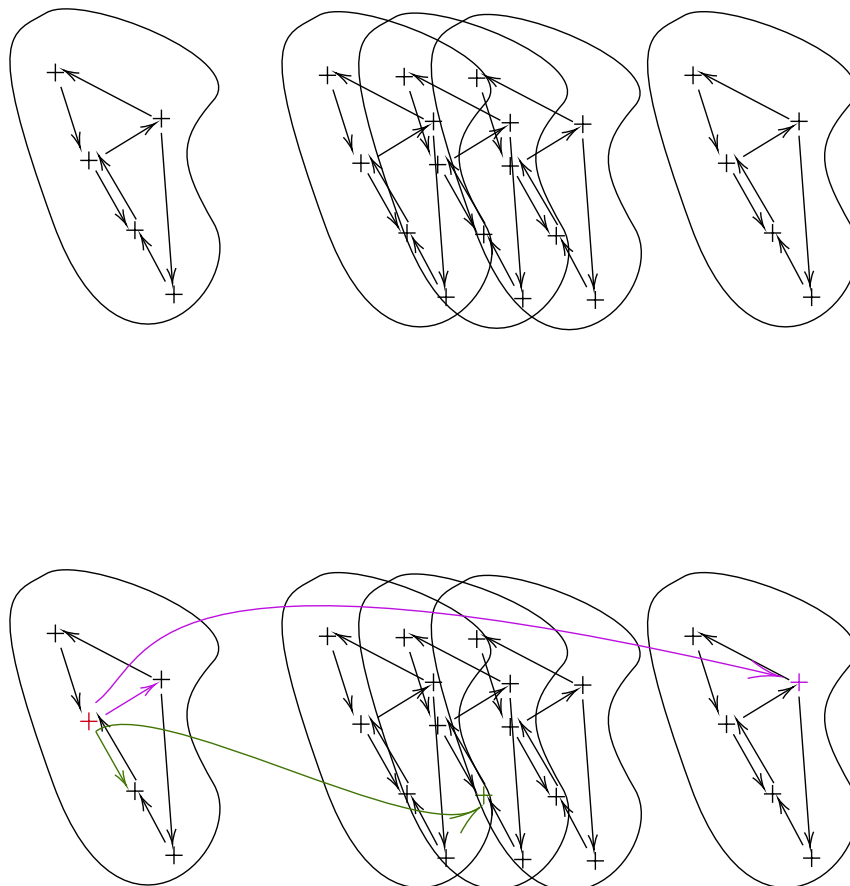
Another mean-field-type model is the routing-mean-field, which is the most direct inspiration for the replica-mean-field approach. Developed mostly in queueing theory by (among others) Dobrushin, Rybko and Borovkov [48] [16], it consists of a closed system of K queues in which a customer leaving a queue joins another queue selected uniformly at random upon departure. When K goes to infinity, the random routing heuristically leads to asymptotic Poisson-distributed arrivals. Note that the authors cited above also sometimes refer to their models as the thermodynamic mean-field. Unlike the models we refer to in that way, however, interactions are not averaged accross the system; the averaging only appears in expectation due to the random routing. This can be related to "power of two"-type routing models developed by Vvedenskaya [79], in which arriving customers are redirected to the shortest queue among a randomly selected amount of queues.

The main motivation behind the introduction of routing-mean-fields was replacing arrivals to nodes by Poisson processes, which is known as setting the Poisson Hypothesis [49]. The replica-mean-field approach aims to extend the setting of the routing-mean-field to the more general case of network dynamics. It consists in considering a physical system where interactions are uniformly randomly routed in between objects of interest. Where originally, this consisted in considering closed systems of queues where a customer would be randomly routed to another queue upon departing from one, and letting the number of queues go to infinity, in a recent work by Baccelli and Taillefumier [8], a replica-mean-field version of the Galves-Löcherbach network dynamics was introduced by considering a physical system consisting of randomly interacting replicas of the network of interest, and studying the limit when the number of replicas goes to infinity. This approach was later used by Coletti and Papageorgiou to study the asymptotic behavior of the Elephant Random Walk [26], assuming that the Poisson Hypothesis holds at the limit.

From the point of view of a single node of the network in one of the replicas,

the influence of the replica structure on its state is as follows: the aggregations to its state consist of the departures from all the other nodes in all the other replicas that have been routed to the given replica.

The M-replica dynamics



The main difference between the routing-mean-field and the replica-mean-field approaches lies in the fact that unlike the routing-mean-field, the scaling parameter in the replica-mean-field is the number of replicas, which is distinct from the number of objects in the network of interest that is fixed throughout.

Heuristically, the following observations explains why it is natural to conjecture that the Poisson Hypothesis holds at the limit of the infinite number of replicas. First, note that because of the uniform random routing, the probability that two nodes in two given replicas interact is inversely proportional to the number of replicas. This gives a strong indication that the replicas should be asymptotically

independent, which as we have seen previously is more often than not required to obtain tractability. Second, note that the arrivals to a given node can be seen as a random sum of Bernoulli random variables, which should under proper conditions converge, at least in a weak sense, to a Poisson distribution.

We shall see in Chapter 3 that for a general class of continuous-time dynamics, this convergence does indeed take place. A key property that allows this is the independence of the routings targeting a given node represented as Bernoulli random variables when conditioned on the number of departures from the other nodes in the other replicas. This is primarily due to the fact that arrivals to any node are always aggregated across replicas, which erases the eventual dependencies due to the routing choosing one replica rather than another.

1.7 Primer on epidemiological models and migration-contagion processes

Up to now, the main application for agent-based models that we have focused on concerned intensity-based models from computational neuroscience. Now, we present a different class of models of interest in life sciences, namely compartmental epidemiological models.

Compartmental epidemiological models

In compartmental epidemiological models, the idea is to study the spread of an epidemic in a population where individuals are split into different classes based on their current condition with regards to the epidemic. Popular classes are Susceptible (S), comprised of individuals that are not sick but that can get sick at a certain given rate, Infected (I), comprised of sick individuals which can heal from the sickness at another given rate and Recovered (R), comprised of individuals who had the sickness but no longer have it. Depending on the model, recovered individuals can sometimes become susceptible again at another given rate. More elaborate versions of these models incorporate other classes, such as Vaccinated (V), comprised of individuals immune to the infection but which become susceptible with a given rate, or Exposed (E), in which susceptible individuals must become exposed before being able to become infected.

In the infinite population limit, by assuming individuals as independent, the behavior of the system can often be reduced to a system of ODEs. Consider, for example, the base SIR model. Denoting by x_S, x_I and x_R the proportions of susceptible, infected and recovered individuals respectively, their evolution is

described by the following system of equations:

$$\begin{cases} \frac{dx_S}{dt}(t) &= -\beta x_S(t)x_I(t) \\ \frac{dx_I}{dt}(t) &= \beta x_S(t)x_I(t) - \gamma x_I(t) \\ \frac{dx_R}{dt}(t) &= \gamma x_I(t), \end{cases} \quad (1.3)$$

where β is the infection rate and γ is the recovery rate.

We refer to this an infinite population limit due to the following observation: in a fully connected finite population, consider a given agent is chosen to update, if the agent is susceptible for example, which occurs with probability $x_S(t)$. Then, they interact with another agent chosen uniformly at random, and if the agent is infected, which occurs with probability $x_I(t)$, the agent changes its state from S to I with probability β , and stays in the same state otherwise. Thus, the probability that a given agent changes its state from S to I is given by $\beta x_S(t)x_I(t)$. A similar interpretation hold for the I \rightarrow R transition. Then, when the number of agents goes to infinity, heuristically, the changes become infinitesimal with respect to the population, thus giving rise to the system of equations (1.3). This heuristic can be made rigorous [54] when the initial conditions of the finite system converge.

One question of interest for compartmental epidemiological models is the addition of a spatial component for the positions of the individuals, in which case an agent becomes infected with a rate dependent on the proportion of infected individuals among their neighbors instead of across the full network. Various works and approaches have been developed, we will mention a few here and refer to [65] for a comprehensive review of epidemic processes on networks. Pemanle and Stacey [66] have studied SIS dynamics (under the name contact process) on Galton-Watson trees and nonhomogeneous trees. They have shown conditions and bounds for phase transitions for the contact process. Degree-based mean-field models have been studied by Boguña and Pastor-Satorras [15] in which the SIS model is described in terms of the probability that a node of certain degree is infected at a given time, assuming the statistical equivalence of all nodes of same degree. Individual-based mean-field models were studied by many authors, see [61] or [22], in which the states of neighboring individuals are considered independent to obtain expressions for epidemic thresholds.

We shall see in the remainder of this section how queueing theory can be used to obtain another framework for epidemics propagation.

A quick reminder from queueing theory

Migration processes arise from another field of mathematics, namely queueing theory. We will recall a few basic notions and notation before presenting a class of

models combining these migration processes with epidemiological dynamics seen previously.

A queueing model consists of a service node treating customers arriving to a waiting area according to some arrivals process, and exiting the node after being serviced. The main questions of interest in queueing theory are stability of the system of queues, that is, whether the number of waiting customers stays finite when time becomes large, and the existence of a stationary state for the system.

A useful notation to describe different standard configurations of queues is Kendall's notation.

Definition 1.7.1. *A queueing model may be represented using Kendall's notation in the following fashion:*

$$A/S/c(/K/N/D),$$

where

- *A represents the distribution of the time between arrivals;*
- *S represents the service time distribution;*
- *c represents the number of service channels;*
- *K represents the capacity of the queue (the default value is ∞);*
- *N represents the number of jobs to be served (the default value is ∞);*
- *D represents the queueing discipline (the default is First In, First Out, or FIFO for short).*

The last three parameters are omitted by default and assumed to be set to their default values. Some common options for A and S are M (Markovian, corresponding to exponential inter-arrival times or service times) and G or GI (General Independent, corresponding to i.i.d. service times). c is most often 1 or infinity, although finite tandems of queues are also extensively studied.

Migration-contagion processes

Migration-contagion processes, introduced by Baccelli, Foss and Shneer in [6], propose to model epidemics by considering a simple migration model, a closed network of $\cdot/M/\infty$ queues, on which simple contagion dynamics, namely SIS dynamics, take place.

In a migration-contagion process, individuals, or customers, travel in between stations, called *reactors* when considering them as single open stations. Each customer is marked with their state, infected (I) or susceptible (S). Inside each

reactor, the customer states change according to mechanisms depending on the migration-contagion process chosen. In [6], several reactors were introduced in order to study the steady state of the queues. We will introduce here the main $M/M/\infty$ SIS reactor and present the main results pertaining to the proportion of infected (I) individuals in the stationary output point process of the reactor in the steady state. Afterwards, we will present two variant reactors, for which, unlike the SIS reactor, the stationary regimes admit closed-form expressions.

In the SIS reactor, consider a $M/M/\infty$ queue with input rate λ and service rate μ . Let $X(t)$ be the number of susceptible customers at time t , and $Y(t)$ be the number of infected customers at time t . Let $X(0) = x, Y(0) = y$. Input customers are infected with probability p , and susceptible with probability $q = 1 - p$. Inside the queue, any susceptible customer becomes infected with rate $\alpha Y(t)$, and any infected customer becomes susceptible with rate β .

In this setting, neither of the processes $X(t)$ or $Y(t)$ is Poisson. There is no product form, the means are unknown and there is no known closed form solution for the wave-type PDE verified by $\Phi(x, y) = \mathbf{E}[x^X y^Y]$:

$$\begin{aligned} \lambda q(1-x) + \lambda p(1-y)\Phi(x, y) &= \mu(1-x)\Phi_x(x, y) + (\mu(1-y) + \beta(x-y))\Phi_y(x, y) \\ &\quad + \alpha y(y-x)\Phi_{x,y}(x, y). \end{aligned}$$

One can, however, consider a closed network of SIS reactors. In this setting, queues have no input or output. Instead, when a customer leaves a queue, they choose the next station they go to uniformly at random. This network model can be simplified by taking the mean-field limit, namely, by letting the number N of stations go to infinity and, denoting K the total number of customers, $\frac{K}{N} \rightarrow \eta \in (0, \infty)$.

Under the thermodynamic limit ansatz, the stations become asymptotically independent and the Poisson Hypothesis holds: the arrivals to a given station become asymptotically Poisson. The proof of the ansatz is a current open research problem. We give the following definitions:

Definition 1.7.2. • We say that there is survival if there exists a steady state distribution with a fraction $0 < p < 1$ of susceptibles;

- We say there is weak extinction if there is no such p .
- We say there is strong extinction if for all initial conditions, there is convergence to a regime with no infected customers.

Then, the following result has been shown:

Theorem 1.7.3. *Provided that the thermodynamic limit ansatz holds, in the thermodynamic limit, there exists a constant η_C^s such that:*

- *there is survival if $\eta > \eta_C^s$;*
- *there is strong extinction if $\eta \leq \eta_C^s$.*

Note that there is no known closed form for η_C^s .

We now introduce two variants of the SIS reactor that have the benefit of admitting closed forms for the stationary state. The first will be of particular relevance for the study in Chapter 4. The first variant is called SIS-DOCS, as in Departure On Change of State: if a customer changes states, they immediately leave the reactor. In this setting, the PDE verified by the stationary generating function can be solved explicitly. Moreover, the stationary distribution of Y is Poisson with parameter $\frac{\lambda p}{\mu + \beta}$. Similarly to the SIS case, it is possible to consider the thermodynamic limit of N SIS-DOCS stations. Once again, the rigorous derivation of this limit is left as a conjecture. In the steady state, it can be shown that there exists a η_c^d playing a similar role to η_C^s .

The second variant is called SIS-AIR, as in Averaged Infection Rate. The reactor consists of an open network of two $M/M/\infty$ stations with input rate λ , service rate of any customer μ and state transition rates β from Infected to Susceptible and αy from Susceptible to Infected, where y is a parameter. A SIS-AIR reactor can thus be seen as an open Jackson network admitting a product-form stationary distribution. As before, it is possible to consider the thermodynamic limit of a closed SIS-AIR network with N stations, and to show the existence of $\eta_c^a = \frac{\beta}{\alpha}$ separating survival from (this time) weak extinction.

Chapter 2

Replica-mean-field limits for discrete-time processes

2.1 Introduction

This chapter consists in the article [5], published in *Advances in Applied Probability* in 2022 and written in collaboration with François Baccelli and Thibaud Taillefumier. Note that some notation has been changed for harmonization purposes between chapters and to follow the recommendation given at the end of Section 1.3.

Abstract

Network dynamics with point-process-based interactions are of paramount modeling interest. Unfortunately, most relevant dynamics involve complex graphs of interactions for which an exact computational treatment is impossible. To circumvent this difficulty, the replica-mean-field approach focuses on randomly interacting replicas of the networks of interest. In the limit of an infinite number of replicas, these networks become analytically tractable under the so-called “Poisson Hypothesis”. However, in most applications, this hypothesis is only conjectured. Here, we establish the Poisson Hypothesis for a general class of discrete-time, point-process-based dynamics, that we propose to call fragmentation-interaction-aggregation processes, and which are introduced in the present paper. These processes feature a network of nodes, each endowed with a state governing their random activation. Each activation triggers the fragmentation of the activated node state and the transmission of interaction signals to downstream nodes. In turn, the signals received by nodes are aggregated to their state. Our main contribution is a proof of the Poisson Hypothesis for the replica-mean-field version of any network in this class. The proof is obtained by establishing the propagation

of asymptotic independence for state variables in the limit of an infinite number of replicas. Discrete time Galves-Löcherbach neural networks are used as a basic instance and illustration of our analysis.

2.2 Discrete-time replica-mean-field limit for FIAPs

2.2.1 Introduction

Epidemics propagation, chemical reactions, opinion dynamics, flow control in the Internet, and even neural computations can all be modelled via punctuate interactions between interconnected agents [65][40][4][7][73]. The phenomena of interest in this context are idealized as network dynamics on a graph of agents which interact via point processes: edges between agents are the support of interactions, with edge-specific point processes registering the times at which these interactions are exerted. Such point-process-based network dynamics constitute a very versatile class of models able to capture phenomena in natural sciences, engineering, social sciences and economics. However, this versatility comes at the cost of tractability as the mathematical analysis of these dynamics is impossible except for the simplest network architectures. As a result, one has to resort to simplifying assumptions to go beyond numerical simulations.

Generic point-process-based networks are computationally untractable because their stochastic dynamics does not appear to belong to any known parametric class of point processes. Replica mean-field (RMF) limits are precisely meant to circumvent this obstacle [8]. The RMF limit of a given network is an extension of this network built in such a way that interaction point processes are parametric, e.g., Poisson. This extended network is made of infinitely many replicas of the initial network, all with the same basic structure, but with randomized interactions across replicas. The interest in RMF limits stems from the fact that they offer tractable version of the original dynamics that retain some of its most important features. The fact that Poisson point processes arise in the RMF version of a network is called the Poisson Hypothesis. Thus formulated, the Poisson Hypothesis originates from communication network theory [51] and is distinct from replica approaches developed in statistical physics [19].

Although intuitively clear and despite its usefulness, the Poisson Hypothesis is often only conjectured and/or numerically validated. The purpose of this work is to rigorously establish the Poisson Hypothesis for the RMF limits of a broad class of point-process-based network dynamics in discrete time introduced in the present paper. This class, which will be referred to as fragmentation-interaction-aggregation processes (FIAPs) below, includes important classes of queuing net-

works as special cases, as well as discrete time Galves-Löcherbach (GL) neural networks.

Galves-Löcherbach networks can be viewed as coupled Hawkes processes with spike-triggered memory resets. Because of these memory resets, it can be shown that the dynamics of finite-size Galves-Löcherbach networks is Markovian [70]. The RMF limit of the GL case was studied from a computational standpoint in [8] but in continuous time. In the next paragraph, we use results established in [8] to illustrate how the Poisson Hypothesis yields tractable mean-field equations for the stationary dynamics of these RMF limits. This is done for the simplest example of GL networks, called the “counting-neuron” model.

Illustration from the study of spiking neural networks

The counting-neuron model consists of a fully-connected network of K exchangeable neurons with homogeneous synaptic weights μ . For each neuron i , $1 \leq i \leq K$, the continuously time-indexed stochastic intensity λ_i increases by $\mu > 0$ upon reception of a spike and resets upon spiking to its base rate $b > 0$. Thus, its stochastic intensity is $\lambda_i(t) = b + \mu C_i(t)$, where $C_i(t)$ is the number of spikes received at time t since the last reset. It can be shown that the network state $\{C_1(t), \dots, C_K(t)\}$ has a well-defined stationary distribution. Despite of the simplicity of the model, analytic characterizations of the stationary state, including the stationary spiking rate, are hindered by the fact that the law of the point process of spike receptions is not known.

To circumvent this hindrance, the RMF setting proposes to compute stationary spiking rates in infinite networks that are closely related to the original finite-size networks. Informally, the counting-model RMF is constructed as follows: for a K -neuron counting model and for an integer $M > 0$, the M -replica model consists of M replicas, each comprising K counting neurons. Upon spiking, a neuron i in replica m , indexed by (m, i) , delivers spikes with synaptic weight μ to the $K - 1$ neurons (v_j, j) , $j \neq i$, where the replica destination v_j is chosen uniformly at random for all j . RMF networks are defined in the limit of an infinite number of replicas, namely infinite M but fixed and finite K . The Poisson Hypothesis then states that the dynamics of replicas become asymptotically independent in the limit $M \rightarrow \infty$, and that each neuron receives spikes from independent Poisson point processes. It is shown in [8] that as a consequence of this Poisson property, the stationary state is characterized by a single ODE bearing on G , the probability-generating function (PGF) of a neuron count C :

$$\beta - \mu z G'(z) + (\beta(K-1)(z-1) - b)G(z) = 0. \quad (2.1)$$

The simplifications warranted by the Poisson Hypothesis in the above ODE characterization comes at the cost of introducing the spiking rate β as an unknown

parameter in (2.1). As the ODE (2.1) is otherwise analytically tractable, characterizing the RMF stationary state amounts to specifying the unknown firing rate β . Then, the challenge of the RMF approach consists in specifying the unknown firing rate via purely analytical considerations about a parametric system of ODEs. It turns out that requiring that the solution of (2.1) be analytic, as any PGF shall be, is generally enough to exhibit self-consistent relations about the stationary rates. For instance, the RMF stationary spiking rate β of the RMF counting model is shown to be determined as the unique solution of

$$\beta = \frac{\mu c^a e^{-c}}{\gamma(a, c)} \quad \text{with} \quad a = \frac{(K-1)\beta + b}{\mu} \quad \text{and} \quad c = \frac{(K-1)\beta}{\mu}, \quad (2.2)$$

where γ denotes the lower incomplete Euler Gamma function.

We have shown that the above approach generalizes to networks with continuous state space, heterogeneous network structures [8], and including pairwise correlations [9]. In all cases, the Poisson Hypothesis is the cornerstone of a computational treatment. As a key step toward establishing the Poisson Hypothesis for continuous-time network dynamics, the goal of this work is to prove it for a broad class of discrete-time dynamics, which we refer to as fragmentation-interaction-aggregation processes.

Fragmentation-interaction-aggregation processes

In FIAPs, agents are graph nodes endowed with a state that evolves over time. The nodes are coupled via point processes which model punctuate interactions. Specifically, each node's state evolves in response to its input point process, and generates an output point process in a state-dependent manner. In all generality, the transformation of input into output point process can be viewed as a random map. In FIAPs, this map is defined through the following dynamics: (i) The *fragmentation* process is triggered by local *activation* events taking place on each node and which occur with a probability that depends on the state of the node. (ii) Each fragmentation event in turn triggers *interactions* between the nodes by creating input events in the neighboring nodes. (iii) Finally, the *aggregation* process consists in the integration of the input point processes to the states of each node.

Thus broadly defined, FIAPs offer a simple albeit general framework to analyze the phenomena alluded to above. The precise definition of FIAPs is given as follows:

Definition 2.2.1. *An instance of the class \mathcal{C} of discrete fragmentation-interaction-aggregation processes is determined by:*

- *An integer K representing the number of nodes;*

- A collection of initial conditions for the integer-valued state variables at step zero, which we denote by $\{X_i\}$, where $i \in \{1, \dots, K\}$;
- A collection of fragmentation random variables $\{U_i\}$, which are i.i.d. uniform in $[0, 1]$ and independent from $\{X_i\}$, where $i \in \{1, \dots, K\}$;
- A collection of fragmentation functions $\{g_{1,i} : \mathbb{N} \rightarrow \mathbb{N}\}_{i \in \{1, \dots, K\}}$ and $\{g_{2,i} : \mathbb{N} \rightarrow \mathbb{N}\}_{i \in \{1, \dots, K\}}$;
- A collection of bounded interaction functions $\{h_{j \rightarrow i} : \mathbb{N} \rightarrow \mathbb{N}\}_{i,j \in \{1, \dots, K\}}$;
- A collection of activation probabilities $\{\sigma_i(0), \sigma_i(1), \dots\}_{i \in \{1, \dots, K\}}$ verifying the conditions $\sigma_i(0) = 0$, and $0 < \sigma_i(1) \leq \sigma_i(2) \leq \dots \leq 1$ for all i .

The associated dynamics take as input the initial integer-valued state variables $\{X_i\}$ and define the state variables at the next step as

$$Y_i = g_{1,i}(X_i) \mathbb{1}_{\{U_i < \sigma_i(X_i)\}} + g_{2,i}(X_i) \mathbb{1}_{\{U_i > \sigma_i(X_i)\}} + A_i, \quad \forall i = 1, \dots, K, \quad (2.3)$$

with arrival processes

$$A_i = \sum_{j \neq i} h_{j \rightarrow i}(X_j) \mathbb{1}_{\{U_j < \sigma_j(X_j)\}}, \quad \forall i = 1, \dots, K. \quad (2.4)$$

The interpretation is as follows: node i activates with probability $\sigma_i(k)$ if its state X_i is equal to k . The state of this node is fragmented to $g_{1,i}(X_i)$ upon activation and to $g_{2,i}(X_i)$ otherwise. The activation of node i triggers an input of $h_{i \rightarrow j}(X_i)$ units to node j . Hence, the interaction functions encode the structure of the graph. The variable A_i gives the total number of arrivals to node i . This variable is aggregated to the state of the node as seen in (2.3). Note that considering $\sigma_i(0) = 0$ for all i ensures that state variables in state 0 cannot be fragmented.

The FIAP class \mathcal{C} encompasses many network dynamics relevant to queuing theory and mathematical biology. For example, taking $g_{1,i}(k) = k - 1$, $g_{2,i}(k) = k$ and $h_{j \rightarrow i}(k) = \mathbb{1}_{\{i=j+1 \pmod K\}}$, we recover an instance of Gordon-Newell queuing networks [50]. Taking $g_{1,i}(k) = 0$, $g_{2,i}(k) = k$ and $h_{j \rightarrow i}(k) = \mu_{j \rightarrow i} \in \mathbb{N}$ defines a discrete instance of Galves-Löcherbach dynamics for neural networks introduced above. Taking $g_{1,i}(k) = \lfloor \frac{k}{2} \rfloor$ and $g_{2,i}(k) = k + 1$, corresponds to aggregation-fragmentation processes modelling, e.g., TCP communication networks [7]. The class \mathcal{C} also includes certain discrete time Hawkes processes. Namely, if for each coordinate of a Hawkes vector process, we define its state as the sum over time of all its variations, then all discrete Hawkes processes that are Markov with respect to their so-defined state are in \mathcal{C} . Thus, the results of the present paper have potential computational implications in a wide set of application domains beyond the neural network setting used above to illustrate them.

The present paper is focused on discrete time versions of this type of dynamics as in, e.g., [72] [21]; note that continuous instances were also considered in the literature such as in [29], [59].

Replica models for fragmentation-interaction-aggregation networks

Finite RMF models are defined as a coupling of replicas of the network of interest by randomized routing decisions. For a FIAP, the state of its M -replica model is thus specified by a collection of state variables $X_{m,i}^M$, where m is the index of the replica and i corresponds to the index of the node in the original network. Instead of interacting with nodes within the same replica, an activated node i in replica m interacts with a downstream node j from a replica n chosen uniformly at random and independently. This randomization preserves essential features of the original dynamics such as the magnitude of interactions between nodes but degrades the dependence structure between nodes. Indeed, over a finite period of time, the probability for a particular node to receive an activation from another given node scales as $1/M$. Thus, as the number of replicas increases, interactions between distinct replicas become ever scarcer, intuitively leading to replica independence when $M \rightarrow \infty$. This asymptotic independence is the root of RMF computational tractability.

Here is the precise definition of the finite-replica version of a FIAP:

Definition 2.2.2. *For any process in \mathcal{C} , the associated M -replica dynamics is entirely specified by*

- *A collection of initial conditions for the integer-valued state variables at step zero, which we denote by $\{X_{m,i}^M\}$, where $m \in \{1, \dots, M\}$ and $i \in \{1, \dots, K\}$, such that for all M, m and i , $X_{m,i}^M = X_i$;*
- *A collection of fragmentation random variables $\{U_{m,i}\}$, which are i.i.d. uniform in $[0, 1]$ and independent from $\{X_{m,i}^M\}$, where $m \in \{1, \dots, M\}$ and $i \in \{1, \dots, K\}$;*
- *A collection of i.i.d. routing random variables $\{R_{(n,j) \rightarrow i}^M\}$ independent from $\{X_{m,i}^M\}$ and $\{U_{m,i}\}$, uniformly distributed on $\{1, \dots, M\} \setminus \{n\}$ for all $i, j \in \{1, \dots, K\}$ and $n \in \{1, \dots, M\}$. In other words, if $R_{(n,j) \rightarrow i}^M = m$, then an eventual activation of node j in replica n at step 0 induces an arrival of size $h_{j \rightarrow i}(X_{n,j}^M)$ in node i of replica m , and m is chosen uniformly among replicas and independently from the state variables. Note that these variables are defined regardless of the fact that an activation actually occurs. Also note that for $i' \neq i$, the activation in question will induce an arrival in node i' of replica m' , with m' sampled in the same way but independently of m .*

Then, the integer-valued state variables at step one, denoted by $\{Y_{m,i}^M\}$, are given by the M -RMF equations

$$Y_{m,i}^M = g_{1,i}(X_{m,i}^M) \mathbb{1}_{\{U_{m,i} < \sigma_i(X_{m,i}^M)\}} + g_{2,i}(X_{m,i}^M) \mathbb{1}_{\{U_{m,i} > \sigma_i(X_{m,i}^M)\}} + A_{m,i}^M, \quad (2.5)$$

where $g_{1,i}$, $g_{2,i}$ denotes fragmentation functions, σ_i denotes activation probabilities, and where

$$A_{m,i}^M = \sum_{n \neq m} \sum_{j \neq i} h_{j \rightarrow i}(X_{n,j}^M) \mathbb{1}_{\{U_{n,j} < \sigma_j(X_{n,j}^M)\}} \mathbb{1}_{\{R_{(n,j) \rightarrow i}^M = m\}} \quad (2.6)$$

is the number of arrivals to node i of replica m via the interaction functions $h_{j \rightarrow i}$.

RMF models are only expected to become tractable when individual replicas become independent. This happens in the limit of an infinite number of replicas, i.e., in the so-called *RMF limit* [8]. In this RMF limit, asymptotic independence between replicas follows from the more specific Poisson Hypothesis. The Poisson Hypothesis states that spiking deliveries to distinct replicas shall be asymptotically distributed as independent Poisson (or compound) point processes. Such a hypothesis, which has been numerically validated for certain RMF networks, has been conjectured for linear Galves-Löcherbach dynamics in [8]. Proving the validity of the Poisson Hypothesis for the RMF limits of the much more general FIAPS is the purpose of the present work.

Methodology for proving the Poisson Hypothesis

Classical mean-field approximations of a given network are obtained by considering the limit of the original network when a certain characteristic of the network – typically the number of nodes – goes to infinity. When the dynamics of the nodes are synchronous, one gets a discrete time dynamical system. The term mean-field comes from the fact that in such network limits, the effect that individual nodes have on one another are approximated by a single averaged effect, typically an empirical mean. In the limit, this empirical mean usually converges to an expectation term through a *propagation of chaos* result [75] which leads to analytical tractability. In replica mean-fields, there is no such empirical mean over the nodes of the network; the mean-field simplification comes from the random routing operations between replicas. The input point process in the M -replica model consists in a superposition of M rare point processes, which informally explains why Poisson (or compound Poisson) processes arise at the limit. For classical mean-fields, different techniques have been developed to prove the existence and the convergence to the mean-field limit. Standard techniques include the use of the theory of nonlinear Markov processes [78] and stochastic approximation algorithms [11] for continuous time dynamics, and induction techniques which assume the existence of limits at time zero and extend the result by induction [56] for dynamics in discrete time. Refinements to the latter approach can be made in order to obtain explicit rates of convergence [36]. The approach developed for the RMF case belongs in spirit to the third class of techniques. We suppose that the property of asymptotic independence (see Definition 2.2.3) holds for the state variables at time zero. We then

prove that this property is preserved by the dynamics of the M -replica model and thus holds by induction for any finite time. Thus, we focus hereafter on the one-step transition of the model from time 0 to time 1. We show that this asymptotic independence hypothesis implies both convergence in distribution and an ergodic type property that we call the *triangular law of large numbers*. We apply this law of large numbers to the input process to a single node to show that Poisson (or compound Poisson) processes appear in the replica mean-field limit indeed. Let us stress that this proof is by induction. The fact that the main difficulty consists in proving the induction step should not hide the fact that the result relies in crucial way on the assumption that at step 0, the initial state variables satisfy the asymptotic independence property. Whether the result can be extended to more general initial conditions is an open question at this stage.

Structure of the Paper

For the sake of clarity in exposition, we start with the proof of the Poisson Hypothesis for the special case of neural networks first before extending it to general FIAPs. More precisely, we first consider the symmetric neural network case, which is a fully symmetric Galves-Löcherbach model [35] in discrete time. We introduce the model in Section 2.2.2 and prove the Poisson Hypothesis in Section 2.2.3. We then extend the proof to the class of FIAPs defined above. We first consider the symmetric case in Section 2.2.4 and then the general case in Section 2.2.5. Finally, some extensions are discussed in Section 2.2.6.

2.2.2 The symmetric Galves-Löcherbach model

The symmetric RMF network model

We consider a network of K spiking neurons. We suppose that the behavior of each neuron is determined by a random variable representing the membrane potential of the neuron. Each neuron spikes at a rate depending on its state variable. Let $X = \{X_i\}$ be the integer-valued state variables at step 0, where $i \in \{1, \dots, K\}$. Let $Y = \{Y_i\}$ be the integer-valued state variables at time one. The system continues to evolve in discrete time with all corresponding state variables defined by induction.

Let $\sigma : \mathbb{N} \rightarrow [0, 1]$ be the spiking probabilities of the neurons. Namely, $\sigma(k)$ is the probability that a neuron in state k spikes. We consider that $\sigma(0) = 0$, accounting for the fact that a neuron in state 0 never spikes. We also consider that $\sigma(1) > 0$ and that σ is non-decreasing. Let $\{U_i\}$ be uniformly distributed i.i.d. random variables independent from $\{X_i\}$. We then write the following evolution

equation for the state of the system:

$$Y_i = \mathbb{1}_{\{U_i > \sigma(X_i)\}} X_i + A_i, \quad (2.7)$$

where

$$A_i = \sum_{j \neq i} \mathbb{1}_{\{U_j < \sigma(X_j)\}} \quad (2.8)$$

is the number of arrivals to neuron i .

Here, the fragmentation is complete if $U_i < \sigma(X_i)$, namely if there is a spike, in which case the state variable is reset (jumps to 0). Otherwise there is no fragmentation at all and the state variable is left unchanged. In both cases, the arrivals A_i are aggregated to the state.

The RMF model described below is a discrete time version of the model introduced in [8]. Namely, we consider a collection of M identically distributed replicas of the initial set of K neurons. Let $X = \{X_{m,i}^M\}$ be the integer-valued state variables at step 0, where $m \in \{1, \dots, M\}, i \in \{1, \dots, K\}$. Let $Y = \{Y_{m,i}^M\}$ be the integer-valued state variables at time one. Let $U = \{U_{m,i}\}$ be uniformly i.i.d. random variables in $[0, 1]$ independent from $\{X_{m,i}^M\}$. Let $R = \{R_{(n,j) \rightarrow i}^M\}$ be i.i.d. *routing* random variables independent from $\{X_{m,i}^M\}$ and $\{U_{m,i}\}$, uniformly distributed on $\{1, \dots, M\} \setminus \{n\}$ for all $i, j \in \{1, \dots, K\}$ and $n \in \{1, \dots, M\}$. The replica model has the following evolution equation:

$$Y_{m,i}^M = \mathbb{1}_{\{U_{m,i} > \sigma(X_{m,i}^M)\}} X_{m,i}^M + A_{m,i}^M, \quad (2.9)$$

where

$$A_{m,i}^M = \sum_{n \neq m} \sum_{j \neq i} \mathbb{1}_{\{U_{n,j} < \sigma(X_{n,j}^M)\}} \mathbb{1}_{\{R_{(n,j) \rightarrow i}^M = m\}} \quad (2.10)$$

is the number of arrivals to neuron i of replica m .

Pairwise asymptotic independence and consequences

Our goal is to show the propagation of chaos and the Poisson Hypothesis in this system. In other words, we want to show that the arrivals to two distinct replicas are asymptotically independent and the number of arrivals to one replica is asymptotically Poisson distributed. We begin by considering the fully exchangeable case with equal weights, but we will consider the general case later. In order to do so, we choose to characterize the propagation of chaos through the following properties:

Definition 2.2.3. *Given $M \in \mathbb{N}$, given an array of integer-valued random variables $Z = \{Z_{m,i}^M\}_{1 \leq m \leq M, 1 \leq i \leq K}$ such that for all fixed M , the random variables $Z_{m,i}^M$*

are exchangeable in m and i , we say that the variables $Z_{m,i}^M$ are pairwise asymptotically independent, which we will denote $\text{PAI}(Z)$, if there exists an integer-valued random variable \tilde{Z} such that for all $(m, i) \neq (n, j)$, for all $u, v \in [0, 1]$,

$$\lim_{N \rightarrow \infty} \mathbf{E}[u^{Z_{m,i}^N} v^{Z_{n,j}^N}] = \mathbf{E}[u^{\tilde{Z}}] \mathbf{E}[v^{\tilde{Z}}]. \quad (2.11)$$

Definition 2.2.4. Given $M \in \mathbb{N}$, given an array of integer-valued random variables $Z = \{Z_m^M\}_{m \in \{1, \dots, M\}}$ such that for all fixed M , the random variables Z_m^M are exchangeable in m , we say that Z verifies the triangular law of large numbers, denoted by $\text{TLLN}(Z)$, if there exists an integer-valued random variable \tilde{Z} such that for all functions $f : \mathbb{N} \rightarrow \mathbb{R}$ with compact support, we have the following limit in L^2 :

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f(Z_n^N) = \mathbf{E}[f(\tilde{Z})]. \quad (2.12)$$

Here are a few remarks about these definitions. First, note that if an array of random variables Z satisfies $\text{PAI}(Z)$, then for all m and i , $Z_{m,i}^M$ converges in distribution to \tilde{Z} as $M \rightarrow \infty$. This can be seen by taking $v = 1$ in the definition. By considering the case where $Z_n^M = Z_1^1$ for all n and M , we see that the convergence in distribution of Z_n^M does not imply $\text{TLLN}(Z)$. However, we show below that for all arrays of random variables $Z = \{Z_{m,i}^M\}_{m \in \{1, \dots, M\}, i \in \{1, \dots, K\}}$ satisfying $\text{PAI}(Z)$, for all i , $Z_i = \{Z_{m,i}^M\}_{m \in \{1, \dots, M\}}$ satisfies $\text{TLLN}(Z_i)$. In other words, pairwise asymptotic independence of an array of random variables implies that these random variables verify the triangular law of large numbers. Finally, note that an array of integer-valued random variables Z satisfies $\text{PAI}(Z)$ iff the random variables are asymptotically independent in the sense that for all $(m, i) \neq (n, j)$

$$\mathbf{P}(Z_{m,i}^M \in B_1, Z_{n,j}^M \in B_2) \rightarrow \mathbf{P}(\tilde{Z} \in B_1) \mathbf{P}(\tilde{Z} \in B_2) \quad (2.13)$$

when $M \rightarrow \infty$ for $B_1, B_2 \in \mathcal{B}(\mathbb{R})$.

The following characterization of L^2 convergence will be used throughout this paper:

Lemma 2.2.5. *Let (X_n) be random variables with finite second moments. Then there exists a constant c such that $X_n \rightarrow c$ in L^2 when $n \rightarrow \infty$ iff*

1. $\mathbf{E}[X_n] \rightarrow c$ when $n \rightarrow \infty$
2. $\text{Var}(X_n) \rightarrow 0$ when $n \rightarrow \infty$.

This follows directly from the definition of L^2 convergence.

The following lemma describes the relation between pairwise asymptotic independence and the triangular law of large numbers.

Lemma 2.2.6. *Let $M \in \mathbb{N}$, let $Z = \{Z_{m,i}^M\}_{m \in \{1, \dots, M\}, i \in \{1, \dots, K\}}$ be an array of integer valued random variables verifying PAI(Z). Then, for all i , $Z_i = \{Z_{m,i}^M\}_{m \in \{1, \dots, M\}}$ satisfies TLLN(Z_i).*

Proof. Let $f : \mathbb{N} \rightarrow \mathbb{R}$ be a function with compact support. We use Lemma 2.2.5. We fix $i \in \{1, \dots, K\}$ that we omit in the rest of the proof. We have

$$\begin{aligned} \text{Var} \left(\frac{1}{M} \sum_{n=1}^M f(Z_n^M) \right) &= \frac{1}{M^2} \left(\sum_{n=1}^M \text{Var} (f(Z_n^M)) + \sum_{p \neq q} \text{cov}[f(Z_p^M), f(Z_q^M)] \right) \\ &= \frac{1}{M} \text{Var} (f(Z_1^M)) + \frac{M(M-1)}{M^2} \text{cov}[f(Z_1^M), f(Z_2^M)], \end{aligned}$$

the last equality holding by exchangeability between replicas. Both terms on the right hand side go to 0 when $M \rightarrow \infty$. For the first term, this follows from the boundedness of f . For the second, we first show the result for indicator functions. Let $B \in \mathcal{B}(\mathbb{R})$ and let f be defined by $f(n) = \mathbb{1}_{\{n \in B\}}$. Then we have

$$\text{cov}[f(Z_1^M), f(Z_2^M)] = \mathbf{P}(Z_1^M \in B, Z_2^M \in B) - \mathbf{P}(Z_1^M \in B)\mathbf{P}(Z_2^M \in B), \quad (2.14)$$

which goes to 0 when $M \rightarrow \infty$ by PAI(Z). This immediately extends to functions with compact support since they only take a finite number of values. Moreover, for all such functions, $\mathbf{E}[\frac{1}{M} \sum_{n=1}^M f(Z_n^M)] \rightarrow \mathbf{E}[f(\tilde{Z})]$ when $M \rightarrow \infty$ as a direct consequence of the fact that for integer-valued random variables, convergence in distribution of Z^M to \tilde{Z} is equivalent to the convergence $\mathbf{P}(Z^M = k) \rightarrow \mathbf{P}(\tilde{Z} = k)$ for all $k \in \mathbb{N}$. This concludes the proof. \square

For our subsequent needs, we also establish the following result: we show that pairwise asymptotic independence implies a property that is slightly more general than the triangular law of large numbers, where we allow the function f to depend on an array of i.i.d. random variables $U = \{U_{m,i}\}_{m \in \{1, \dots, M\}, i \in \{1, \dots, K\}}$, independent from the rest of the dynamics.

Lemma 2.2.7. *Let $M \in \mathbb{N}$, let $Z = \{Z_{m,i}^M\}_{m \in \{1, \dots, M\}, i \in \{1, \dots, K\}}$ be an array of integer valued random variables verifying PAI(Z). Then for all bounded functions $f : \mathbb{N} \times [0, 1] \rightarrow \mathbb{R}$ with compact support, for all i.i.d. sequences of random variables $U = \{U_{m,i}\}_{m \in \{1, \dots, M\}, i \in \{1, \dots, K\}}$ independent from Z , there exists \tilde{U} independent from \tilde{Z} and Z such that, for all $i \in \{1, \dots, K\}$, we have the following limit in L^2 :*

$$\lim_{M \rightarrow \infty} \frac{1}{M} \sum_{m=1}^M f(Z_{m,i}^M, U_{m,i}) = \mathbf{E}[f(\tilde{Z}, \tilde{U})]. \quad (2.15)$$

Note that compared to Definition 2.2.4, we consider that the functions are bounded, a condition that was automatically fulfilled for functions with compact support on \mathbb{N} .

Proof. We proceed as in the last lemma, conditioning on the $U_{m,i}$ when necessary. Let $M \in \mathbb{N}$, let $i \in \{1, \dots, K\}$. We will omit this index in the rest of the proof. By exchangeability between replicas, defining $\tilde{U} = U_1$, we have

$$\mathbf{E} \left[\frac{1}{M} \sum_{m=1}^M f(Z_m^M, U_m) \right] = \mathbf{E}[f(Z_1^M, U_1)] = \mathbf{E}[f(Z_1^M, \tilde{U})]$$

Since Z_1^M converges in distribution to \tilde{Z} when $M \rightarrow \infty$, and since Z_1^M is integer-valued and f is bounded, for all $u \in [0, 1]$, $\mathbf{E}[f(Z_1^M, u)] \rightarrow \mathbf{E}[f(\tilde{Z}, u)]$ when $M \rightarrow \infty$. Therefore, since \tilde{U} is independent from Z and \tilde{Z} , $\mathbf{E}[f(Z_1^M, \tilde{U})] \rightarrow \mathbf{E}[f(\tilde{Z}, \tilde{U})]$ when $M \rightarrow \infty$ a.s.. Finally,

$$\mathbf{E} \left[\frac{1}{M} \sum_{m=1}^M f(Z_m^M, U_m) \right] \rightarrow \mathbf{E} [f(\tilde{Z}, \tilde{U})] \quad (2.16)$$

when $M \rightarrow \infty$. Moreover,

$$\begin{aligned} \text{Var} \left(\frac{1}{M} \sum_{m=1}^M f(Z_m^M, U_m) \right) &= \frac{1}{M^2} \sum_{m=1}^M \text{Var} (f(Z_m^M, U_m)) + \frac{1}{M^2} \sum_{m \neq m'} \text{cov} [f(Z_m^M, U_m), f(Z_{m'}^M, U_{m'})] \\ &= \frac{1}{M} \text{Var} (f(Z_1^M, U_1)) + \frac{M(M-1)}{M^2} \text{cov} [f(Z_1^M, U_1), f(Z_2^M, U_2)], \end{aligned}$$

the last equality stemming from exchangeability between replicas. When $M \rightarrow \infty$, the first term goes to 0 because f is bounded. For the second term, since the $\{Z_m^M\}$ and the $\{U_m\}$ are independent and the $\{U_m\}$ are i.i.d., we can proceed as above. Namely, let $B, C \in \mathcal{B}(\mathbb{R})$. Let f be defined by $f(m, t) = \mathbb{1}_{\{m \in B\}} \mathbb{1}_{\{t \in C\}}$. Then we have

$$\begin{aligned} \text{cov}[f(Z_1^M, U_1), f(Z_2^M, U_2)] &= \mathbf{P}(Z_1^M \in B, Z_2^M \in B, U_1 \in C, U_2 \in C) \\ &\quad - \mathbf{P}(Z_1^M \in B, U_1 \in C) \mathbf{P}(Z_2^M \in B, U_2 \in C) \\ &= (\mathbf{P}(Z_1^M \in B, Z_2^M \in B) - \mathbf{P}(Z_1^M \in B) \mathbf{P}(Z_2^M \in B)) \mathbf{P}(U_1 \in C) \mathbf{P}(U_2 \in C), \end{aligned}$$

the last equality holding by independence between Z and $\{U_{m,i}\}_{m \in \{1, \dots, M\}, i \in \{1, \dots, K\}}$. The right hand term goes to 0 when $M \rightarrow \infty$ by PAI(Z). This generalizes to bounded functions with compact support, which concludes the proof. \square

Main result

Our goal is to show that if $X = \{X_{m,i}^M\}$ are asymptotically independent, then $Y = \{Y_{m,i}^M\}$ are as well. In other words, if we choose initial conditions that verify a certain property, this property will hold by induction at any finite discrete time.

Theorem 2.2.8. *Let $M \in \mathbb{N}$, let $X = \{X_{m,i}^M\}_{m \in \{1, \dots, M\}, i \in \{1, \dots, K\}}$ be an array of integer valued random variables (the “state variables”). Suppose that $\text{PAI}(X)$ holds. Then $\text{PAI}(Y)$ holds as well, where Y is defined by (2.9). Moreover, the arrivals to a given node $A_{m,i}^M$ converge in distribution to a Poisson random variable when $M \rightarrow \infty$.*

Note that the result depends on a choice of initial conditions verifying $\text{PAI}(X)$, a typical example of which is i.i.d. initial conditions stable in law, in the sense that their law does not depend on M . The question of whether given an arbitrary initial condition, the dynamics become pairwise asymptotically independent after some (finite or infinite) amount of time, is still open. Note also that this shows that we have convergence in distribution of the exchangeable variables $\{Y_m^M\}$ when $M \rightarrow \infty$.

2.2.3 The proof

In the following proof, since K is always finite and all considered random variables are exchangeable, as above, we will sometimes omit the neuron index $i \in \{1, \dots, K\}$ in order to simplify notation. Tilde superscripts will refer to objects in the infinite replica limit. Hat superscripts will refer to fragmentation processes.

Step one: fragmentation

Lemma 2.2.9. *Let $\hat{X} = \{\hat{X}_{m,i}^M = X_{m,i}^M \mathbb{1}_{\{U_{m,i} > \sigma(X_{m,i}^M)\}}\}$. Then $\text{PAI}(X)$ implies $\text{PAI}(\hat{X})$.*

Proof. We have for $u, v \in [0, 1]$,

$$\mathbf{E}[u^{\hat{X}_1^M} v^{\hat{X}_2^M}] = \sum_{k, l \in \mathbb{N}} \mathbf{P}(\hat{X}_1^M = k, \hat{X}_2^M = l) u^k v^l. \quad (2.17)$$

For $k, l > 0$, we have

$$\mathbf{P}(\hat{X}_1^M = k, \hat{X}_2^M = l) = \mathbf{P}(X_1^M = k, X_2^M = l)(1 - \sigma(k))(1 - \sigma(l)). \quad (2.18)$$

Similarly,

$$\begin{aligned} \mathbf{P}(\hat{X}_1^M = k, \hat{X}_2^M = 0) &= \sum_{l \in \mathbb{N}} \mathbf{P}(X_1^M = k, X_2^M = l)(1 - \sigma(k))\sigma(l), \forall k > 0 \\ \mathbf{P}(\hat{X}_1^M = 0, \hat{X}_2^M = 0) &= \sum_{k, l \in \mathbb{N}} \mathbf{P}(X_1^M = k, X_2^M = l)\sigma(k)\sigma(l). \end{aligned}$$

Since PAI(X) holds, for all $k, l \in \mathbb{N}$, $\mathbf{P}(X_1^M = k, X_2^M = l) \rightarrow \mathbf{P}(\tilde{X} = k)\mathbf{P}(\tilde{X} = l)$ when $M \rightarrow \infty$. Since all considered functions are bounded by 1, we have that for all $k, l \in \mathbb{N}$,

$$\mathbf{P}(\hat{X}_1^M = k, \hat{X}_2^M = l) \rightarrow \mathbf{P}(\tilde{\hat{X}} = k)\mathbf{P}(\tilde{\hat{X}} = l)$$

when $M \rightarrow \infty$, where $\tilde{\hat{X}} = \tilde{X} \mathbb{1}_{\{U > \sigma(\tilde{X})\}}$. This shows that

$$\mathbf{E}[u^{\hat{X}_1^M} v^{\hat{X}_2^M}] \rightarrow \mathbf{E}[u^{\tilde{\hat{X}}}] \mathbf{E}[v^{\tilde{\hat{X}}}] \quad (2.19)$$

when $M \rightarrow \infty$, which concludes the proof. \square

Step two: asymptotic behavior of the arrivals processes

We now show that the number of arrivals $A_{m,i}^M$, defined in (2.10) is asymptotically Poisson as the number of replicas goes to infinity. This is precisely the Poisson Hypothesis introduced in [51].

Lemma 2.2.10. *Supposing that PAI(X) holds, when $M \rightarrow \infty$, we have the convergence in distribution $A_{m,i}^M \rightarrow Poi((K-1)\theta)$ where $\theta = \mathbf{E}[\sigma(\tilde{X})]$.*

Proof. Let $z \in [0, 1]$. Then

$$\begin{aligned} \mathbf{E}[z^{A_{m,i}^M}] &= \mathbf{E} \left[z^{\sum_{n \neq m} \sum_{j \neq i} \mathbf{1}_{\{U_{n,j} < \sigma(X_{n,j}^M)\}} \mathbf{1}_{\{R_{(n,j)}^M \rightarrow i = m\}}} \right] \\ &= \mathbf{E} \left[\prod_{n \neq m} \prod_{j \neq i} \mathbf{E} \left[z^{\mathbf{1}_{\{U_{n,j} < \sigma(X_{n,j}^M)\}} \mathbf{1}_{\{R_{(n,j)}^M \rightarrow i = m\}}} \middle| X_{n,j}^M, U \right] \right] \\ &= \mathbf{E} \left[\prod_{n \neq m} \prod_{j \neq i} \left(\left(1 - \frac{1}{M-1} \right) + \frac{1}{M-1} z^{\mathbf{1}_{\{U_{n,j} < \sigma(X_{n,j}^M)\}}} \right) \right] \\ &= \mathbf{E} \left[e^{\sum_{n \neq m} \sum_{j \neq i} \log \left(1 - \frac{1}{M-1} \left(1 - z^{\mathbf{1}_{\{U_{n,j} < \sigma(X_{n,j}^M)\}}} \right) \right)} \right]. \end{aligned}$$

We now give an upper and lower bound for this expression. Since $\log(1-x) \leq -x$ for $x \leq 1$, we have

$$\mathbf{E} \left[z^{A_{m,i}^M} \right] \leq \mathbf{E} \left[e^{-\frac{1}{M-1} \sum_{n \neq m} \sum_{j \neq i} \left(1 - z^{\mathbf{1}_{\{U_{n,j} < \sigma(X_{n,j}^M)\}} \right)} \right].$$

Using the generalized TLLN given in Lemma 2.2.7,

$$\frac{1}{M-1} \sum_{n \neq m} \sum_{j \neq i} (1 - z^{\mathbf{1}_{\{U_{n,j} < \sigma(X_{n,j}^M)\}}}) \rightarrow (K-1)(1 - \Phi(z))$$

in L^2 when $M \rightarrow \infty$ with $\Phi(z) = \mathbf{E}[z^{\mathbf{1}_{U < \sigma(\tilde{X})}}]$, where U is any $U_{n,j}$.

We have $\Phi(z) = z \int_0^1 \mathbf{P}(\sigma(\tilde{X}) > t) dt + (1 - \int_0^1 \mathbf{P}(\sigma(\tilde{X}) > t) dt) = (z-1)\theta + 1$.

Therefore, since L^2 convergence implies convergence in distribution and thus convergence of the Laplace transforms,

$$\mathbf{E} \left[e^{-\frac{1}{M-1} \sum_{n \neq m} \sum_{j \neq i} \left(1 - z^{\mathbf{1}_{\{U_{n,j} < \sigma(X_{n,j}^M)\}} \right)} \right] \rightarrow e^{-\theta(1-z)(K-1)}$$

when $M \rightarrow \infty$. Thus,

$$\limsup_{M \rightarrow \infty} \mathbf{E}[z^{A_{m,i}^M}] \leq e^{-\theta(1-z)(K-1)}. \quad (2.20)$$

Similarly, since $\log(1-x) \geq -x - \frac{x^2}{2}$ for $x \leq 1$, we have

$$\mathbf{E} \left[z^{A_{m,i}^M} \right] \geq \mathbf{E} \left[e^{-\frac{1}{M-1} \sum_{n \neq m} \sum_{j \neq i} \left(1 - z^{\mathbf{1}_{\{U_{n,j} < \sigma(X_{n,j}^M)\}} \right)} e^{-\frac{1}{2(M-1)^2} \sum_{n \neq m} \sum_{j \neq i} \left(1 - z^{\mathbf{1}_{\{U_{n,j} < \sigma(X_{n,j}^M)\}} \right)^2} \right].$$

Using once again Lemma 2.2.7, as the second term goes to 0 when $M \rightarrow \infty$, by the same reasoning as previously, we get

$$\liminf_{M \rightarrow \infty} \mathbf{E}[z^{A_{m,i}^M}] \geq e^{-\theta(1-z)(K-1)}. \quad (2.21)$$

Combining (2.20) and (2.21), the result follows. \square

Now, we show that the arrivals to different replicas become pairwise asymptotically independent:

Lemma 2.2.11. *For all $(m, i) \neq (n, j)$, $A_{m,i}^M$ and $A_{n,j}^M$ are pairwise asymptotically independent.*

Proof. We first show the result in the case $m \neq n$ and $i \neq j$. Let $u, v \in [0, 1]$. Then

$$\begin{aligned}
\mathbf{E} \left[u^{A_{m,i}^M} v^{A_{n,j}^M} \right] &= \mathbf{E} \left[u^{\sum_{m' \neq m, i' \neq i} \mathbf{1}_{\{U_{m',i'} < \sigma(X_{m',i'}^M)\}} \mathbf{1}_{\{R_{(m',i')}^M \rightarrow i = m\}}} \right. \\
&\quad \left. v^{\sum_{n' \neq n, j' \neq j} \mathbf{1}_{\{U_{n',j'} < \sigma(X_{n',j'}^M)\}} \mathbf{1}_{\{R_{(n',j')}^M \rightarrow j = n\}}} \right] \\
&= \mathbf{E} \left[\prod_{m' \neq m, i' \neq i} u^{\mathbf{1}_{\{U_{m',i'} < \sigma(X_{m',i'}^M)\}} \mathbf{1}_{\{R_{(m',i')}^M \rightarrow i = m\}}} \right. \\
&\quad \left. \prod_{n' \neq n, j' \neq j} v^{\mathbf{1}_{\{U_{n',j'} < \sigma(X_{n',j'}^M)\}} \mathbf{1}_{\{R_{(n',j')}^M \rightarrow j = n\}}} \right] \\
&= \mathbf{E} \left[\mathbf{E} \left[\prod_{m' \neq m, i' \neq i} u^{\mathbf{1}_{\{U_{m',i'} < \sigma(X_{m',i'}^M)\}} \mathbf{1}_{\{R_{(m',i')}^M \rightarrow i = m\}}} \right. \right. \\
&\quad \left. \left. \prod_{n' \neq n, j' \neq j} v^{\mathbf{1}_{\{U_{n',j'} < \sigma(X_{n',j'}^M)\}} \mathbf{1}_{\{R_{(n',j')}^M \rightarrow j = n\}}} \middle| X^M, U \right] \right] \\
&= \mathbf{E} \left[\prod_{m' \neq m, i' \neq i} \left[\left(1 - \frac{1}{M-1} \right) + \frac{1}{M-1} u^{\mathbf{1}_{\{U_{m',i'} < \sigma(X_{m',i'}^M)\}}} \right] \right. \\
&\quad \left. \prod_{n' \neq n, j' \neq j} \left[\left(1 - \frac{1}{M-1} \right) + \frac{1}{M-1} v^{\mathbf{1}_{\{U_{n',j'} < \sigma(X_{n',j'}^M)\}}} \right] \right] \\
&= \mathbf{E} \left[e^{\sum_{m' \neq m, i' \neq i} \log \left(1 - \frac{1}{M-1} \left(1 - u^{\mathbf{1}_{\{U_{m',i'} < \sigma(X_{m',i'}^M)\}}} \right) \right)} \right. \\
&\quad \left. e^{\sum_{n' \neq n, j' \neq j} \log \left(1 - \frac{1}{M-1} \left(1 - v^{\mathbf{1}_{\{U_{n',j'} < \sigma(X_{n',j'}^M)\}}} \right) \right)} \right].
\end{aligned}$$

The fourth equality above comes from the independence between the routing variables R^M .

Just as in the proof of Lemma 2.2.10, we can give upper and lower bounds of the last right-hand side expression:

$$\mathbf{E} \left[u^{A_{m,i}^M} v^{A_{n,j}^M} \right] \leq \mathbf{E} \left[e^{\frac{-1}{M-1} \sum_{m' \neq m, i' \neq i} \left(2 - u^{\mathbf{1}_{\{U_{m',i'} < \sigma(X_{m',i'}^M)\}}} - v^{\mathbf{1}_{\{U_{m',i'} < \sigma(X_{m',i'}^M)\}}} \right)} \right]$$

and

$$\mathbf{E} \left[u^{A_{m,i}^M} v^{A_{n,j}^M} \right] \geq \mathbf{E} \left[e^{-\frac{1}{M-1} \sum_{m' \neq m, i' \neq i} \left(2^{-u} \mathbf{1}_{\{U_{m',i'} < \sigma(X_{m',i'}^M)\}} - v \mathbf{1}_{\{U_{m',i'} < \sigma(X_{m',i'}^M)\}} \right)} \cdot e^{-\frac{1}{2(M-1)^2} \sum_{m' \neq m, i' \neq i} \left(2^{-u} \mathbf{1}_{\{U_{m',i'} < \sigma(X_{m',i'}^M)\}} - v \mathbf{1}_{\{U_{m',i'} < \sigma(X_{m',i'}^M)\}} \right)^2} \right].$$

The last right-hand side expression goes to $e^{(1-u+1-v)(K-1)\theta}$ when $M \rightarrow \infty$ in both cases, as previously. The result follows from these two bounds as in the proof of Lemma 2.2.10.

The case where $m = n$, i.e., when we consider the arrivals to two different neurons in the same replica, is done in the same way since the routing variables are independent from the neurons chosen. The case where $i = j$, i.e. when we consider the arrivals to the same neuron in two different replicas, is treated in the same way, with the extra step of isolating the terms that are not independent from each other. \square

Step three: propagation of pairwise asymptotic independence

Our goal is now to combine the previous results to show that PAI(Y) holds, assuming PAI(X). We have that for all $i \in \{1, \dots, K\}$ and all $m \in \{1, \dots, M\}$, $Y_{m,i}^M = \hat{X}_{m,i}^M + A_{m,i}^M$. We call \tilde{A} the limit in distribution of $A_{m,i}^M$ (it is Poisson distributed by the previous lemma). It is clear that by exchangeability between replicas, we only require the following lemma:

Lemma 2.2.12. For $i, j \in \{1, \dots, K\}$,

$$\mathbf{E}[u^{Y_{1,i}^M}, v^{Y_{2,j}^M}] \rightarrow \mathbf{E}[u^{\tilde{Y}}] \mathbf{E}[v^{\tilde{Y}}] \quad (2.22)$$

when $M \rightarrow \infty$, where $\tilde{Y} = \tilde{X} + \tilde{A}$.

Proof. Let $u, v \in [0, 1]$. Then, given $i, j \in [1, K]$, with $i \neq j$,

$$\begin{aligned}
\mathbf{E} \left[u^{Y_{1,i}^M} v^{Y_{2,j}^M} \right] &= \mathbf{E} \left[u^{\hat{X}_{1,i}^M} v^{\hat{X}_{2,j}^M} u^{\sum_{m' \neq 1, i' \neq i} \mathbf{1}_{\{U_{m',i'} < \sigma(X_{m',i'}^M)\}} \mathbf{1}_{\{R_{(m',i')}^M \rightarrow i = 1\}}} \right. \\
&\quad \left. v^{\sum_{n' \neq 2, j' \neq j} \mathbf{1}_{\{U_{n',j'} < \sigma(X_{n',j'}^M)\}} \mathbf{1}_{\{R_{(n',j')}^M \rightarrow j = 2\}}} \right] \\
&= \mathbf{E} \left[u^{\hat{X}_{1,i}^M} v^{\hat{X}_{2,j}^M} \prod_{m' \neq 1, i' \neq i} u^{\mathbf{1}_{\{U_{m',i'} < \sigma(X_{m',i'}^M)\}} \mathbf{1}_{\{R_{(m',i')}^M \rightarrow i = 1\}}} \right. \\
&\quad \left. \prod_{n' \neq 2, j' \neq j} v^{\mathbf{1}_{\{U_{n',j'} < \sigma(X_{n',j'}^M)\}} \mathbf{1}_{\{R_{(n',j')}^M \rightarrow j = 2\}}} \right] \\
&= \mathbf{E} \left[\mathbf{E} \left[u^{\hat{X}_{1,i}^M} v^{\hat{X}_{2,j}^M} \prod_{m' \neq 1, i' \neq i} u^{\mathbf{1}_{\{U_{m',i'} < \sigma(X_{m',i'}^M)\}} \mathbf{1}_{\{R_{(m',i')}^M \rightarrow i = 1\}}} \right. \right. \\
&\quad \left. \left. \prod_{n' \neq 2, j' \neq j} v^{\mathbf{1}_{\{U_{n',j'} < \sigma(X_{n',j'}^M)\}} \mathbf{1}_{\{R_{(n',j')}^M \rightarrow j = 2\}}} \middle| X^M, U \right] \right] \\
&= \mathbf{E} \left[u^{\hat{X}_{1,i}^M} v^{\hat{X}_{2,j}^M} \prod_{m' \neq 1, i' \neq i} \left(\frac{1}{M-1} u^{\mathbf{1}_{\{U_{m',i'} < \sigma(X_{m',i'}^M)\}} + \left(1 - \frac{1}{M-1} \right)} \right) \right. \\
&\quad \left. \prod_{n' \neq 2, j' \neq j} \left(\frac{1}{M-1} v^{\mathbf{1}_{\{U_{n',j'} < \sigma(X_{n',j'}^M)\}} + \left(1 - \frac{1}{M-1} \right)} \right) \right] \\
&= \mathbf{E} \left[\varphi_1^M(u, v) \varphi_2^M(u, v) \right],
\end{aligned}$$

where

$$\begin{aligned}
\varphi_1^M(u, v) &= u^{\hat{X}_{1,i}^M} \left(1 - \frac{1}{M-1} + \frac{1}{M-1} v^{\mathbf{1}_{\{U_{1,i} < \sigma(X_{1,i}^M)\}}} \right) \\
&\quad v^{\hat{X}_{2,j}^M} \left(1 - \frac{1}{M-1} + \frac{1}{M-1} u^{\mathbf{1}_{\{U_{2,j} < \sigma(X_{2,j}^M)\}}} \right)
\end{aligned}$$

and

$$\begin{aligned}
\varphi_2^M(u, v) &= e^{\sum_{m' \neq 1; i' \neq i; (m', i') \neq (2, j)} \log \left(1 - \frac{1}{M-1} \left(1 - u^{\mathbf{1}_{\{U_{m',i'} < \sigma(X_{m',i'}^M)\}}} \right) \right)} \\
&\quad e^{\sum_{n' \neq 2; j' \neq j; (n', j') \neq (1, i)} \log \left(1 - \frac{1}{M-1} \left(1 - v^{\mathbf{1}_{\{U_{n',j'} < \sigma(X_{n',j'}^M)\}}} \right) \right)}.
\end{aligned}$$

When $M \rightarrow \infty$, by Lemmas 2.2.9 and 2.2.11, $\varphi_1^M(u, v)$ and $\varphi_2^M(u, v)$ are pairwise asymptotically independent. Since in $\varphi_2^M(u, v)$, the contribution of the missing terms in the sum is negligible, when $M \rightarrow \infty$, we have

$$\mathbf{E} \left[\varphi_1^M(u, v) \varphi_2^M(u, v) \right] \rightarrow \mathbf{E} \left[u^{\tilde{X}} \right] \mathbf{E} \left[v^{\tilde{X}} \right] \mathbf{E} \left[u^{\tilde{A}} \right] \mathbf{E} \left[v^{\tilde{A}} \right]. \quad (2.23)$$

This shows that (2.22) holds. \square

Thus, $\text{PAI}(X)$ implies $\text{PAI}(Y)$, which concludes the proof of the theorem. Note that Lemma 2.2.12 also shows that \tilde{X} and \tilde{A} are independent.

2.2.4 The symmetric fragmentation-interaction-aggregation process

Our goal is to show that propagation of chaos and the Poisson Hypothesis hold in the more general setting of symmetric FIAPs under mild hypotheses on the dynamics of the system. The symmetrical evolution equations read

$$Y_i = g_1(X_i) \mathbb{1}_{\{U_i < \sigma(X_i)\}} + g_2(X_i) \mathbb{1}_{\{U_i > \sigma(X_i)\}} + A_i \quad (2.24)$$

where

$$A_i = \sum_{j \neq i} h(X_j) \mathbb{1}_{\{U_j < \sigma(X_j)\}} \quad (2.25)$$

and $g_1, g_2, h : \mathbb{N} \rightarrow \mathbb{N}$ are functions such that h is bounded.

We now introduce the corresponding replica dynamics. Let $\{X_{m,i}^M\}$ be the integer-valued state variables at step 0, where $m \in \{1, \dots, M\}$ and $i \in \{1, \dots, K\}$. Let $\{Y_{m,i}^M\}$ be the integer-valued state variables at time one. Let $\{U_{m,i}\}$ be i.i.d. random variables independent from $\{X_{m,i}^M\}$ uniformly distributed in $[0, 1]$. We introduce again the i.i.d. routing variables $R_{(n,j) \rightarrow i}^M$, independent from $\{U_{m,i}\}$ and $\{X_{m,i}^M\}$ and uniformly distributed in $\{1, \dots, M\} \setminus \{n\}$. The M -replica equations read:

$$Y_{m,i}^M = g_1(X_{m,i}^M) \mathbb{1}_{\{U_{m,i} < \sigma(X_{m,i}^M)\}} + g_2(X_{m,i}^M) \mathbb{1}_{\{U_{m,i} > \sigma(X_{m,i}^M)\}} + A_{m,i}^M, \quad (2.26)$$

where

$$A_{m,i}^M = \sum_{n \neq m} \sum_{j \neq i} h(X_{n,j}^M) \mathbb{1}_{\{U_{n,j} < \sigma(X_{n,j}^M)\}} \mathbb{1}_{\{R_{(n,j) \rightarrow i}^M = m\}} \quad (2.27)$$

is the number of arrivals in node i of replica m .

We also recall the definition of a compound Poisson distribution:

Definition 2.2.13. *The random variable X is said to follow a compound Poisson distribution if there exist a Poisson(λ) random variable N and i.i.d. random variables $(X_i)_{i \in \mathbb{N}^*}$ independent from N such that $X = \sum_{i=1}^N X_i$. The generating function of X , denoted φ_X , is given by*

$$\varphi_X(t) = e^{\lambda(\varphi(t)-1)}, \quad (2.28)$$

where $\varphi(t)$ is the generating function of X_1 .

We have the following theorem:

Theorem 2.2.14. *For all symmetric RMF FIAP dynamics, $\text{PAI}(X)$ implies $\text{PAI}(Y)$. Moreover, the arrivals to a given node are asymptotically compound Poisson distributed.*

We will require the following lemmas. The following result replaces Lemma 2.2.9:

Lemma 2.2.15. *Let $\hat{X}^1 = \{\hat{X}_{m,i}^{1,M} = g_1(X_{m,i}^M) \mathbb{1}_{\{U_{m,i} < \sigma(X_{m,i}^M)\}}\}$. Let $\hat{X}^2 = \{\hat{X}_{m,i}^{2,M} = g_2(X_{m,i}^M) \mathbb{1}_{\{U_{m,i} > \sigma(X_{m,i}^M)\}}\}$. Then $\text{PAI}(X)$ implies $\text{PAI}(\hat{X}^1)$, $\text{PAI}(\hat{X}^2)$ and $\text{PAI}(\hat{X})$, where $\hat{X} = \hat{X}^1 + \hat{X}^2$.*

Proof. We proceed exactly as in Lemma 2.2.9. We write here only the proof for \hat{X}^2 , the others being identical except for the numerical expressions involved. We have for $u, v \in [0, 1]$,

$$\mathbf{E}[u^{\hat{X}_1^{2,M}} v^{\hat{X}_2^{2,M}}] = \sum_{k,l \in \mathbb{N}} \mathbf{P}(\hat{X}_1^{2,M} = k, \hat{X}_2^{2,M} = l) u^k v^l. \quad (2.29)$$

For $k, l > 0$, we have

$$\begin{aligned} \mathbf{P}(\hat{X}_1^{2,M} = k, \hat{X}_2^{2,M} = l) &= \sum_{p,q \in \mathbb{N}} \mathbf{P}(g_2(X_1^M) = k, g_2(X_2^M) = l, X_1^M = p, X_2^M = q) \\ &\quad (1 - \sigma(p))(1 - \sigma(q)). \end{aligned} \quad (2.30)$$

Since $\text{PAI}(X)$ holds,

$\mathbf{P}(g_2(p) = k, g_2(q) = l, X_1^M = p, X_2^M = q) \rightarrow \mathbf{P}(\tilde{X} = p, g_2(p) = k) \mathbf{P}(\tilde{X} = q, g_2(q) = l)$ when $M \rightarrow \infty$. Hence, $\mathbf{E}[u^{\hat{X}_1^{2,M}} v^{\hat{X}_2^{2,M}}] \rightarrow \mathbf{E}[u^{g_2(\tilde{X})} \mathbf{1}_{\{U < \sigma(\tilde{X})\}}] \mathbf{E}[v^{g_2(\tilde{X})} \mathbf{1}_{\{U < \sigma(\tilde{X})\}}]$ when $M \rightarrow \infty$. The cases where k and/or l are equal to 0 are handled in the same way. This proves the result. \square

We now prove the following result, which replaces Lemma 2.2.10:

Lemma 2.2.16. *Supposing that $\text{PAI}(X)$ holds, when $M \rightarrow \infty$, we have the convergence in distribution $A_{m,i}^M \rightarrow \tilde{A}$, where \tilde{A} follows a compound Poisson distribution.*

Proof. We still have, just like in the proof of Lemma 2.2.10, that for $z \in [0, 1]$, $i \in \{1, \dots, K\}$, $m \in \{1, \dots, M\}$,

$$\mathbf{E}[z^{A_{m,i}^M}] = \mathbf{E} \left[e^{\sum_{n \neq m} \sum_{j \neq i} \log \left(1 - \frac{1}{M-1} \left(1 - z^{h(X_{n,j}^M) \mathbf{1}_{\{U_{n,j} < \sigma(X_{n,j}^M)\}}} \right) \right)} \right]. \quad (2.31)$$

Using the same arguments as before, we have when $M \rightarrow \infty$

$$\mathbf{E} \left[z^{A_{m,i}^M} \right] \rightarrow e^{(K-1)(\Phi(z)-1)}, \quad (2.32)$$

where $\Phi(z) = \mathbf{E} \left[z^{h(\tilde{X})} \mathbf{1}_{\{U < \sigma(\tilde{X})\}} \right]$, which is precisely of the form (2.28), that is, a generating function of a random variable with a compound Poisson distribution. \square

We now combine these results to prove Theorem 2.2.14.

Proof. We follow the outline of the previous section. Lemmas 2.2.5, 2.2.6 and 2.2.7 still apply as previously. Lemma 2.2.15 replaces Lemma 2.2.9. Lemma 2.2.16 replaces Lemma 2.2.10. Lemmas 2.2.11 and 2.2.12 still hold, with only differences in the limiting expressions. \square

2.2.5 The general fragmentation-interaction-aggregation process

The previously introduced exchangeable dynamics allow for simpler computations at the expense of realistic modeling. For example, neuron populations are not homogeneous and are not fully connected. In order to account for such a geometry, we now generalize the previous result to the case where the functions governing the information received by a node when another node activates depend on the nodes involved. Specifically, recall the class \mathcal{C} of discrete FIAP defined in Section 1.

For any process in \mathcal{C} , we can define a replica mean field model as in the previous sections: we consider a collection of M identically distributed replicas of a set of K nodes, which could be neurons, particles, queues or other objects, depending on context. As previously, let $\{X_{m,i}^M\}$ be the integer-valued state variables at step 0, where $m \in \{1, \dots, M\}$ and $i \in \{1, \dots, K\}$. Let $\{Y_{m,i}^M\}$ be the integer-valued state variables at time one. Let $\{U_{m,i}\}$ be uniformly distributed on $[0, 1]$ i.i.d. random variables independent from $\{X_{m,i}^M\}$. Let $\{R_{(n,j) \rightarrow i}^M\}$ be i.i.d. routing random variables independent from $\{X_{m,i}^M\}$ and $\{U_{m,i}\}$, uniformly distributed on $\{1, \dots, M\} \setminus \{n\}$ for all $i, j \in \{1, \dots, K\}$ and $n \in \{1, \dots, M\}$. Recall that the M -RMF equations read

$$Y_{m,i}^M = g_{1,i}(X_{m,i}^M) \mathbb{1}_{\{U_{m,i} < \sigma_i(X_{m,i}^M)\}} + g_{2,i}(X_{m,i}^M) \mathbb{1}_{\{U_{m,i} > \sigma_i(X_{m,i}^M)\}} + A_{m,i}^M, \quad (2.33)$$

where

$$A_{m,i}^M = \sum_{n \neq m} \sum_{j \neq i} h_{j \rightarrow i}(X_{n,j}^M) \mathbb{1}_{\{U_{n,j} < \sigma_j(X_{n,j}^M)\}} \mathbb{1}_{\{R_{(n,j) \rightarrow i}^M = m\}} \quad (2.34)$$

is the number of arrivals in node i of replica m . We now show that the result from the previous section carries over to this more general setting with only minor modifications.

First, we must slightly modify the definition of pairwise asymptotic independence in order to take into account the dependence on the node of the limiting distribution. As a simplification, we keep the same notations for this modified definition.

Definition 2.2.17. *Given $M \in \mathbb{N}$, given an array of integer-valued random variables $Z = \{Z_{m,i}^M\}_{1 \leq m \leq M, 1 \leq i \leq K}$ such that for all fixed M , the random variables $Z_{m,i}^M$ are exchangeable in m , we say that the variables $Z_{m,i}^M$ are pairwise asymptotically independent, which we will denote $\text{PAI}(Z)$, if there exist integer-valued random variables $(\tilde{Z}_i)_{i \in \{1, \dots, K\}}$ such that $\forall (m, i) \neq (n, j), \forall u, v \in [0, 1]$,*

$$\lim_{M \rightarrow \infty} \mathbf{E}[u^{Z_{m,i}^M} v^{Z_{n,j}^M}] = \mathbf{E}[u^{\tilde{Z}_i}] \mathbf{E}[v^{\tilde{Z}_j}]. \quad (2.35)$$

For clarity of exposition, we also recall here the definition of the triangular law of large numbers, even though it is left unchanged:

Definition 2.2.18. *Given $M \in \mathbb{N}$, given an array of integer-valued random variables $Z = \{Z_m^M\}_{m \in \{1, \dots, M\}}$ such that for all fixed M , the random variables Z_m^M are exchangeable in m , we say that Z verifies the triangular law of large numbers $\text{TLLN}(Z)$ if there exist an integer-valued random variable \tilde{Z} such that for all functions $f : \mathbb{N} \rightarrow \mathbb{R}$ with compact support, we have the following limit in L^2 :*

$$\lim_{M \rightarrow \infty} \frac{1}{M} \sum_{m=1}^M f(Z_m^M) = \mathbf{E}[f(\tilde{Z})]. \quad (2.36)$$

Then, we obtain the same result:

Theorem 2.2.19. *Using previously defined notations, $\text{PAI}(X)$ implies $\text{PAI}(Y)$. Moreover, the arrivals to a given node are asymptotically compound Poisson distributed and are independent of the states of the nodes.*

We once again require the following lemmas for the proof.

We replace Lemma 2.2.7 with the following similar result, taking into account the fact that the limiting distribution now depends on the node:

Lemma 2.2.20. *Let $M \in \mathbb{N}$, let $Z = \{Z_{m,i}^M\}_{m \in \{1, \dots, M\}, i \in \{1, \dots, K\}}$ be an array of integer valued random variables verifying $\text{PAI}(Z)$. Then for all bounded functions $f : \mathbb{N} \times [0, 1] \rightarrow \mathbb{R}$ with compact support, for all i.i.d. sequences of random variables $U = \{U_{m,i}\}_{m \in \{1, \dots, M\}, i \in \{1, \dots, K\}}$ independent from Z , there exists U independent from*

$(\tilde{Z}_i)_{i \in \{1, \dots, K\}}$ and Z such that for all $i \in \{1, \dots, K\}$, we have the following limit in L^2 :

$$\lim_{M \rightarrow \infty} \frac{1}{M} \sum_{m=1}^M f(Z_{m,i}^M, U_{m,i}) = \mathbf{E}[f(\tilde{Z}_i, U)]. \quad (2.37)$$

The proof is exactly the same as for Lemma 2.2.6.

We must replace Lemma 2.2.10 with the following result:

Lemma 2.2.21. *Supposing that PAI(X) holds, when $M \rightarrow \infty$: $A_{m,i}^M \rightarrow \tilde{A}_i$ in distribution, where \tilde{A}_i follows a compound Poisson distribution.*

Proof. We have for $z \in [0, 1]$, $i \in \{1, \dots, K\}$, $m \in \{1, \dots, M\}$, that

$$\mathbf{E} \left[z^{A_{m,i}^M} \right] = \mathbf{E} \left[e^{\sum_{n \neq m} \sum_{j \neq i} \log \left(1 - \frac{1}{M-1} \left(1 - z^{h_{j \rightarrow i}(X_{n,j}^M)} \mathbf{1}_{\{U_{n,j} < \sigma_j(X_{n,j}^M)\}} \right) \right)} \right]. \quad (2.38)$$

Using the same arguments as before, we have when $M \rightarrow \infty$

$$\mathbf{E} \left[z^{A_{m,i}^M} \right] \rightarrow e^{\Phi_i(z)}, \quad (2.39)$$

where $\Phi_i(z) = - \sum_{j \neq i} \mathbf{E} \left[1 - z^{h_{j \rightarrow i}(\tilde{X}_i)} \mathbf{1}_{\{U < \sigma_j(\tilde{X}_i)\}} \right]$. Therefore,

$$\mathbf{E} \left[z^{A_{m,i}^M} \right] \rightarrow e^{- \sum_{j \neq i} \left(1 - \mathbf{E} \left[z^{h_{j \rightarrow i}(\tilde{X}_i)} \mathbf{1}_{\{U < \sigma_j(\tilde{X}_i)\}} \right] \right)}. \quad (2.40)$$

The expression is of the form (2.28), which proves Lemma 2.2.21. \square

We now prove Theorem 2.2.19.

Proof. We use the same reasoning as previously.

Lemma 2.2.6 still holds (the replicas are still exchangeable, only the nodes are not). Lemma 2.2.20 replaces Lemma 2.2.7. Since the functions $g_{1,i}$ and $g_{2,i}$ only depend on the node and not on the replica index, an equivalent result to Lemma 2.2.15 still holds. Lemma 2.2.10 is replaced by Lemma 2.2.21. For asymptotic independence, we have, using the same arguments as in the proof of Lemma 2.2.11, that for $u, v \in [0, 1]$, for $m \neq n$ and $i \neq j$,

$$\mathbf{E} \left[u^{A_{m,i}^M} v^{A_{n,j}^M} \right] \rightarrow e^{- \sum_{i' \neq i} \left(1 - \mathbf{E} \left[u^{h_{i,i'}(\tilde{X}_i)} \mathbf{1}_{\{U < \sigma_j(\tilde{X}_i)\}} \right] \right) - \sum_{j' \neq j} \left(1 - \mathbf{E} \left[v^{h_{j,j'}(\tilde{X}_j)} \mathbf{1}_{\{U < \sigma_i(\tilde{X}_j)\}} \right] \right)}, \quad (2.41)$$

when $M \rightarrow \infty$. The other cases ($m = n$ and $i = j$) are also valid. Lemma 2.2.12 also still holds, with only minor differences in the limit expressions. \square

Note that once again, this proves that the limit processes \tilde{X}_i and \tilde{A}_j are independent for all $i, j \in \{1, \dots, K\}$.

As an application, let us apply this result to the model from Section 2 with the addition of nonexchangeable interactions. Namely, we consider $h_{j \rightarrow i}(X_{m,j}^M) = \mu_{j \rightarrow i}$ with $\mu_{j \rightarrow i} \in \mathbb{N}$ (potentially zero). In this case, Theorem 2.2.19 proves the propagation of chaos in this system, and the limit distributions of arrivals at the different nodes are characterized by, for $i \in \{1, \dots, K\}$ and $z \in [0, 1]$,

$$\mathbf{E} \left[z^{\tilde{A}_i} \right] = e^{\theta_i \sum_{j \neq i} (z^{\mu_{j \rightarrow i}} - 1)} = \prod_{j \neq i} e^{\theta_i (z^{\mu_{j \rightarrow i}} - 1)}, \quad (2.42)$$

where $\theta_i = \mathbf{E} \left[\sigma_i(\tilde{X}_i) \right]$. Note that as expected, when all $\mu_{j \rightarrow i}$ are equal to one, we obtain the result from Section 2.

2.2.6 Extensions

There are several ways of extending the FIAP framework while preserving the basic properties proved in the present paper (propagation of chaos and Poisson Hypothesis). We decided not to include them in the general framework in order to keep notation and exposition light. A few natural extensions of this type are nevertheless discussed below.

Random Interactions The functions $h_{j \rightarrow i}(k)$ can be replaced by randomized functions of the type $h_{j \rightarrow i}(k, V_{i,j})$ where the random variables $\{V_{i,j}\}_{1 \leq i, j \leq K}$ are uniform in $[0, 1]$ and i.i.d.. This allows one to represent, e.g., the queuing theory scenario where a customer leaving a queue is randomly routed to an other queue of the network according to some stochastic routing matrix $\{p_{j,i}\}_{1 \leq i, j \leq K}$, namely a customer leaving queue i is routed to queue j with probability $p_{j,i}$. If the random variables $\{V_{i,j}\}$ are independent of $\{X_i\}_i$, then the main results still hold.

Time in-homogeneous dynamics The general setting of the paper implicitly suggests to use the same (activation, fragmentation, and interaction) functions at all time steps for a given node. There is no difficulty extending the results to the time in-homogeneous case where these functions depend on the time step. In the neural network case, this for instance happens in certain learning dynamics where the synaptic weights evolve over time.

Exogenous input and output To the *endogenous* arrivals A_i to node i given in Equation (2.3), we add arbitrary *exogenous* arrivals $B_i \in \mathbb{N}$. In the special case where the random variables $\{B_i\}$ are independent, Poisson, and independent of

the state variables $\{X_i\}_i$ then the same results still hold. Note that one can also define an *exogenous output* for node i through the relation

$$D_i = h_{o,i}(X_i) \mathbb{1}_{\{U_i < \sigma_i(X_i)\}},$$

where $h_{o,i}$ is a given output function $\mathbb{N} \rightarrow \mathbb{N}$.

Exogenous input point processes are needed for modeling reasons in a variety of contexts (e.g., to represent requests from end-users in a computer network, or input signals from sensors in a neural network). Exogenous output processes are useful in, e.g., a two-layer network where the first layer feeds the second one, but not conversely. Exogenous input and output point processes are instrumental in the partition scheme of the vector state example discussed below. In that example, the exogenous input variables are neither necessarily Poisson nor independent.

Vector State - Example This extension is first described through a simple neural network example. We partition the set of neurons of a discrete Galves-Löcherbach network in pairs (this assumes that K is even). Each pair of the partition is a node of the network. If (i, j) is one of these nodes, it has a two-dimensional vector state (X_i, X_j) (rather than a one dimensional state in the initial model). We let this pair (as well as each other pair in the partition) evolve as a two-node Galves-Löcherbach network with some *vector* exogenous input and output. More precisely, conditionally on $(X_i, X_j) = (k, l)$, neurons i and j spike independently with probability $\sigma_i(k)$ and $\sigma_j(l)$ respectively. If none of them spikes, the state (Z_i, Z_j) after its endogenous evolution is still (X_i, X_j) . If only i (resp. j) spikes and the other neuron of the pair does not spike, then (Z_i, Z_j) is equal $(0, X_j + r_{j,i})$ (resp. $(X_i + r_{i,j}, 0)$). If both spike, then $(Z_i, Z_j) = (r_{i,j}, r_{j,i})$, with $r_{k,l}$ equal to 1 if there is a directed edge from l to k and 0 otherwise. Therefore, if (B_i, B_j) denotes the vector exogenous input, the state of this pair at time one is $(Z_i + B_i, Z_j + B_j)$, by combining the endogenous evolution and the exogenous arrivals.

Define now the *exogenous output of type $k \notin \{i, j\}$ of node (i, j)* by

$$D_k(i, j) = \mathbb{1}_{\{U_i < \sigma_i(X_i)\}} r_{k,i} + \mathbb{1}_{\{U_j < \sigma_j(X_j)\}} r_{k,j}, \quad (2.43)$$

The extension of interest here is that where we take the following exogenous input to node (i, j) :

$$B_i = \sum_{k \notin \{i, j\}} D_i(k, l(k)) \mathbb{1}_{\{k < l(k)\}}, \quad B_j = \sum_{k \notin \{i, j\}} D_j(k, l(k)) \mathbb{1}_{\{k < l(k)\}}, \quad (2.44)$$

with $l(k)$ the neuron paired with k .

Note that with this definition, when neuron i spikes, the effect on pair (k, l) with $l = l(k)$ is as follows: no effect if $r_{k,i} = r_{l,i} = 0$; one arrival in k and none in

l (resp. one in l and none in k) if $r_{k,i} = 1$ and $r_{l,i} = 0$ (resp. $r_{l,i} = 1$ and $r_{k,i} = 0$); a simultaneous arrival in both l and k otherwise. This defines a network which does not belong to the FIAP class, were it only because the state is now a vector. The M -RMF model features M replicas of this network with $K/2$ (vector state) nodes each. In this M -RMF model, the exogenous output of node/pair (i, j) in replica m is randomly sent to a replica chosen at random. More precisely, for all exogenous output type k paired with l ,

$$D_k^m(i, j) = \mathbb{1}_{\{U_i^m < \sigma_i(X_i^m)\}} r_{k,i} + \mathbb{1}_{\{U_j^m < \sigma_j(X_j^m)\}} r_{k,i},$$

$$\text{(resp. } D_l^m(i, j) = \mathbb{1}_{\{U_i^m < \sigma_i(X_i^m)\}} r_{l,i} + \mathbb{1}_{\{U_j^m < \sigma_j(X_j^m)\}} r_{l,i}\text{)}$$

units are sent to k (resp. l) of another replica selected uniformly at random where they are aggregated to the coordinates of the state variable of this pair. It can be shown that when M tends to infinity, (1) the random state vectors (X_i^m, X_j^m) and $(X_{i'}^m, X_{j'}^m)$, where (i, j) and (i', j') are two different pairs, are asymptotically independent (although the two coordinates of each vector are in general dependent); (2) the exogenous arrivals to any coordinate of a pair in a typical replica tends to an independent compound Poisson variable.

Vector State - General Case Consider a FIAP F with K nodes. Let S_1, S_2, \dots, S_l be a partition of $[1, \dots, K]$. Let $K_p, 1 \leq p \leq l$ denote the cardinality of set S_p , and let F_p be the restriction of F to the coordinates of S_p . Let \mathfrak{F}_p be the FIAP combining the endogenous dynamics of F_p and exogenous input $(\mathfrak{B}_{p,i}, i \in S_p)$. Let $\mathfrak{X}_{p,i}$ denote the state variables in \mathfrak{F}_p . For all $k \notin S_p$, define the *exogenous output* of type k of \mathfrak{F}_p as

$$\mathfrak{D}_p(k) = \sum_{i \in S_p} \mathbb{1}_{\{U_{p,i} < \sigma_i(\mathfrak{X}_{p,i})\}} h_{k,i}(\mathfrak{X}_{p,i}). \quad (2.45)$$

Note that $\mathfrak{D}_p(k)$ is also what coordinate k receives as exogenous input from S_p . That is, if we take

$$\mathfrak{B}_{p,i} = \sum_{q \neq p} \mathfrak{D}_q(i), \quad i \in S_p, \quad (2.46)$$

we get another (more complex) representation of the dynamics of F based on the point processes describing the interactions between the sets of the partition. The M -RMF model associated with this partition features M replicas of this network with q (vector state) nodes each. In this M -RMF model, the *exogenous* output of node $i \in S_p$ of replica m is randomly sent to replicas chosen at random. More precisely, for all $q \neq p$, the vector $(\mathfrak{D}_p^m(k), k \in S_q)$, with $\mathfrak{D}_p^m(k)$ defined as in (2.45), is sent to one replica chosen at random, and this is done independently for all $q \neq p$. This in turn defines new exogenous input point processes $\mathfrak{B}_{p,i}^m$ as in

(2.46). Let $(\mathfrak{X}_{p,i}^m, i \in S_p, p = 1, \dots, l, m = 1, \dots, M)$ denote the state variables in this M -RMF model. It can be shown that, when M tends to infinity,

1. for all $p \neq q$, the random state vectors $(\mathfrak{X}_{p,i}^m, i \in S_p)$ and $(\mathfrak{X}_{q,j}^m, j \in S_q)$, are asymptotically independent (although the coordinates of each vector are in general dependent);
2. for all p , and for all m , the exogenous arrivals $(\mathfrak{B}_{p,i}^m, i \in S_p)$ to set S_p tend to an independent multivariate compound Poisson variable with multivariate generating function

$$\exp \left(\sum_{q \neq p} \sum_{n_i \in \mathbb{N}, i \in S_q} \sum_{s \subset S_q} \pi_{q,s,(n_i)} \left(1 - \prod_{i \in s} \prod_{k \in S_q} z_k^{h_{k,i}(n_i)} \right) \right).$$

In this last equation,

$$\pi_{q,s,(n_i)} = \mathbf{P}[\tilde{\mathfrak{X}}_{q,i} = n_i, i \in S_q] \prod_{j \in s} \sigma_j(n_j) \prod_{j' \in S_q \setminus s} (1 - \sigma_{j'}(n_{j'})),$$

where $(\tilde{\mathfrak{X}}_{q,i})$ denotes random variables with the limiting joint distribution assumed in the vector generalization of PAI.

Conclusion

A new class of discrete time dynamics involving point process based interactions between interconnected nodes was introduced. The Poisson Hypothesis was proved for the RMF version of such dynamics. The proof is based on the property of pairwise asymptotic independence between replicas and is by induction over time. The key point is that randomized routing decisions on exchangeable events which are asymptotically independent lead to Poisson point processes. As for future research, a natural question is whether these results extend to continuous-time and continuous-space versions of FIAP dynamics. The extension to continuous-time FIAPs will follow from similar arguments as presented here under the condition that the replica-limit/time-limit diagram commutes for FIAPs. The extension to continuous-space FIAPs appears to require distinct analytical tools than those presented here. Finally, another question of interest is whether the Poisson Hypothesis can be shown for the RMF limits of other classes of systems besides FIAPs.

Chapter 3

Replica-mean-field limits for continuous-time processes

3.1 Introduction

This chapter consists of two articles, [27] and [28]. The first has been accepted in *Annals of Applied Probability* and both are available on ArXiv. They can be seen as companion papers, as the second aims to generalize the results proven in the first in a wider framework. Once again, notation has been slightly modified for harmonization purposes.

3.1.1 Replica-mean-field limits for intensity-based neural networks: abstract

Abstract

Neural computations arising from myriads of interactions between spiking neurons can be modeled as network dynamics with punctuate interactions. However, most relevant dynamics do not allow for computational tractability. To circumvent this difficulty, the Poisson Hypothesis regime replaces interaction times between neurons by Poisson processes. We prove that the Poisson Hypothesis holds at the limit of an infinite number of replicas in the replica-mean-field model, which consists of randomly interacting copies of the network of interest. The proof is obtained through a novel application of the Chen-Stein method to the case of a random sum of Bernoulli random variables and a fixed point approach to prove a law of large numbers for exchangeable random variables.

3.1.2 Replica-mean-field limit of continuous-time fragmentation-interaction-aggregation processes: abstract

Abstract

Many phenomena can be modeled as network dynamics with punctuate interactions. However, most relevant dynamics do not allow for computational tractability. To circumvent this difficulty, the Poisson Hypothesis regime replaces interaction times between nodes by independent Poisson processes, allowing for tractability in several cases, such as intensity-based models from computational neuroscience. This hypothesis is usually only conjectured, or numerically validated. In this work, we introduce a class of processes in continuous time called continuous fragmentation-interaction-aggregation processes, by analogy with previously introduced processes in discrete time. The state of each node, described by the stochastic intensity of an associated point process, aggregates arrivals from its neighbors and is fragmented upon departure. We consider the replica-mean-field version of such a process, which is a physical system consisting of randomly interacting copies of the network of interest. Generalizing results proved in discrete time and in the particular case of excitatory intensity-based neural dynamics, we prove that the Poisson Hypothesis holds at the limit of an infinite number of replicas.

3.2 Replica-mean-field limits for intensity-based neural networks

3.2.1 Introduction

Many phenomena in a variety of fields can be modeled as punctuate interactions between agents. Whether it be opinion dynamics [4], epidemics propagation [65], flow control on the internet [7] or neural computations [73], an agent-based approach is a versatile way to describe the phenomenon of interest through the behavior of each agent.

In such an approach, each agent is seen as a node in a network in which edges represent the possibility of interactions, and point processes associated to each node register the times at which these interactions happen. These point processes idealize the stochasticity inherent in the phenomena of interest. The state of the system can then be given by a set of stochastic differential equations, each describing the state of an agent. In intensity-based models, used extensively for example

in computational neuroscience [77], [68], this state is given by the stochastic intensity of the point process associated with the agent. Neuronal stochastic intensities model the instantaneous firing rate of a neuron as a function of the spiking inputs received from other neurons.

However, the versatility of this "microscopic" approach comes at a cost, namely, that of computational tractability. Indeed, except for the simplest network architectures, such as systems of 1 or 2 agents, an analytical expression characterizing the law of the GL model in the stationary regime is not in general available. To go beyond numerical simulations, it then becomes imperative to resort to some simplifying assumption.

As the complexity of the dynamics resides in the dependencies between agents due to interactions, it is natural to choose a simplified model in which the agents are considered independent. One such classical approach is called the mean-field regime. Introduced originally by McKean [60] and developed, among others, by Dobrushin[31] and Sznitman[75], it consists in approximating the interactions received by any one particle by an empirical mean of the interactions over the whole network. The mean-field regime arises at the limit with infinitely many agents, as the empirical mean typically converges to an expectation and gives rise to a nonlinear ordinary differential equation describing the behavior of an agent at a macroscopic level and allowing for tractability. This convergence, when it takes place, is linked to the concept of propagation of chaos, mainly in reference to the asymptotic independence between agents that arises at the limit.

In classical mean-field models such as [34] or [56], the network considered must be assumed fully connected, the effect of interactions on the state of a given agent must be small, typically inversely proportional to the number of agents, in order to prevent explosion in finite time in the system. These assumptions represent significant constraints on the architectures and sizes of the networks and thus on the types of phenomena for which a mean-field approximation is relevant.

To circumvent these limitations, different approaches have been explored in recent years. To apply mean-field approximation to small-sized networks (with less than 100 agents, for example), the refined-mean-field approach [36], [3] adds a corrective term to the macroscopic ODE. Mean-field models have also been studied in other scalings, for example diffusive, where the effect of interactions on a particle is inversely proportional to the square root of the number of agents in the system. So-called conditional propagation of chaos properties have been proven in that setting [32]. To incorporate heterogeneity, the properties of graphons (large dense graphs) have been used to derive new limit equations [1]. In this setting, the limit object is an infinite system of ODEs. However, this approach is only valid for dense networks; when the average degree of a node is of order lesser than the amount of nodes in the network, as is the case for example in the human brain, graphon theory

does not apply. Another approach to incorporate heterogeneity circumvents mean-field models altogether, relying instead on conditional independence properties and local weak limits to obtain local convergence [55].

In this work, we are interested in another tractable regime for agent-based models: the Poisson Hypothesis. First formulated by Kleinrock for large queueing systems [49], it states that the flow of arrivals to a given node can be approximated by a Poisson flow with rate equal to the average rate of the original flow of arrivals. In agent-based models, the flow of arrivals corresponds to the effect of interactions on a given node. Under the Poisson Hypothesis, the behavior of each agent is still described by a stochastic differential equation, but the agents are considered independent and interaction times are replaced by Poisson clocks, which allows for tractability. This regime has been studied for queueing models by Rybko, Shlosman and others [78] and by Baccelli and Taillefumier for intensity-based models from computational neuroscience [8].

Hereafter, we focus on the continuous-time Galves-Löcherbach model introduced in [35] and studied under the Poisson Hypothesis in [8]. We introduce a physical system, called the replica-mean-field, derived from the initial model that converges under a certain scaling to the dynamics under the Poisson Hypothesis. The replica-mean-field was first introduced by Dobrushin to study queueing models [79], and adapted to a network setting in [8]. However, in their work, the convergence of the replica-mean-field dynamics to dynamics under the Poisson Hypothesis is only assumed, and not proven. The crux of this article is proving that a propagation of chaos-type convergence does take place for the replica-mean-field model derived from the Galves-Löcherbach model.

In the recent work [27], we have introduced a class of discrete-time processes on a discrete space, called fragmentation-interaction-aggregation processes (FIAPs), that include among others discrete versions of the Galves-Löcherbach model, and we have proven the propagation of chaos property for a replica-mean-field model associated with a FIAP for any finite time. Our aim is to generalize these results to a model in continuous time and with a continuous state space.

Structure of the paper

After this general introduction, we formally define all the models that we will be considering, namely, the Galves-Löcherbach model, its replica-mean-field version and its dynamics under the Poisson Hypothesis. We then state the main result of the paper, namely propagation of chaos in the replica-mean-field model on compacts of time, which we then prove in Section 3.2.2. Finally, in Section 3.2.3, we generalize the main result to weak convergence on the half-line through a tightness argument.

The continuous-time Galves-Löcherbach (GL) model

Let us formally present the Galves-Löcherbach model mentioned above: we consider a finite collection of K neurons whose spiking activities are given by the realization of a system of simple point processes without any common points $\mathbf{N} = \{N_i\}_{1 \leq i \leq K}$ on \mathbb{R} defined on some measurable space (Ω, \mathcal{F}) . For each neuron $1 \leq i \leq K$, we denote by $(T_{i,n})_{n \in \mathbb{Z}}$ the sequence of successive spiking times with the convention that, almost surely, $T_{i,0} \leq 0 < T_{i,1}$ and $T_{i,n} < T_{i,n+1}$ for all n .

To model the interactions due to spikes within the system, we consider that the spiking rate of neuron i depends on the times at which neuron i and the other neurons $j \neq i$ have spiked in the past. Formally, we introduce the network spiking history $(\mathcal{F}_t)_{t \in \mathbb{R}}$ as an increasing collection of σ -fields such that

$$\mathcal{F}_t^{\mathbf{N}} = \{\sigma(N_1(B_1), \dots, N_K(B_K)) | B_i \in \mathbf{B}(\mathbb{R}), B_i \subset (-\infty, t]\} \subset \mathcal{F}_t,$$

where $\mathcal{F}_t^{\mathbf{N}}$ is the internal history of the spiking process \mathbf{N} .

Recall that the \mathcal{F}_t -stochastic intensity $\{\lambda_i(t)\}_{t \in \mathbb{R}}$ of the associated point process N_i is the \mathcal{F}_t -predictable random process satisfying for all $s < t \in \mathbb{R}$:

$$\mathbf{E}[N_i(s, t) | \mathcal{F}_s] = \mathbf{E}\left[\int_s^t \lambda_i(u) du | \mathcal{F}_s\right], \quad (3.1)$$

where \mathcal{F}_t is the network history. We will hereafter refer to (3.1) as the stochastic intensity property.

We consider the \mathcal{F}_t -stochastic intensities $\lambda_1, \dots, \lambda_K$ associated with the point processes N_1, \dots, N_K .

In the Galves-Löcherbach model, the evolution in time of these intensities is given by the following system of stochastic differential equations:

$$\lambda_i(t) = \lambda_i(0) + \frac{1}{\tau_i} \int_0^t (b_i - \lambda_i(s)) ds + \sum_{j \neq i} \mu_{j \rightarrow i} \int_0^t N_j(ds) + \int_0^t (r_i - \lambda_i(s)) N_i(ds), \quad (3.2)$$

where $\tau_i, b_i, r_i > 0$ and $\mu_{j \rightarrow i} \geq 0$ are given constants and $\lambda_i(0)$ is assumed to be greater than r_i and b_i .

Let us make each term more explicit. The first integral term shows that without any spikes, the intensity exponentially decays to its base rate b_i with a relaxation time τ_i . The second integral term represents the aggregation of all the spikes received from the other neurons in the system: a spike received from neuron j causes a jump of $\mu_{j \rightarrow i}$ units in the intensity of neuron i . Finally, the third integral term is the reset that occurs when neuron i spikes: λ_i is then reset to r_i , which is a value such that $0 < r_i \leq b_i$. Taking $r_i \leq b_i$ models a refractory period that occurs after a spike during which the neuron enters a transient phase.

It has been shown in [8] that (3.2) defines a piecewise deterministic Markov process. When there is no exponential decay, that is, when $\tau_i = \infty$ for all i , the process becomes a pure jump Markov process that is Harris-ergodic and thus has a unique invariant measure.

It has also been shown that the moment generating function (MGF) at stationarity $u \rightarrow L(u) = \mathbf{E}[\exp(\sum_{i=1}^K u_i \lambda_i)]$ satisfies the following differential equation:

$$\left(\sum_i \frac{u_i b_i}{\tau_i} \right) L - \sum_i \left(1 + \frac{u_i}{\tau_i} \right) \partial_{u_i} L + \sum_i e^{(u_i r_i + \sum_{j \neq i} u_j \mu_{j \rightarrow i})} \partial_{u_i} L|_{u_i=0} = 0. \quad (3.3)$$

This equation is not solvable except for some very special cases ($K \leq 2$, for example). It is thus necessary to use approximating schemes or truncating moments, both of which neglect correlations due to the finite size of the network. Here, we introduce a different physical system that allows to obtain closed forms for equations similar to (3.3).

Replica-mean-fields of GL models

In replica-mean-field models, we consider M replicas of the initial set of K neurons. When neuron i in replica m spikes, for each neuron j that would receive something from the spike, a replica n is uniformly and independently chosen among the other $M - 1$ replicas, and neuron i sends $\mu_{i \rightarrow j}$ to it.

Formally, for $1 \leq m \leq M, 1 \leq i, j \leq K$, let $\{V_{(m,i) \rightarrow j}^M(t)\}_{t \in \mathbb{R}}$ be \mathcal{F}^N -predictable stochastic processes such that, for each spiking time T , i.e., each point of $N_{m,i}^M$, the random variables $\{V_{(m,i) \rightarrow j}^M(T)\}_j$ are mutually independent, independent from the past, and uniformly distributed on $\{1, \dots, M\} \setminus \{m\}$. Here, $V_{(m,i) \rightarrow j}^M(T)$ gives the index of the replica to which the spike of neuron i in replica m at time T is sent to neuron j .

The stochastic intensities associated with the point processes will then solve the following system of stochastic differential equations:

$$\begin{aligned} \lambda_{m,i}^M(t) &= \lambda_{m,i}^M(0) + \frac{1}{\tau_i} \int_0^t (b_i - \lambda_{m,i}^M(s)) ds \\ &+ \sum_{j \neq i} \mu_{j \rightarrow i} \sum_{n \neq m} \int_0^t \mathbb{I}_{\{V_{(n,j) \rightarrow i}^M(s)=m\}} N_{n,j}^M(ds) + \int_0^t (r_i - \lambda_{m,i}^M(s)) N_{m,i}^M(ds). \end{aligned} \quad (3.4)$$

These equations, which we will hereafter refer to as the RMF dynamics, characterize the dynamics of the M -replica system. As before, for all M , these dynamics

form a pure jump Harris-ergodic Markov process with a unique invariant measure since we are under Assumption 3.2.1.

The infinitesimal generator of the M -replica dynamics is given by

$$\begin{aligned} \mathcal{A}[f](\lambda) &= \sum_{i=1}^K \sum_{m=1}^M \left(\frac{b_i - \lambda_{m,i}}{\tau_i} \partial_{\lambda_{m,i}} f(\lambda) \right) \\ &+ \sum_{i=1}^K \sum_{m=1}^M \frac{1}{|\mathcal{V}_{m,i}^M|} \sum_{v \in \mathcal{V}_{m,i}^M} (f(\lambda + \mu_{m,i,v}(\lambda)) - f(\lambda)) \lambda_{m,i}, \end{aligned}$$

where the update due to the spiking of neuron (m, i) is defined by

$$[\mu_{m,i,v}(\lambda)]_{n,j} = \begin{cases} \mu_{j \rightarrow i} & \text{if } j \neq i, n = v_j \\ r_i - \lambda_{m,i} & \text{if } j = i, n = m \\ 0 & \text{otherwise .} \end{cases}$$

As before, we can establish an equation for the MGF $u \rightarrow L(u) = \mathbf{E}[e^{(\sum_{m=1}^M \sum_{i=1}^K u_{m,i} \lambda_{m,i})}]$ in the stationary regime:

$$\begin{aligned} &\left(\sum_m \sum_i \frac{u_{m,i} b_i}{\tau_i} \right) L - \sum_m \sum_i \left(1 + \frac{u_{m,i}}{\tau_i} \right) \partial_{u_{m,i}} L \\ &+ \sum_m \sum_i \frac{1}{(M-1)^K} \sum_{v \in \mathcal{V}_{m,i}^M} e^{(u_{m,i} r_i + \sum_{j \neq i} u_{v_j, j} \mu_{j \rightarrow i})} \partial_{u_{m,i}} L|_{u_i=0} = 0, \end{aligned}$$

where

$$\mathcal{V}_{m,i}^M = \{v \in \{1, \dots, M\}^K | v_i = m \text{ and } v_j \neq m, j \neq i\}.$$

Once again, this equation is not easily solvable. However, a closed form has been obtained in [8] by setting the Poisson Hypothesis, that is, by considering that at the limit when $M \rightarrow \infty$, the replicas become asymptotically independent and the arrivals process to a given replica becomes asymptotically Poisson.

This hypothesis is often conjectured or numerically validated and not proven, as was the case in [8]. The aim of this work is to give a proof of the Poisson Hypothesis in the RMF limit in continuous time, by analogy of the work done in discrete time in [27].

In the rest of the paper, we make the following assumptions:

Assumption 3.2.1. For all $i \in \{1, \dots, K\}$, $\tau_i = \infty$ (no exponential decay).

Assumption 3.2.2. There exists $\xi_0 > 0$ such that for all $1 \leq m \leq M, 1 \leq i \leq K$ and all $0 < \xi \leq \xi_0$, $\mathbf{E}[e^{\xi \lambda_{m,i}(0)}] < \infty$.

Assumption 3.2.1 is introduced mainly to simplify notation and reasonings. While we do not rigorously prove it, we strongly believe that all results remain valid without it. Assumption 3.2.2 restricts the initial conditions of the system to allow for propagation of chaos to take place. We shall see that this assumption allows us to have bounds for the moments of the state process and later to apply Chernoff's inequality at a crucial juncture.

The limit dynamics

In this section, we aim to define the limit dynamics of the RMF GL model when the number of replicas goes to infinity. As previously mentioned, intuitively, the arrivals from each neuron should become a Poisson process, whereas the reset term should remain unchanged.

As such, we introduce the following system of SDEs which is the natural candidate for the limit dynamics and to which we will hereafter refer to as the limit process. We will denote with tildes everything pertaining to it. We consider point processes $\tilde{N}_1, \dots, \tilde{N}_K$ on \mathbb{R}^+ with respective (\mathcal{F}_t) -stochastic intensities $\tilde{\lambda}_1, \dots, \tilde{\lambda}_K$, where \mathcal{F}_t is the internal spiking history of the network defined as previously, verifying the following stochastic differential equations: for $t > 0$, for $1 \leq i \leq K$,

$$\tilde{\lambda}_i(t) = \tilde{\lambda}_i(0) + \sum_{j \neq i} \mu_{j \rightarrow i} \tilde{A}_{j \rightarrow i}(t) + \int_0^t (r_i - \tilde{\lambda}_i(s)) \tilde{N}_i(ds), \quad (3.5)$$

where $\tilde{A}_{j \rightarrow i}$ are independent inhomogeneous Poisson point processes with intensities $a_j(t) = \int_0^t \mathbf{E}[\tilde{\lambda}_j(s)] ds = \mathbf{E}[\tilde{N}_j([0, t])]$ and $(\tilde{\lambda}_1(0), \dots, \tilde{\lambda}_K(0))$ verify Assumption 3.2.2.

The existence and uniqueness of the solution to this equation comes from the general theory of [70], and is derived analogously to the existence and uniqueness of the solution to (3.4), which is done in [8]. Note that (3.5) is a nonlinear equation in the sense of McKean-Vlasov [60], as the process $\tilde{\lambda}_i$ depends on its own law through the presence of the terms $\mathbf{E}[\tilde{\lambda}_j(t)]$ in the intensities of the Poisson processes.

The main result

Recall the following definition of convergence in total variation:

Definition 3.2.3. *Let P and Q be two probability measures on a probability space (Ω, \mathcal{F}) . We define the total variation distance by*

$$d_{TV}(P, Q) = \sup_{A \in \mathcal{F}} |P(A) - Q(A)|.$$

When Ω is countable, an equivalent definition is

$$d_{TV}(P, Q) = \frac{1}{2} \sum_{\omega \in \Omega} |P(\omega) - Q(\omega)|.$$

Note that certain authors use a multiplicative constant 2 when defining the total variation distance. We will also abusively say that random variables converge in total variation when their distributions do.

The following theorem is the main result of this work:

Theorem 3.2.4. *There exists $T > 0$ such that, if $t \in [0, T]$ and if*

$$A_{m,i}^M(t) = \sum_{j \neq i} \mu_{j \rightarrow i} \sum_{n \neq m} \int_0^t \mathbb{1}_{\{V_{(n,j) \rightarrow i}^M(s) = m\}} N_{n,j}^M(ds),$$

with $N_{m,i}^M$ defined in (3.4), and

$$\tilde{A}_i(t) = \sum_{j \neq i} \mu_{j \rightarrow i} \tilde{A}_{j \rightarrow i}(t),$$

with $\tilde{A}_{j \rightarrow i}(t)$ defined in (3.5), then,

1. the processes $(\tilde{A}_1, \dots, \tilde{A}_K)$ are independent, as are the processes $(\tilde{\lambda}_1, \dots, \tilde{\lambda}_K)$;
2. for all $(m, i) \in \{1, \dots, M\} \times \{1, \dots, K\}$, the random variable $A_{m,i}^M(t)$ converges in total variation to $\tilde{A}_i(t)$ when $M \rightarrow \infty$;
3. for all $(m, i) \in \{1, \dots, M\} \times \{1, \dots, K\}$, the random variable $\lambda_{m,i}^M(t)$ defined by (3.4) converges in total variation to $\tilde{\lambda}_i(t)$ defined in (3.5) when $M \rightarrow \infty$;
4. let \mathcal{N} be a finite subset of \mathbb{N}^* , for all $i \in \{1, \dots, K\}$, the processes $(A_{m,i}^M(\cdot))_{m \in \mathcal{N}}$ and $(\lambda_{m,i}^M(\cdot))_{m \in \mathcal{N}}$ weakly converge in the Skorokhod space $D([0, T])^{\text{card}(\mathcal{N})}$ endowed with the product Skorokhod metric to $\text{card}(\mathcal{N})$ independent copies of the corresponding limit processes $(\tilde{A}_i(\cdot))$ and $(\tilde{\lambda}_i(\cdot))$ when $M \rightarrow \infty$.

Here are a few remarks on this result.

First, note that for each $i, j \in \{1, \dots, K\}$, the variable $\mu_{j \rightarrow i} \tilde{A}_{j \rightarrow i}(t)$ is a scaled Poisson random variable and is thus a special case of a compound Poisson random variable. As such, unless all $\mu_{j \rightarrow i}$ are equal, $\tilde{A}_i(t)$ does not follow any standard Poisson or compound Poisson distribution.

Note also that we do not aim to prove \mathcal{L}^1 convergence, which we believe does not hold in this model. This marks a significant difference with classical mean field models for Hawkes processes, see [34]. From a computational point of view,

this is due to the fact that the averaging factor $\frac{1}{M-1}$ only appears in expectation in the interaction term $A_{m,i}^M(t)$.

Finally, note that statement 4 of the theorem is a consequence of statements 2 and 3, as weak convergence of point processes is implied by weak convergence of their void probabilities (see [45], Theorem 2.2), which comes directly from statements 2 and 3. More precisely, the convergence in total variation of $\lambda_{m,i}$ and $A_{m,i}$ implies the weak convergence of the void probabilities of $N_{m,i}$ and of the point process of arrival times, which in turn implies weak convergence of the point processes.

Methodology for the proof

In contrast with classical mean-field models presented in the beginning of the paper, in replica mean-fields, the mean-field simplification comes from the random routing operations between replicas. The input point process in the M -replica model consists in a superposition of M rare point processes, which informally explains why Poisson (or compound Poisson) processes arise at the limit. This point of view leads us to fix an instant $t \in \mathbb{R}^+$ and to consider the random variable of inputs up to time t as a random sum of Bernoulli random variables with means $\frac{1}{M-1}$. The Chen-Stein method is a natural candidate to obtain explicit bounds in the total variation metric between this random sum and a Poisson random variable. We generalize it to account for the fact that the amount of Bernoulli random variables in the sum is random. As far as the author is aware, this application is novel. The bound obtained through the Chen-Stein method does not easily converge to 0 when M goes to infinity. Namely, we obtain a term similar to the L^1 norm of an empirical mean of centered random variables that are not independent. In order to circumvent the difficult direct analysis of such a term, we uncouple the inputs and outputs of the dynamics by considering the replica-mean-field dynamics (3.4) as the fixed point of some function on the space of probability distributions on the space of càdlàg functions endowed with a metric rendering it complete. This procedure is often used in the study of stochastic differential equations to prove the existence and uniqueness of solutions to these equations, see, for example, [75] or [18]. Here, we use it to prove that a certain property, namely the convergence of an empirical mean, holds at the fixed point by proving that the property is preserved by the function and that the iterates of the function converge to its fixed point.

3.2.2 Proof of the Poisson Hypothesis for the RMF GL model

The aim of this section is to prove Theorem 3.2.4. We organize the proof as follows. First, we recall some well-known facts about Poisson embedding representations

for real-valued point processes and derive asymptotic independence from them. Then, we state and prove some properties of the RMF and limit dynamics that will be useful in the following steps of the proof. Afterwards, we present the Chen-Stein method, which we use to derive conditions for the validity of the Poisson approximation that we aim to prove. Finally, we interpret the RMF dynamics as the fixed point of some function on the space of probability measures on the space of càdlàg functions and we show that this function has properties that allow the aforementioned conditions to hold at the fixed point, thus proving the Poisson approximation result.

Poisson embedding representation and independence of the limit processes

First, recall the following result from [17] about Poisson embeddings:

Lemma 3.2.5. *Let N be a point process on \mathbb{R} . Let (\mathcal{F}_t) be an internal history of N . Suppose N admits a (\mathcal{F}_t) -stochastic intensity $\{\mu(t)\}_{t \in \mathbb{R}}$. Then there exists a Poisson point process \hat{N} with intensity 1 on \mathbb{R}^2 such that, for all $C \in \mathcal{B}(\mathbb{R})$,*

$$N(C) = \int_{C \times \mathbb{R}} \mathbb{I}_{[0, \mu(s)]}(u) \hat{N}(ds \times du).$$

This result states that any point process admitting a stochastic intensity can be embedded in a Poisson point process with intensity 1 on the positive half-plane by considering the points of the Poisson process which lie below the curve given by the stochastic intensity of the process. We now proceed to apply this in our model, constructing all the state processes coupled through their Poisson embeddings.

For $m \geq 1, M \geq 1, 1 \leq i \leq K$, let $\hat{N}_{m,i}$ be i.i.d. Poisson point processes on $\mathbb{R}^+ \times \mathbb{R}^+$ with intensity 1.

Let $\Omega = (\mathbb{R}^+ \times ((\mathbb{R}^+)^2)^{\mathbb{N}^*})^{\mathbb{N}^*}$ be a probability space endowed with the probability measure $(\mu_0 \otimes P)^{\otimes \mathbb{N}^*}$, where μ_0 is the law of the initial conditions and P is the law of a Poisson process with intensity 1 on $(\mathbb{R}^+)^2$. We construct on Ω the following processes:

- The processes $(N_{m,i}^M(t)), m \geq 1, M \geq 1, 1 \leq i \leq K$, with stochastic intensities $(\lambda_{m,i}^M(t))$ verifying

$$\begin{aligned} \lambda_{m,i}^M(t) = & \sum_{n \neq m} \sum_{j \neq i} \int_0^t \int_0^{+\infty} \mu_{j \rightarrow i} \mathbb{I}_{\{V_{(n,j)}^M(s)=m\}} \mathbb{I}_{[0, \lambda_{n,j}^M(s)]}(u) \hat{N}_{n,j}(ds \times du) \\ & + \int_0^t \int_0^{+\infty} (r_i - \lambda_{m,i}^M(s)) \mathbb{I}_{[0, \lambda_{m,i}^M(s)]}(u) \hat{N}_{m,i}(ds \times du) + \lambda_{m,i}^M(0), \end{aligned} \tag{3.6}$$

with $\lambda_{m,i}^M(0) = Z_i$ for all $m \in \mathbb{N}^*$ and where, for all M , $(V_{(n,j) \rightarrow i}^M(t))_j$ are càdlàg stochastic processes such that for each point T of $\hat{N}_{n,j}$, the random variables $(V_{(n,j) \rightarrow i}^M(T))_j$ are independent of the past, mutually independent and uniformly distributed on $\{1, \dots, M\} \setminus \{n\}$, considered as marks of the Poisson point process $\hat{N}_{n,j}$. Namely, to each point of the Poisson embedding, we attach a mark that is an element of $(\mathbb{N}^K)^{\mathbb{N}^*}$, where the M th term of the sequence corresponds to $(V_{(n,j) \rightarrow i}^M(T))_j$.

- The processes $(\tilde{N}_i(t))$, $1 \leq i \leq K$, with stochastic intensities $(\tilde{\lambda}_i(t))$ verifying

$$\begin{aligned} \tilde{\lambda}_i(t) &= \tilde{\lambda}_i(0) + \sum_{j \neq i} \int_0^t \int_0^{+\infty} \mu_{j \rightarrow i} \mathbb{I}_{[0, \mathbf{E}[\tilde{\lambda}_j(s)]]}(u) \hat{N}_{j,i}(ds \times du) \\ &\quad + \int_0^t \int_0^{+\infty} (r_i - \tilde{\lambda}_i(s)) \mathbb{I}_{[0, \tilde{\lambda}_i(s)]}(u) \hat{N}_{i,i}(ds \times du), \end{aligned} \quad (3.7)$$

with $\tilde{\lambda}_i(0) = Z_i$.

In other words, we construct the M -replica dynamics (3.4) and the limit processes (3.5) with the same initial conditions and Poisson embeddings $(\hat{N}_{m,i})$. We require that the law of the initial conditions (Z_i) verifies Assumption 3.2.2, in other words, we require it to have uniform polynomial bounds of its moments.

This representation allows us to derive the following, which is statement 1 of Theorem 3.2.4.

Lemma 3.2.6. *The processes $(\tilde{A}_{j \rightarrow i})_{1 \leq i, j \leq K}$ are independent, as are the processes $(\tilde{\lambda}_1, \dots, \tilde{\lambda}_K)$.*

Proof. For all $t \in [0, T]$, we can write using the construction above

$$\tilde{A}_{j \rightarrow i}(t) = \int_0^t \int_0^{+\infty} \mathbb{I}_{[0, \mathbf{E}[\tilde{\lambda}_j(s)]]}(u) \hat{N}_{j,i}(ds \times du).$$

Therefore, all the randomness in \tilde{A}_i is contained in the Poisson embeddings $(\hat{N}_{k,i})_{1 \leq k \leq K}$. Thus, for $i \neq j$, \tilde{A}_i and \tilde{A}_j are independent. The independence of the processes $(\tilde{\lambda}_1, \dots, \tilde{\lambda}_K)$ follows in the same manner. \square

Properties of the RMF and limit processes

In this section, we prove several properties of the RMF and limit dynamics that will be used throughout the proof.

In what follows, we will often omit the M superscript in the notations $N_{m,i}^M$, $A_{m,i}^M$ and $\lambda_{m,i}^M$ to increase readability.

Note that the arrival process $A_{m,i}(t)$ can be represented as a random sum of Bernoulli random variables with parameters $\frac{1}{M-1}$. Indeed, for $n \neq m$ and $j \neq i$, if $S \in \text{Supp}(N_{n,j}|_{[0,T]})$, we define $B_{S,(n,j) \rightarrow (m,i)}^M$ the random variable equal to 1 if the routing between replicas at time S caused by a spike in neuron j in replica n chose the replica m for the recipient i of the spike, and 0 otherwise. As such, it is clear that we can write for all $t \in [0, T]$, $m \in \{1, \dots, M\}$ and $i \in \{1, \dots, K\}$,

$$A_{m,i}(t) = \sum_{n \neq m} \sum_{j \neq i} \mu_{j \rightarrow i} \sum_{k \leq N_{n,j}([0,t])} B_{k,(n,j) \rightarrow (m,i)}^M. \quad (3.8)$$

Note that when m, n, i and j are fixed, the random variables $(B_{k,(n,j) \rightarrow (m,i)}^M)_{k \leq N_{n,j}([0,T])}$ are i.i.d. Also note that when n, j, i and k are fixed, the joint distribution of $(B_{k,(n,j) \rightarrow (m,i)}^M)_m$ with $m \in \{1, \dots, M\}$ is that of Bernoulli random variables with parameter $\frac{1}{M-1}$ such that exactly one of them is equal to 1, all the others being equal to 0. Combining these two observations allows us to prove a lemma that highlights the core of the replica-mean-field approach:

Lemma 3.2.7. *Fix $(m, i) \in \{1, \dots, M\} \times \{1, \dots, K\}$. Keeping notation from (3.8), let*

$$N = (N_{n,j}([0, t]))_{n \neq m, j \neq i} \in \mathbb{N}^{(K-1)(M-1)}.$$

Conditionally on the event $\{N = q\}$, for $q = (q_{n,j})_{n \neq m, j \neq i} \in \mathbb{N}^{(K-1)(M-1)}$, the random variables $(B_{k,(n,j) \rightarrow (m,i)}^M)_{n \neq m, j \neq i, k \in \{1, \dots, q_{n,j}\}}$ are independent Bernoulli random variables with parameter $\frac{1}{M-1}$.

Proof. The structure of the proof goes as follows: since N is entirely determined by the Poisson embeddings $(\hat{N}_{n,j})_{j \neq i}$ and the arrivals to the nodes (n, j) from all the nodes $h \neq j$ across replicas, it is sufficient to show that these arrivals and the routing variables $(B_{k,(n,j) \rightarrow (m,i)}^M)_{k \leq \hat{N}_{n,j}([0,t] \times \mathbb{R}^+)}$ are independent. Intuitively, this holds because arrivals are aggregated across all replicas, which will erase the eventual dependencies due to the routing variables to nodes i choosing one replica rather than another.

In order to transcribe this intuition rigorously, we first show that the total number of departures from nodes i up to time t , that is, $\sum_{l=1}^M N_{l,i}([0, t])$, and the routing variables $(B_{k,(n,j) \rightarrow (m,i)}^M)_{k \leq \hat{N}_{n,j}([0,t] \times \mathbb{R}^+)}$ are independent. Indeed, using the representation given by Lemma 3.2.5, due to the structure of the Poisson embeddings $(\hat{N}_{l,i})_{l \in \{1, \dots, M\}}$, there is a point of $\sum_{l=1}^M N_{l,i}$ in some interval I iff there is a point of the superposition of the corresponding Poisson embeddings such that the x-coordinate is in I and the y-coordinate is under the curve of the function $t \rightarrow \sum_{l=1}^M \lambda_{l,i}(t)$. In turn, the last event does not depend on $(B_{k,(n,j) \rightarrow (m,i)}^M)_{k \leq \hat{N}_{n,j}([0,t] \times \mathbb{R}^+)}$, as the symmetry inherent to the replica structure ensures that all arrivals increment $t \rightarrow \sum_{l=1}^M \lambda_{l,i}(t)$ by the same amount, which concludes the proof of this preliminary remark.

For all (n, j) such that $n \neq m$ and $j \neq i$, let

$$A_{i \rightarrow (n, j)}(t) = \sum_{l \neq n} \sum_{k \leq N_{l, i}([0, t])} B_{k, (l, i) \rightarrow (n, j)}^M.$$

Note that $A_{i \rightarrow (n, j)}(t)$ represents the arrivals to node j in replica n from all nodes i across replicas. As such, it is clear that we can write

$$A_{i \rightarrow (n, j)}(t) = \sum_{k \leq \sum_{l \neq n} N_{l, i}([0, t])} B_{k, (i) \rightarrow (n, j)}^M,$$

where $(B_{k, (i) \rightarrow (n, j)}^M)$ are independent Bernoulli random variables with parameter $\frac{1}{M-1}$ such that they and $(B_{k, (n, j) \rightarrow (m, i)}^M)$ are independent. Then by the previous observation, $A_{i \rightarrow (n, j)}(t)$ and $(B_{k, (n, j) \rightarrow (m, i)}^M)$ are independent. Therefore, N , which is entirely determined by the Poisson embeddings $(\hat{N}_{n, j})$ and the arrivals $(A_{h \rightarrow (n, j)}(t))_{h \neq j}$, and $(B_{k, (n, j) \rightarrow (m, i)}^M)_{k \leq \hat{N}_{n, j}([0, T] \times \mathbb{R}^+)}$, are independent. Thus, conditioning on N does not break independence between the variables $(B_{k, (n, j) \rightarrow (m, i)}^M)$. \square

We will now give bounds on the moments of both the M-replica and limit processes, using the bounds on the moments of the initial conditions.

Lemma 3.2.8. *Suppose the initial conditions verify Assumption 3.2.2. Then, for all $p \geq 1$, for all $(m, i) \in \{1, \dots, M\} \times \{1, \dots, K\}$, for all $t \in [0, T]$, there exists $Q_p(T) \in \mathbb{R}_p[X]$ a polynomial of degree exactly p such that*

$$\mathbf{E}[\lambda_{m, i}^p(t)] \leq Q_p(\mathbf{E}[\lambda_{m, i}(0)]). \quad (3.9)$$

Proof. We first prove the result for $p = 1$ in the special case where exchangeability also holds between neurons and where there are no resets. Namely, we temporarily consider here that for all $i, j \in \{1, \dots, K\}$, $\mu_{j \rightarrow i} = 1$. Let $t \in [0, T]$. Thus, we have

$$\mathbf{E}[\lambda_{m, i}(t)] = \mathbf{E}[\lambda_{m, i}(0)] + \sum_{n \neq m} \sum_{j \neq i} \mathbf{E} \left[\int_0^t \mathbb{1}_{\{V_{(n, j) \rightarrow i}(s) = m\}} N_{n, j}(ds) \right].$$

Denoting by $((T_{n, j})_r)_{r \in \mathbb{Z}}$ the points of $N_{n, j}$, we have

$$\begin{aligned} \mathbf{E}[\lambda_{m, i}(t)] &= \mathbf{E}[\lambda_{m, i}(0)] \\ &+ \sum_{n \neq m} \sum_{j \neq i} \mathbf{E} \left[\sum_{r \in \mathbb{Z}} \mathbf{E}[\mathbb{1}_{\{V_{(n, j) \rightarrow i}((T_{n, j})_r) = m\}} \mathbb{1}_{\{(T_{n, j})_r \in [0, t]\}} | \mathcal{F}_{(-\infty, (T_{n, j})_r)}^{N_{n, j}}] \right]. \end{aligned}$$

Using the predictability w.r.t. the network history of the processes $V_{(n,j) \rightarrow i}^M$, we have

$$\mathbf{E}[\lambda_{m,i}(t)] = \mathbf{E}[\lambda_{m,i}(0)] + \sum_{n \neq m} \sum_{j \neq i} \mathbf{E} \left[\sum_{r \in \mathbb{Z}} \mathbb{1}_{\{V_{(n,j) \rightarrow i}((T_{n,j})_r) = m\}} \mathbb{1}_{\{(T_{n,j})_r \in [0,t)\}} \right].$$

Using the property of stochastic intensity, we can rewrite this as

$$\mathbf{E}[\lambda_{m,i}(t)] = \mathbf{E}[\lambda_{m,i}(0)] + (K-1) \mathbf{E} \left[\int_0^t \lambda_{m,i}(s) ds \right].$$

Therefore,

$$\mathbf{E}[\lambda_{m,i}(t)] = \mathbf{E}[\lambda_{m,i}(0)] + (K-1) \int_0^t \mathbf{E}[\lambda_{m,i}(s)] ds.$$

This gives

$$\mathbf{E}[\lambda_{m,i}(t)] \leq \mathbf{E}[\lambda_{m,i}(0)] e^{(K+r_i-1)T} := Q_1(\mathbf{E}[\lambda_{m,i}(0)]). \quad (3.10)$$

Now, let us write the differential equation for $\lambda_{m,i}^2$, still considering the dynamics without resets and with equal weights:

$$\lambda_{m,i}^2(t) = \lambda_{m,i}^2(0) + \sum_{n \neq m} \sum_{j \neq i} \int_0^t \mathbb{1}_{\{V_{(n,j) \rightarrow i}(s) = m\}} (2\lambda_{n,j}(s) + 1) N_{n,j}(ds).$$

Therefore, we have

$$\mathbf{E}[\lambda_{m,i}^2(t)] = \mathbf{E}[\lambda_{m,i}^2(0)] + (K-1) \int_0^t (2\mathbf{E}[\lambda_{m,i}^2(s)] + \mathbf{E}[\lambda_{m,i}(s)]) ds$$

which gives using (3.10) the bound

$$\mathbf{E}[\lambda_{m,i}^2(t)] \leq \mathbf{E}[\lambda_{m,i}^2(0)] + (K-1)Q_1((\mathbf{E}[\lambda_{m,i}(0)])T) + 2(K-1) \int_0^t \mathbf{E} \left[\sup_{u \in [0,s]} \lambda_{m,i}^2(u) \right] ds.$$

By applying Gronwall's lemma and using the assumption on the initial conditions, we get

$$\mathbf{E}[\lambda_{m,i}^2(t)] \leq ((\mathbf{E}[\lambda_{m,i}(0)]^2 + (K-1)Q_1((\mathbf{E}[\lambda_{m,i}(0)])T)) e^{2(K-1)T} := Q_2((\mathbf{E}[\lambda_{m,i}(0)]).$$

This reasoning can be extended by induction to all $p \geq 3$, which proves the result for the case of exchangeable interactions.

Now, to get the result in the general case where all $\mu_{j \rightarrow i}$ are not necessarily all equal to 1, note that by monotonicity the dynamics that we consider (in both the exchangeable and nonexchangeable cases) are stochastically dominated by the same dynamics without the reset terms. Finally, note that nonexchangeable dynamics without the reset terms are stochastically dominated by exchangeable dynamics without the reset terms. This shows that the moment bounds still hold in the general case. \square

Finally, note that the exact same reasoning can be applied to obtain an equivalent result for the limit process, which we will only state:

Lemma 3.2.9. *For all $p \geq 1$, for all $i \in \{1, \dots, K\}$, for all $t \in [0, T]$, there exists a polynomial $\tilde{Q}_p \in \mathbb{R}_p[X]$ a polynomial of degree exactly p such that*

$$\mathbf{E}[\tilde{\lambda}_i^p(t)] \leq \tilde{Q}_p[\mathbf{E}[\tilde{\lambda}_i(0)]]. \quad (3.11)$$

Lemma 3.2.8 allows us to prove the following result, which states that Assumption 3.2.2 can be propagated to any time t less than some fixed T .

Lemma 3.2.10. *There exists $T > 0$ and $\xi_0 > 0$ (which is the same as in Assumption 3.2.2) such that for $t \in [0, T]$ and $\xi \leq \xi_0$,*

$$\mathbf{E}[e^{\xi\lambda_{m,i}(t)}] < \infty \text{ and } \mathbf{E}[e^{\xi\tilde{\lambda}_i(t)}] < \infty. \quad (3.12)$$

Proof. To prove this result, we once again consider exchangeable dynamics without resets, using the same observation as previously, namely that nonexchangeable dynamics with resets are stochastically dominated by exchangeable dynamics without resets, to generalize the result. Let us thus assume $\mu_{j \rightarrow i} = \mu$ for all i, j . Let ξ_0 as in Assumption 3.2.2. Let $t \in [0, T]$. Let us write out the equation verified by $e^{\xi\lambda_{m,i}(t)}$:

$$e^{\xi\lambda_{m,i}(t)} = e^{\xi\lambda_{m,i}(0)} + \sum_{j \neq i} \sum_{n \neq m} \int_0^t \mathbb{I}_{\{V_{(n,j) \rightarrow i}^M(s) = m\}} e^{\xi\lambda_{m,i}(s)} (e^{\xi\mu} - 1) N_{n,j}(ds).$$

Taking the expectation and using the stochastic intensity property, we get

$$\mathbf{E}[e^{\xi\lambda_{m,i}(t)}] = \mathbf{E}[e^{\xi\lambda_{m,i}(0)}] + \frac{1}{M-1} \sum_{j \neq i} \sum_{n \neq m} \int_0^t \mathbf{E}[e^{\xi\lambda_{m,i}(s)} (e^{\xi\mu} - 1) \lambda_{n,j}(s)] ds.$$

Using exchangeability between replicas, this boils down to

$$\mathbf{E}[e^{\xi\lambda_{m,i}(t)}] = \mathbf{E}[e^{\xi\lambda_{m,i}(0)}] + \sum_{j \neq i} \int_0^t \mathbf{E}[e^{\xi\lambda_{m,i}(s)} (e^{\xi\mu} - 1) \lambda_{m,j}(s)] ds.$$

Since we are looking at dynamics without resets, $\lambda_{m,i}(s)$ and $\lambda_{m,j}(s)$ are positively correlated. Therefore, we have

$$\mathbf{E}[e^{\xi\lambda_{m,i}(t)}] \leq \mathbf{E}[e^{\xi\lambda_{m,i}(0)}] + (e^{\xi\mu} - 1) \sum_{j \neq i} \int_0^t \mathbf{E}[e^{\xi\lambda_{m,i}(s)}] \mathbf{E}[\lambda_{m,j}(s)] ds.$$

By Lemma 3.2.8 and Assumption 3.2.2, we have the existence of a constant $B > 0$ such that

$$\mathbf{E}[e^{\xi\lambda_{m,i}(t)}] \leq \mathbf{E}[e^{\xi\lambda_{m,i}(0)}] + (e^{\xi\mu} - 1)(K-1)B \int_0^t \mathbf{E}[e^{\xi\lambda_{m,i}(s)}] ds.$$

The desired result follows from Grönwall's lemma. \square

The final lemma we will state and prove in this section concerns the means of the RMF and limit processes. Namely, we show that the replica mean-field construction preserves the mean as M varies and that the mean of the limit process coincides with the mean of the RMF process.

Lemma 3.2.11. *For all $M \geq 2$, for all $(m, i) \in \{1, \dots, M\} \times \{1, \dots, K\}$, if the initial conditions are such that for all $(m, i) \in \{1, \dots, M\} \times \{1, \dots, K\}$,*

$$\lambda_{m,i}(0) = \tilde{\lambda}_i(0),$$

then there exists $T \in \mathbb{R}$, such that for all $t \in [0, T]$, we have

$$\mathbf{E}[A_{m,i}(t)] = \mathbf{E}[\tilde{A}_i(t)]$$

and

$$\mathbf{E}[\lambda_{m,i}(t)] = \mathbf{E}[\tilde{\lambda}_i(t)].$$

Proof. Using both the property of stochastic intensity and exchangeability between replicas, we have as previously,

$$\mathbf{E}[A_{m,i}(t)] = \sum_{j \neq i} \mu_{j \rightarrow i} \mathbf{E}[N_{m,j}([0, t])] = \sum_{j \neq i} \mu_{j \rightarrow i} \int_0^t \mathbf{E}[\lambda_{m,j}(s)] ds.$$

Similarly, we have

$$\mathbf{E}[\tilde{A}_i(t)] = \sum_{j \neq i} \mu_{j \rightarrow i} \int_0^t \mathbf{E}[\tilde{\lambda}_j(s)] ds.$$

Therefore, we see that

$$\left| \mathbf{E}[A_{m,i}(t) - \tilde{A}_i(t)] \right| = \left| \sum_{j \neq i} \mu_{j \rightarrow i} \int_0^t \mathbf{E}[\lambda_{m,j}(s) - \tilde{\lambda}_j(s)] ds \right|. \quad (3.13)$$

Thus, it is sufficient to show that $\mathbf{E}[\lambda_{m,j}(s) - \tilde{\lambda}_j(s)] = 0$.

Now, let $t \in [0, T]$. Let $i(t) = \arg \max_{j \in \{1, \dots, K\}} \left| \mathbf{E}[\lambda_{m,j}(t) - \tilde{\lambda}_j(t)] \right|$ for $t \in [0, T]$.

Now, using (3.4) and (3.5), we have

$$\begin{aligned}
\left| \mathbf{E}[\lambda_{m,i(t)}(t) - \tilde{\lambda}_{i(t)}(t)] \right| &\leq \left| \mathbf{E}[\lambda_{m,i(t)}(0) - \tilde{\lambda}_{i(t)}(0)] + \mathbf{E}[A_{m,i(t)}(t) - \tilde{A}_{i(t)}(t)] \right. \\
&\quad \left. + \mathbf{E} \left[\int_0^t (r_i - \lambda_{m,i(t)}(s)) \lambda_{m,i(t)}(s) - (r_i - \tilde{\lambda}_{i(t)}(s)) \tilde{\lambda}_{i(t)}(s) ds \right] \right| \\
&\leq 0 + \sum_{j \neq i(t)} \mu_{i(t),j} \int_0^t \left| \mathbf{E}[\lambda_{m,j}(s) - \tilde{\lambda}_j(s)] \right| ds \\
&\quad + r_i \int_0^t \left| \mathbf{E}[\lambda_{m,i(t)}(s) - \tilde{\lambda}_{i(t)}(s)] \right| ds \\
&\quad + \left| \int_0^t \mathbf{E}[\lambda_{m,i(t)}^2(s) - \tilde{\lambda}_{i(t)}^2(s)] ds \right|,
\end{aligned}$$

using the coupling on the initial conditions and (3.13).

Let $C > 0$. Let $A_C(t, [0, T]) := \{(\omega, t) \in \Omega \times [0, T] \mid \max(\lambda_{m,i(t)}(t, \omega), \tilde{\lambda}_{i(t)}(t, \omega)) < C\}$.

Then, we write

$$\begin{aligned}
\left| \int_0^t \mathbf{E}[\lambda_{m,i(t)}^2(s) - \tilde{\lambda}_{i(t)}^2(s)] ds \right| &= \left| \int_0^t \int_{A_C(s, [0, T])} (\lambda_{m,i(t)}^2(s) - \tilde{\lambda}_{i(t)}^2(s)) ds P(d\omega) \right. \\
&\quad \left. + \int_0^t \int_{A_C^c(s, [0, T])} (\lambda_{m,i(t)}^2(s) - \tilde{\lambda}_{i(t)}^2(s)) ds P(d\omega) \right|.
\end{aligned}$$

Since we have $\mathbf{E}[\lambda_{m,i(t)}^2(s) - \tilde{\lambda}_{i(t)}^2(s)] = \mathbf{E}[(\lambda_{m,i(t)}(s) - \tilde{\lambda}_{i(t)}(s))(\lambda_{m,i(t)}(s) + \tilde{\lambda}_{i(t)}(s))]$, by definition of $A_C([0, T])$ the first term can be bounded by

$$\left| \int_0^t \int_{A_C(s, [0, T])} (\lambda_{m,i(t)}^2(s) - \tilde{\lambda}_{i(t)}^2(s)) ds P(d\omega) \right| \leq 2C \left| \int_0^t \mathbf{E}[\lambda_{m,i(t)}(s) - \tilde{\lambda}_{i(t)}(s)] ds \right|. \tag{3.14}$$

For the second term, we write

$$\begin{aligned}
&\left| \int_0^t \int_{A_C^c(s, [0, T])} (\lambda_{m,i(t)}^2(s) - \tilde{\lambda}_{i(t)}^2(s)) ds P(d\omega) \right| = \\
&\left| \int_0^t \mathbf{E}[(\lambda_{m,i(t)}^2(s) - \tilde{\lambda}_{i(t)}^2(s)) \mathbb{I}_{A_C^c(s, [0, T])}] ds \right|.
\end{aligned}$$

Now, using the Cauchy-Schwarz inequality,

$$\left| \mathbf{E}[(\lambda_{m,i(t)}^2(s) - \tilde{\lambda}_{i(t)}^2(s)) \mathbb{I}_{A_C^c(s, [0, T])}] \right| \leq \sqrt{\mathbf{E}[(\lambda_{m,i(t)}^2(s) - \tilde{\lambda}_{i(t)}^2(s))^2] \mathbf{E}[\mathbb{I}_{A_C^c(s, [0, T])}]}.$$

Using Lemmas 3.2.8 and 3.2.9 that give bounds on the moments of the considered processes, we have the existence of a positive constant $D_1(T)$ such that

$$\sqrt{\mathbf{E}[(\lambda_{m,i(t)}^2(s) - \tilde{\lambda}_{i(t)}^2(s))^2]} \leq D_1(T).$$

For the last term, we have by definition of $A_C^c(s, [0, T])$,

$$\mathbf{E}[\mathbb{1}_{A_C^c(s, [0, T])}] = \mathbf{P}(\max(\lambda_{m,i(t)}(s), \tilde{\lambda}_{i(t)}(s)) > C).$$

By Lemma 3.2.10, there exists $T > 0$ such that $\mathbf{E}[e^{6T \max(\lambda_{m,i(t)}(s), \tilde{\lambda}_{i(t)}(s))}]$ is finite. Applying the Chernoff inequality, we have

$$\mathbf{P}(e^{6T \max(\lambda_{m,i(t)}(s), \tilde{\lambda}_{i(t)}(s))} > e^{6CT}) \leq \mathbf{E}[e^{6T \max(\lambda_{m,i(t)}(s), \tilde{\lambda}_{i(t)}(s))}] e^{-6CT}.$$

Using Lemma 3.2.8, this shows that there exists a constant $D_2(T) > 0$ such that

$$\sqrt{\mathbf{E}[\mathbb{1}_{A_C^c(s, [0, T])}]} \leq D_2(T)e^{-3CT}.$$

Combining the previous bounds, we finally obtain

$$\left| \int_0^t \int_{A_C^c(s, [0, T])} (\lambda_{m,i(t)}^2(s) - \tilde{\lambda}_{i(t)}^2(s)) ds P(d\omega) \right| \leq D_1(T)D_2(T)Te^{-3CT}. \quad (3.15)$$

Combining (3.14) and (3.15), we have

$$\left| \int_0^t \mathbf{E}[(\lambda_{m,i(t)}^2(s) - \tilde{\lambda}_{i(t)}^2(s))] ds \right| \leq 2C \left| \int_0^t \mathbf{E}[\lambda_{m,i(t)}(s) - \tilde{\lambda}_{i(t)}(s)] ds \right| + D_1(T)D_2(T)Te^{-3CT}.$$

Therefore, we have

$$\begin{aligned} \left| \mathbf{E}[\lambda_{m,i(t)}(t) - \tilde{\lambda}_{i(t)}(t)] \right| &\leq D_1(T)D_2(T)Te^{-3CT} \\ &+ (r_i + 2C) \int_0^t \left| \mathbf{E}[\lambda_{m,i(t)}(s) - \tilde{\lambda}_{i(t)}(s)] \right| ds \\ &+ \sum_{j \neq i(t)} \mu_{i(t),j} \int_0^t \left| \mathbf{E}[\lambda_{m,j}(s) - \tilde{\lambda}_j(s)] \right| ds. \end{aligned}$$

By definition of $i(t)$, we then have

$$\begin{aligned}
\left| \mathbf{E}[\lambda_{m,i(t)}(t) - \tilde{\lambda}_{i(t)}(t)] \right| &\leq D_1(T)D_2(T)e^{-3CT}T \\
&+ (r_i + 2C) \int_0^t \left| \mathbf{E}[\lambda_{m,i(s)}(s) - \tilde{\lambda}_{i(s)}(s)] \right| ds \\
&+ \sum_{j \neq i(t)} \mu_{i(t),j} \int_0^t \left| \mathbf{E}[\lambda_{m,i(s)}(s) - \tilde{\lambda}_{i(s)}(s)] \right| ds. \\
&\leq D_1(T)D_2(T)e^{-3CT}T \\
&+ \left(r_i + 2C + \sum_{j \neq i(t)} \mu_{i(t),j} \right) \int_0^t \left| \mathbf{E}[\lambda_{m,i(s)}(s) - \tilde{\lambda}_{i(s)}(s)] \right| ds.
\end{aligned}$$

By Gronwall's lemma, we therefore get

$$\left| \mathbf{E}[\lambda_{m,i(t)}(t) - \tilde{\lambda}_{i(t)}(t)] \right| \leq D_1(T)D_2(T)e^{-3CT}T e^{(r_i + 2C + \sum_{j \neq i(t)} \mu_{i(t),j})T}.$$

Finally, note that for any $\varepsilon > 0$, we can now choose $C > 0$ such that

$$D_1(T)D_2(T)T e^{-3CT} e^{(r_i + 2C + \sum_{j \neq i(t)} \mu_{i(t),j})T} \leq \varepsilon.$$

By the choice of $i(t)$, the result follows for any $1 \leq i \leq K$. \square

Poisson approximation bound using the Chen-Stein method

The goal of this section is to use the Chen-Stein method, which we will briefly recall, to obtain a bound in total variation distance between the arrivals term (3.8) and the limit sum of Poisson random variables. Recall that (3.8) states that for all $t \in [0, T]$, $m \in \{1, \dots, M\}$, $i \in \{1, \dots, K\}$,

$$A_{m,i}(t) = \sum_{n \neq m} \sum_{j \neq i} \mu_{j \rightarrow i} \sum_{k \leq N_{n,j}([0,t])} B_{k,(n,j) \rightarrow (m,i)}^M.$$

Recall that if Z is a random variable taking values in \mathbb{N} with $\mathbf{E}[Z] < \infty$, Z is a Poisson random variable iff the distribution of $Z + 1$ is equal to the distribution of the size-biased version of Z , in other words, iff for all bounded functions f on \mathbb{N} ,

$$\mathbf{E}[Z] \mathbf{E}[f(Z + 1)] = \mathbf{E}[f(Z)Z]. \tag{3.16}$$

The key principle of the Chen-Stein method is to say that if (3.16) holds approximately for some r.v. Z for any bounded function f on \mathbb{N} , then Z approximately has a Poisson distribution. In the case of a sum of Bernoulli random variables that are not necessarily independent, we have the following result:

Lemma 3.2.12. *Let $l \in \mathbb{N}$. Consider $W = \sum_{i=1}^l Y_i$, where Y_i are Bernoulli random variables with respective means p_i , without any independence assumptions. Let Z be a Poisson distributed random variable with mean $\mathbf{E}[W] = \sum_i p_i$. For $1 \leq k \leq l$, let U_k and V_k be random variables on the same probability space such that U_k has the same distribution as W and $1 + V_k$ has the same distribution as W conditioned on the event $Y_k = 1$ (with the convention $V_k = 0$ if $\mathbf{P}(Y_k = 1) = 0$). Then*

$$d_{TV}(W, Z) \leq \left(1 \wedge \frac{1}{\mathbf{E}[W]}\right) \sum_{i=1}^l p_i \mathbf{E}[|U_i - V_i|].$$

Then, it suffices to exhibit a coupling of U_i and V_i such that $\mathbf{E}[|U_i - V_i|]$ is small. We refer to [13], [58] or [10] for a comprehensive overview of the Chen-Stein method.

We now adapt the Chen-Stein method to the replica-mean-field framework, generalizing the method to the case of a random sum of Bernoulli random variables.

We first give a result that will be an immediate corollary of the lemma we prove afterwards to compare it with Lemma 3.2.12:

Lemma 3.2.13. *Let L be a \mathbb{N} -valued random variable such that $\mathbf{E}[L] < \infty$. Let $(Y_i)_{i \in \mathbb{N}}$ be random variables such that for any $l \in \mathbb{N}$, conditionally on the event $\{L = l\}$, Y_i are Bernoulli random variables with respective means p_i . Consider $W = \sum_{i=1}^L Y_i$. Let Z be a Poisson distributed random variable with mean $\mathbf{E}[W] = \mathbf{E}[\sum_i p_i]$. For $k \in \mathbb{N}$, let U_k and V_k be random variables on the same probability space such that for $l \in \mathbb{N}$, conditionally on the event $\{L = l\}$, U_k has the same distribution as W and $1 + V_k$ has the same distribution as W conditioned on the event $Y_k = 1$ (with the convention $V_k = 0$ if $\mathbf{P}(Y_k = 1 | L = l) = 0$). Then*

$$d_{TV}(W, Z) \leq \left(1 \wedge \frac{0.74}{\mathbf{E}[W]}\right) \mathbf{E}[|L - \mathbf{E}[L]|] + \left(1 \wedge \frac{1}{\mathbf{E}[W]}\right) \mathbf{E} \left[\sum_{i=1}^L p_i \mathbf{E}[|U_i - V_i| | L] \right].$$

For our purposes, we will prove a slightly different result with $p_i = \frac{1}{M-1}$ in a vector setting, but it is easy to see that Lemma 3.2.13 can be proven in the same way as what follows.

We will now use notation consistent with (3.8). Since what follows is done with $t \in [0, T]$ fixed, we will additionally write $N_{n,j}([0, t])$ as $N_{n,j}$ in this section, continuing to omit the M superscript to simplify notation.

Lemma 3.2.14. *Let $M > 1$. Let $(m, i) \in \{1, \dots, M\} \times \{1, \dots, K\}$. For $j \in \{1, \dots, K\} \setminus \{i\}$, let $A_{j \rightarrow (m,i)} = \sum_{n \neq m} \sum_{k=1}^{N_{n,j}} B_{k,(n,j) \rightarrow (m,i)}$ and let $\tilde{A}_{j \rightarrow i}$ be indepen-*

dent Poisson random variables with means $\mathbf{E}[N_{1,j}]$ as in (3.5). Then,

$$\begin{aligned} d_{TV}(A_{j \rightarrow (m,i)}, \tilde{A}_{j \rightarrow i}) &\leq \left(1 \wedge \frac{0.74}{\sqrt{\mathbf{E}[N_{1,j}]}}\right) \frac{1}{M-1} \mathbf{E} \left[\left| \sum_{n \neq m} (\mathbf{E}[N_{n,j}] - N_{n,j}) \right| \right] \\ &\quad + \frac{1}{M-1} \left(1 \wedge \frac{1}{\mathbf{E}[N_{1,j}]}\right) \mathbf{E}[N_{1,j}]. \end{aligned} \tag{3.17}$$

A few remarks on this result are in order. First, note that the two terms in the right hand side of (3.17) are very different in nature. The second term goes to 0 when $M \rightarrow \infty$ due to the moment bound obtained in Lemma 3.2.8, whereas the first term is the \mathcal{L}^1 norm of an empirical mean of centered random variables which are not independent. As such, obtaining the convergence to 0 of that term when $M \rightarrow \infty$ requires proving an \mathcal{L}^1 law of large numbers result for non i.i.d. summands, which is not trivial and will be the subject of the next section. Next, note that the two terms can be heuristically interpreted in the following way: the second term represents a Le Cam-type bound [57] between a sum of Bernoulli random variables and a sum of Poisson random variables with equal means, in the case where the amount of summands is random. The first term represents the distance between such a random sum of Poisson random variables and a Poisson random variable whose mean is the mean number of summands, similar to [76].

Finally, note that this lemma only provides a bound for fixed i and $j \in \{1, \dots, K\} \setminus \{i\}$: $A_{j \rightarrow (m,i)}$ represents the arrivals from nodes j across replicas to node i in replica m . Thus Lemma 3.2.14 does not directly give a bound for the approximation of $A_{m,i}$ by $\tilde{A}_i = \sum_{j \neq i} \mu_{j \rightarrow i} \tilde{A}_{j \rightarrow i}$. However, since by Lemma 3.2.6, we have asymptotic independence, it is natural to expect that the eventual convergence in total variation will also take place for the sum, and we shall see later that it does indeed hold.

We now proceed to the proof of Lemma 3.2.14.

Proof. For $B \subset \mathbb{N}$ and $j \in \{1, \dots, K\}$, let g_B be the solution to the following equation, sometimes referred to as the Stein equation, see [23]:

$$\mathbf{E}[N_{1,j}]g_B(k+1) - kg_B(k) = \mathbb{1}_B(k) - \mathbf{P}(\tilde{A}_{j \rightarrow i} \in B),$$

for $k \in \mathbb{N}$, with initial condition $g_B(0) = 0$.

As in the proof of Lemma 3.2.7, let $N = (N_{n,j})_{(n,j) \in (\{1, \dots, M\} \setminus \{m\}) \times (\{1, \dots, K\} \setminus \{i\})}$.

Let $B \subset \mathbb{N}$. We have

$$\begin{aligned}
 & \mathbf{P}(A_{j \rightarrow (m,i)} \in B) - \mathbf{P}(\tilde{A}_{j \rightarrow i} \in B) = \mathbf{E}[\mathbb{1}_{A_{m,i,j} \in B} - \mathbf{P}(\tilde{A}_{j \rightarrow i} \in B)] \\
 & = \mathbf{E} \left[\mathbf{E}[N_{1,j}]g_B(A_{j \rightarrow (m,i)} + 1) - A_{j \rightarrow (m,i)}g_B(A_{j \rightarrow (m,i)}) \right] \text{ by the Stein equation} \\
 & = \mathbf{E} \left[\mathbf{E} \left[\mathbf{E}[N_{1,j}]g_B(A_{j \rightarrow (m,i)} + 1) - \sum_{n \neq m} \sum_{k=1}^{N_{n,j}} B_{k,(n,j) \rightarrow (m,i)}g_B(A_{j \rightarrow (m,i)}) \middle| N \right] \right] \\
 & = \mathbf{E} \left[\mathbf{E} \left[\frac{1}{M-1} \sum_{n \neq m} (\mathbf{E}[N_{n,j}] - N_{n,j})g_B(A_{j \rightarrow (m,i)} + 1) \right. \right. \\
 & \quad \left. \left. + \sum_{n \neq m} \sum_{k=1}^{N_{n,j}} \left(\frac{g_B(A_{j \rightarrow (m,i)} + 1)}{M-1} - B_{k,(n,j) \rightarrow (m,i)}g_B(A_{j \rightarrow (m,i)}) \right) \middle| N \right] \right].
 \end{aligned}$$

For all $n \neq m$ and all $1 \leq k \leq N_{n,j}$, let $U_{k,(n,j) \rightarrow (m,i)}$ and $V_{k,(n,j) \rightarrow (m,i)}$ be random variables on the same probability space such that $U_{k,(n,j) \rightarrow (m,i)} \stackrel{\mathcal{L}}{=} A_{j \rightarrow (m,i)}$ and $\mathbf{P}(V_{k,(n,j) \rightarrow (m,i)} + 1 \in \cdot) \stackrel{\mathcal{L}}{=} \mathbf{P}(A_{j \rightarrow (m,i)} \in \cdot | B_{k,(n,j) \rightarrow (m,i)} = 1)$ conditionally on N .

Using Lemma 3.2.7, we have for all k

$$\mathbf{E}[B_{k,(n,j) \rightarrow (m,i)} | N] = \frac{1}{M-1}.$$

Therefore,

$$\begin{aligned}
 & \mathbf{P}(A_{j \rightarrow (m,i)} \in B) - \mathbf{P}(\tilde{A}_{j \rightarrow i} \in B) = \\
 & \mathbf{E} \left[\mathbf{E} \left[\frac{1}{M-1} \sum_{n \neq m} (\mathbf{E}[N_{n,j}] - N_{n,j})g_B(A_{j \rightarrow (m,i)} + 1) \middle| N \right] \right] \\
 & + \frac{1}{M-1} \mathbf{E} \left[\mathbf{E} \left[\sum_{n \neq m} \sum_{k=1}^{N_{n,j}} (g_B(U_{k,(n,j) \rightarrow (m,i)} + 1) - g_B(V_{k,(n,j) \rightarrow (m,i)} + 1)) \middle| N \right] \right].
 \end{aligned}$$

Thus, we have:

$$\begin{aligned}
 & \left| \mathbf{P}(A_{j \rightarrow (m,i)} \in B) - \mathbf{P}(\tilde{A}_{j \rightarrow i} \in B) \right| \leq \frac{\|g_B\|}{M-1} \mathbf{E} \left[\left| \sum_{n \neq m} (\mathbf{E}[N_{n,j}] - N_{n,j}) \right| \right] \\
 & + \frac{\|\Delta g_B\|}{M-1} \sum_{n \neq m} \mathbf{E} \left[\sum_{k=1}^{N_{n,j}} \mathbf{E} \left[|U_{k,(n,j) \rightarrow (m,i)} - V_{k,(n,j) \rightarrow (m,i)}| \middle| N \right] \right],
 \end{aligned}$$

where for a function f , we denote $\|f\| = \sup_{t \in [0, T]} |f(t)|$. Now, take

$$U_{k, (n, j) \rightarrow (m, i)} = A_{j \rightarrow (m, i)} \text{ and } V_{k, (n, j) \rightarrow (m, i)} = \sum_{l \neq k} B_{l, (n, j) \rightarrow (m, i)}.$$

Then,

$$|U_{k, (n, j) \rightarrow (m, i)} - V_{k, (n, j) \rightarrow (m, i)}| = B_{k, (n, j) \rightarrow (m, i)}.$$

Therefore, using once again Lemma 3.2.7,

$$\mathbf{E} [|U_{k, (n, j) \rightarrow (m, i)} - V_{k, (n, j) \rightarrow (m, i)}| | N] = \frac{1}{M-1}.$$

Moreover, it can be shown (see [10]) that $\|g_B\| \leq 1 \wedge \frac{0.74}{\mathbf{E}[N_{1, j}]}$ and $\|\Delta g_B\| \leq 1 \wedge \frac{1}{\mathbf{E}[N_{1, j}]}$, where for $k \in \mathbb{N}$, $\Delta g_B(k) = g_B(k+1) - g_B(k)$. Combining this yields (3.17). \square

Decoupling arrivals and outputs: a fixed point scheme approach

As we have seen in the previous lemma, for the Poisson approximation to hold, it is sufficient to prove a law of large numbers-type result on the random variables $(N_{n, j})_{n \neq m}$. However, since these random variables themselves depend on the random variables $(A_{j \rightarrow (m, i)})_{m \in \{1, \dots, M\}}$, a direct proof seems difficult to obtain.

As such, we propose to see Equation (3.4) as the fixed point equation of some function on the space of probability laws on the space of càdlàg trajectories. This fixed point exists and is necessarily unique due to the fact that Equation (3.4) admits a unique solution. The main idea goes as follows: if we endow this space with a metric that makes it complete, in order to prove that the law of large numbers holds at the fixed point, it is sufficient to show that, on one hand, if this law of large numbers holds for a given probability law, it also holds for its image by the function; and that on the other hand, the function's iterates form a Cauchy sequence. This approach is similar to the one developed in [27], where propagation of chaos is proven in discrete time by showing that the one-step transition of the discrete dynamics preserves a triangular law of large numbers.

Our goal in this section is to prove the two aforementioned points. We start by introducing the metric space we will be considering and defining the function on it.

Fix $T \in \mathbb{R}$, and let D_T be the space of càdlàg functions on $[0, T]$ endowed with the Billingsley metric [12]: for $x, y \in D_T$, let

$$d_{D_T}(x, y) = \inf_{\theta \in \Theta} \max(\|\theta\|, \|x - y \circ \theta\|),$$

where

$$\Theta = \{\theta : [0, T] \rightarrow [0, T], \text{ s.t. } \theta(0) = 0, \theta(T) = T, \text{ and } \|\theta\| < \infty\},$$

where

$$|||\theta||| = \sup_{s \neq t \in [0, T]} \left| \log \left(\frac{\theta(t) - \theta(s)}{t - s} \right) \right|.$$

Intuitively, Θ represents all possible "reasonable" time shifts allowing one to minimize the effect of the jumps between the two functions x and y , where "reasonable" means that all slopes of θ are close to 1.

We denote by $d_{D_T, U}$ the uniform metric on D_T : for $x, y \in D_T$,

$$d_{D_T, U}(x, y) = \|x - y\|.$$

Note that we have for all $x, y \in D_T$, $d_{D_T}(x, y) \leq d_{D_T, U}(x, y)$, since the uniform metric corresponds precisely to the case where θ is the identity function.

Let $\mathcal{P}(D_T)$ be the space of probability measures on D_T . We endow it with the Kantorovitch metric [46] (sometimes also known as the Wasserstein distance or the earth mover's distance): for $\mu, \nu \in \mathcal{P}(D_T)$, let

$$K_T(\mu, \nu) = \inf_{\Pi \in D_T \times D_T} \mathbf{E}[d_{D_T}(x, y)],$$

where Π is a coupling s.t. $x \stackrel{\mathcal{L}}{=} \mu$ and $y \stackrel{\mathcal{L}}{=} \nu$.

Finally, we fix $K, M \in \mathbb{N}$ and consider the space $(\mathcal{P}(D_T))^{MK}$ endowed with the 1-norm metric: for $\mu, \nu \in (\mathcal{P}(D_T))^{MK}$, let

$$K_T^{MK}(\mu, \nu) = \sum_{m=1}^M \sum_{i=1}^K K_T(\mu_{m,i}, \nu_{m,i}).$$

It is known that (D_T, d_{D_T}) is a complete separable metric space, see [12], and thus that $(\mathcal{P}(D_T), K_T)$ and $(\mathcal{P}(D_T))^{MK}, K_T^{MK}$ are as well, see [14].

We will also need to consider $\mathcal{P}(D_T)$ endowed with a Kantorovitch metric based on the uniform metric: we introduce for $\mu, \nu \in \mathcal{P}(D_T)$,

$$K_{T, U}(\mu, \nu) = \inf_{\Pi \in D_T \times D_T} \mathbf{E}[d_{D_T, U}(x, y)],$$

where Π is a coupling s.t. $x \stackrel{\mathcal{L}}{=} \mu$ and $y \stackrel{\mathcal{L}}{=} \nu$. We also introduce its product version $K_{T, U}^{MK}$ defined analogously to above. Note that even though $(D_T, d_{D_T, U})$ is a complete metric space, it is not separable, therefore $(\mathcal{P}(D_T), K_{T, U})$ is not a priori a complete metric space.

We now define the following mapping:

$$\begin{aligned} \Phi: (\mathcal{P}(D_T))^{MK} &\rightarrow (\mathcal{P}(D_T))^{MK} \\ (\mathcal{L}(M)) &\mapsto \Phi(\mathcal{L}(M)), \end{aligned}$$

where for all $(m, i) \in \{1, \dots, M\} \times \{1, \dots, K\}$, $\Phi(\mathcal{L}(M))_{m,i}$ is the law of the stochastic intensity $\lambda_{m,i}^\Phi$ of a point process $N_{m,i}^\Phi$ such that λ^Φ is the solution of the stochastic differential equation

$$\begin{aligned} \lambda_{m,i}^\Phi(t) &= \lambda_{m,i}^\Phi(0) + \sum_{j \neq i} \mu_{j \rightarrow i} \sum_{n \neq m} \int_0^t \mathbb{I}_{\{V_{(n,j)}^M \rightarrow i(s)=m\}} M_{n,j}(ds) \\ &+ \int_0^t (r_i - \lambda_{m,i}^\Phi(s)) N_{m,i}^\Phi(ds), \end{aligned} \quad (3.18)$$

where $(\lambda_{m,i}^\Phi(0))$ are random variables verifying Assumption 3.2.2. Note that we will exclusively apply the mapping Φ to laws of stochastic intensities of point processes, the image of which by Φ are also by definition of Φ laws of stochastic intensities of point processes.

We formalize the law of large numbers we aim to prove as follows:

Definition 3.2.15. *Let $M \in \mathbb{N}$. Let $(X_n^M)_{1 \leq n \leq M}$ be exchangeable random variables with finite expectation. We say they satisfy an \mathcal{L}^1 triangular law of large numbers, which we denote $\text{TLLN}(X_n^M)$, if when $M \rightarrow \infty$,*

$$\mathbf{E} \left[\left| \frac{1}{M-1} \sum_{n=1}^M (X_n^M - \mathbf{E}[X_n^M]) \right| \right] \rightarrow 0 \quad (3.19)$$

and

$$X_n^M \Rightarrow \tilde{X}, \quad (3.20)$$

where the convergence takes place in distribution.

From (3.17), we know that if the triangular law of large numbers holds for the fixed point of Φ , it allows for convergence in total variation of arrivals across replicas from a given neuron j to a given neuron i to a Poisson random variable. As such, our aim here is twofold:

1. Show that for all $(m, i) \in \{1, \dots, M\} \times \{1, \dots, K\}$, $\text{TLLN}(\overline{N_{m,i}}([0, t]))$ implies $\text{TLLN}(\Phi(\overline{N_{m,i}}([0, t])))$;
2. Show that $(\Phi^l)_{l \in \mathbb{N}^*}$ is a Cauchy sequence that converges to the fixed point, where Φ^l is the l -th iterate of $\Phi : \Phi^l = \Phi \circ \Phi \circ \dots \circ \Phi$ l times.

Since we can choose $\overline{N_{m,i}}([0, t])$ to be i.i.d. to ensure that there exist inputs for which TLLN holds, this will allow us to propagate the property and show that TLLN holds at the fixed point as well.

We will start by proving a lemma that will be key for the second point:

Lemma 3.2.16. *There exists $T > 0$ such that for $\rho, \nu \in (\mathcal{P}(D_T))^{MK}$ that are laws of stochastic intensities of point processes, there exists a constant $C_T > 0$ such that*

$$K_{T,U}^{MK}(\Phi(\rho), \Phi(\nu)) \leq C_T \int_0^T K_{t,U}^{MK}(\rho, \nu) dt. \quad (3.21)$$

Proof. Let $T > 0$. Let $t \in [0, T]$. Fix $(m, i) \in \{1, \dots, M\} \times \{1, \dots, K\}$. Let N^ρ (resp. $N^\nu, N^{\Phi(\rho)}, N^{\Phi(\nu)}$) be a point process admitting ρ (resp. $\nu, \Phi(\rho), \Phi(\nu)$) as a stochastic intensity. We have

$$\begin{aligned} \Phi(\rho)_{m,i}(t) - \Phi(\nu)_{m,i}(t) &= \sum_{j \neq i} \mu_{j \rightarrow i} \sum_{n \neq m} \left(\int_0^t \mathbb{1}_{\{V_{(n,j) \rightarrow i}^M(s) = m\}} (N_{n,j}^\rho(ds) - N_{n,j}^\nu(ds)) \right) \\ &\quad + \int_0^t (r_i - \Phi(\rho)_{m,i}(s)) N_{m,i}^{\Phi(\rho)}(ds) \\ &\quad - \int_0^t (r_i - \Phi(\nu)_{m,i}(s)) N_{m,i}^{\Phi(\nu)}(ds). \end{aligned}$$

Let $(\hat{N}_{m,i})_{(m,i) \in \{1, \dots, M\} \times \{1, \dots, K\}}$ be independent Poisson point processes with intensity 1 on $[0, T] \times \mathbb{R}^+$. Using the Poisson embedding construction, we can write

$$\begin{aligned} \Phi(\rho)_{m,i}(t) - \Phi(\nu)_{m,i}(t) &= \\ &\sum_{j \neq i} \mu_{j \rightarrow i} \sum_{n \neq m} \int_0^t \int_0^{+\infty} \mathbb{1}_{\{V_{(n,j) \rightarrow i}^M(s) = m\}} (\mathbb{1}_{\{u \leq \rho_{n,j}(s)\}} - \mathbb{1}_{\{u \leq \nu_{n,j}(s)\}}) \hat{N}_{n,j}(ds du) \\ &\quad + r_i \int_0^t \int_0^{+\infty} (\mathbb{1}_{\{u \leq \Phi(\rho)_{m,i}(s)\}} - \mathbb{1}_{\{u \leq \Phi(\nu)_{m,i}(s)\}}) \hat{N}_{m,i}(ds du) \\ &\quad + \int_0^t \int_0^{+\infty} (\Phi(\nu)_{m,i}(s) \mathbb{1}_{\{u \leq (\Phi(\nu)_{m,i}(s))\}} - \Phi(\rho)_{m,i}(s) \mathbb{1}_{\{u \leq (\Phi(\rho)_{m,i}(s))\}}) \hat{N}_{m,i}(ds du). \end{aligned}$$

Therefore, we have

$$\begin{aligned} |\Phi(\rho)_{m,i}(t) - \Phi(\nu)_{m,i}(t)| &\leq \\ &\sum_{j \neq i} \mu_{j \rightarrow i} \sum_{n \neq m} \int_0^t \int_0^{+\infty} \mathbb{1}_{\{V_{(n,j) \rightarrow i}^M(s) = m\}} \mathbb{1}_{\{u \leq \sup_{z \in [0, s]} |\rho_{n,j}(z) - \nu_{n,j}(z)|\}} \hat{N}_{n,j}(ds du) \\ &\quad + r_i \int_0^t \int_0^{+\infty} \mathbb{1}_{\{u \leq \sup_{z \in [0, s]} |\Phi(\rho)_{m,i}(z) - \Phi(\nu)_{m,i}(z)|\}} \hat{N}_{m,i}(ds du) \\ &\quad + \int_0^t \int_0^{+\infty} \sup_{z \in [0, s]} |\Phi(\rho)_{m,i}(z) - \Phi(\nu)_{m,i}(z)| \mathbb{1}_{\{u \leq \Phi(\rho)_{m,i}(s) \wedge \Phi(\nu)_{m,i}(s)\}} \hat{N}_{m,i}(ds du) \\ &\quad + \int_0^t \int_0^{+\infty} |\Phi(\rho)_{m,i}(s) \vee \Phi(\nu)_{m,i}(s)| \mathbb{1}_{\{u \leq \sup_{z \in [0, s]} |\Phi(\rho)_{m,i}(z) - \Phi(\nu)_{m,i}(z)|\}} \hat{N}_{m,i}(ds du). \end{aligned}$$

Taking the expectation, using the property of stochastic intensity and proceeding as before to obtain the $\frac{1}{M-1}$ from the routing indicators, we get

$$\begin{aligned}
& \mathbf{E} \left[\sup_{t \in [0, T]} |\Phi(\rho)_{m,i}(t) - \Phi(\nu)_{m,i}(t)| \right] \leq \\
& \frac{1}{M-1} \sum_{j \neq i} \mu_{j \rightarrow i} \sum_{n \neq m} \int_0^T \mathbf{E} \left[\sup_{z \in [0, s]} |\rho_{n,j}(z) - \nu_{n,j}(z)| \right] ds \\
& + r_i \int_0^T \mathbf{E} \left[\sup_{z \in [0, s]} |\Phi(\rho)_{m,i}(z) - \Phi(\nu)_{m,i}(z)| \right] ds \\
& + \int_0^T \mathbf{E} \left[\sup_{z \in [0, s]} |\Phi(\rho)_{m,i}(z) - \Phi(\nu)_{m,i}(z)| (\Phi(\rho)_{m,i}(s) \wedge \Phi(\nu)_{m,i}(s)) \right] ds \\
& + \int_0^T \mathbf{E} \left[\sup_{z \in [0, s]} |\Phi(\rho)_{m,i}(z) - \Phi(\nu)_{m,i}(z)| (\Phi(\rho)_{m,i}(s) \vee \Phi(\nu)_{m,i}(s)) \right] ds.
\end{aligned}$$

Denote $\|\mu\| = \max_{i,j} \mu_{j \rightarrow i}$. We then have

$$\begin{aligned}
& \frac{1}{M-1} \sum_{j \neq i} \mu_{j \rightarrow i} \sum_{n \neq m} \int_0^T \mathbf{E} \left[\sup_{z \in [0, s]} |\rho_{n,j}(z) - \nu_{n,j}(z)| \right] ds \leq \\
& \|\mu\| \sum_{j=1}^K \sum_{n=1}^M \int_0^T d_{D_s, U}(\rho_{n,j}, \nu_{n,j}) ds,
\end{aligned}$$

from which we immediately get by definition of $K_{T,U}^{MK}$

$$\frac{1}{M-1} \sum_{j \neq i} \mu_{j \rightarrow i} \sum_{n \neq m} \int_0^T \mathbf{E} \left[\sup_{z \in [0, s]} |\rho_{n,j}(z) - \nu_{n,j}(z)| \right] ds \leq \|\mu\| \int_0^T K_{s,U}^{MK}(\rho, \nu) ds. \tag{3.22}$$

Let $C > 0$. As before, let $A_C([0, T]) = \{(\omega, t) \in \Omega \times [0, T], \Phi(\rho)_{m,i}(t) \vee \Phi(\nu)_{m,i}(t) > C\}$. Using the exact same reasoning as in Lemma 3.2.11, we have for small enough T the existence of a constant $K_T > 0$ such that

$$\begin{aligned}
& \int_0^t \mathbf{E} \left[\sup_{z \in [0, s]} |\Phi(\rho)_{m,i}(z) - \Phi(\nu)_{m,i}(z)| (\Phi(\rho)_{m,i}(s) \wedge \Phi(\nu)_{m,i}(s)) \right] ds \\
& \leq C \int_0^t \mathbf{E} \left[\sup_{z \in [0, s]} |\Phi(\rho)_{m,i}(z) - \Phi(\nu)_{m,i}(z)| \right] ds + K_T e^{-3CT}.
\end{aligned}$$

Plugging in (3.22) and applying the same reasoning as above to the last integral term, we get the existence of a constant $K'_T > 0$ such that

$$\begin{aligned} \mathbf{E} \left[\sup_{t \in [0, T]} |\Phi(\rho)_{m,i}(t) - \Phi(\nu)_{m,i}(t)| \right] &\leq \|\mu\| \int_0^T K_{s,U}^{MK}(\rho, \nu) \, ds \\ &+ (r_i + 2C) \int_0^T \mathbf{E} \left[\sup_{z \in [0, s]} |\Phi(\rho)_{m,i}(z) - \Phi(\nu)_{m,i}(z)| \right] \, ds \\ &+ (K_T + K'_T) e^{-3CT}. \end{aligned}$$

Applying Grönwall's lemma, we get

$$\begin{aligned} \mathbf{E} \left[\sup_{t \in [0, T]} |\Phi(\rho)_{m,i}(t) - \Phi(\nu)_{m,i}(t)| \right] &\leq \\ &\left(\|\mu\| \int_0^T K_{s,U}^{MK}(\rho, \nu) \, ds + (K_T + K'_T) e^{-3CT} \right) e^{(r_i + 2C)T}. \end{aligned}$$

For any $\varepsilon > 0$, we can choose $C > 0$ such that

$$\mathbf{E} \left[\sup_{t \in [0, T]} |\Phi(\rho)_{m,i}(t) - \Phi(\nu)_{m,i}(t)| \right] \leq \left(\|\mu\| \int_0^T K_{s,U}^{MK}(\rho, \nu) \, ds \right) e^{(r_i + 2C)T} + \varepsilon.$$

Letting ε go to 0 and taking the sum over all coordinates and the infimum across all couplings, we get the result. \square

As previously mentioned, we need to prove convergence of the sequence of iterates of Φ to the fixed point of Φ to prove the triangular law of large numbers. We will now derive this from Lemma 3.2.16.

Lemma 3.2.17. *Let $\rho \in (\mathcal{P}(D_T))^{MK}$ be the law of the stochastic intensity of a point process. The sequence $(\Phi^l(\rho))_{l \in \mathbb{N}^*}$ of iterates of the function Φ is a Cauchy sequence. Moreover, it converges to the unique fixed point of Φ .*

Proof. Let $\rho \in (\mathcal{P}(D_T))^{MK}$ be the law of the stochastic intensity of a point process. By immediate induction, from (3.21), we have, for all $l \in \mathbb{N}^*$,

$$K_{T,U}^{MK}(\Phi^{l+1}(\rho), \Phi^l(\rho)) \leq C_T \frac{T^l}{l!} K_{T,U}^{MK}(\Phi(\rho), \rho).$$

This in turn implies that for any $p < q \in \mathbb{N}^*$,

$$K_{T,U}^{MK}(\Phi^p(\rho), \Phi^q(\rho)) \leq \sum_{l=p}^{q-1} \frac{C_T T^l}{l!} K_{T,U}^{MK}(\Phi(\rho), \rho). \quad (3.23)$$

Since the series on the right hand side is converging, it proves that the sequence $(\Phi^l)_{l \in \mathbb{N}^*}$ is a Cauchy sequence for the $K_{T,U}^{MK}$ metric. The space $(\mathcal{P}(D_T))^{MK}, K_{T,U}^{MK}$ is not complete. However, since for any $\mu, \nu \in D_T, d_{D_T}(\mu, \nu) \leq d_{D_T,U}(\mu, \nu)$, it follows that $(\Phi^l)_{l \in \mathbb{N}^*}$ is a Cauchy sequence for the K_T^{MK} metric as well. By completeness of $(\mathcal{P}(D_T))^{MK}, K_T^{MK}$, $(\Phi^l)_{l \in \mathbb{N}^*}$ converges to the unique fixed point of Φ . \square

All that remains is proving that the triangular law of large numbers is carried over by the function Φ , namely, that if we have some input X that verifies TLLN(X), then we have TLLN($\Phi(X)$).

To do so, the key lemma will be the following law of large numbers:

Lemma 3.2.18. *Let $M \in \mathbb{N}^*$. Let (X_1^M, \dots, X_M^M) be M -exchangeable centered random variables with finite exponential moments. Suppose that for any $N \in \mathbb{N}^*$, $(X_1^M, \dots, X_N^M) \xrightarrow{\mathcal{L}} (\tilde{X}_1, \dots, \tilde{X}_N)$ when $M \rightarrow \infty$, where $(\tilde{X}_i)_{i \in \mathbb{N}^*}$ are i.i.d. random variables and the convergence takes place in distribution. Then*

$$\mathbf{E} \left[\left[\frac{1}{M} \sum_{n=1}^M X_n^M \right] \right] \rightarrow 0 \tag{3.24}$$

when $M \rightarrow \infty$.

Proof. Let

$$U_M = \frac{1}{M} \sum_{n=1}^M X_n^M.$$

Note that $\mathbf{E}[U_M] = 0$. We have

$$\mathbf{E}[U_M^2] = \frac{1}{M^2} \mathbf{E} \left[\sum_{n=1}^M (X_n^M)^2 + \sum_{m \neq n} X_m^M X_n^M \right].$$

Since the exponential moments of (X_n^M) are bounded and they all converge in distribution, we have by asymptotic independence that for any m, n ,

$$\mathbf{E}[X_m^M X_n^M] \rightarrow \mathbf{E}[\tilde{X}_m \tilde{X}_n] = \mathbf{E}[\tilde{X}_m] \mathbf{E}[\tilde{X}_n] = 0$$

when $M \rightarrow \infty$.

Therefore, for $\varepsilon > 0$, for large enough M , each of the terms of the corresponding sums in the equation above is smaller than ε . Combining this with exchangeability, we get the existence of a positive constant C s.t. for large enough M ,

$$\mathbf{E}[U_M^2] \leq \frac{1}{M^2} (CM + M(M-1)\varepsilon).$$

Now, applying Chebychev's inequality, for any $\delta > 0$,

$$\mathbf{P}(|U_M - \mathbf{E}[U_M]| > \delta) \leq \frac{\mathbf{E}[U_M^2]}{\delta^2} \leq \frac{C}{M\delta^2} + \frac{M(M-1)\varepsilon}{M^2\delta^2}.$$

This gives convergence in probability of U_M to 0 when $M \rightarrow \infty$.

Since in addition the second moments are uniformly bounded, \mathcal{L}^1 convergence follows. \square

Now, we can finally prove the following lemma which is the last step needed to prove the main theorem:

Lemma 3.2.19. *Let $(\overline{N_{m,i}})$ be point processes on $[0, T]$ with finite exponential moments. Let $t \in [0, T]$. Suppose $\text{TLLN}((\overline{N_{m,i}}([0, t])))$ holds. Then, $\text{TLLN}(\Phi((\overline{N_{m,i}}([0, t])))$ holds as well.*

Proof. Suppose $\text{TLLN}((\overline{N_{m,i}}([0, t])))$ holds. For $(m, i) \in \{1, \dots, M\} \times \{1, \dots, K\}$, let $\lambda_{m,i}$ be the stochastic intensity of the process $\overline{N_{m,i}}$. We write for all $t \in [0, T]$,

$$\lambda_{m,i}(t) = \lambda_{m,i}(0) + \sum_{j \neq i} \mu_{j \rightarrow i} \overline{A_{j \rightarrow (m,i)}}(t) + \int_0^t (r_i - \lambda_{m,i}(s)) N_{m,i}(ds),$$

where $(\lambda_{m,i}(0))$ verifies Assumption 3.2.2 and

$$\overline{A_{j \rightarrow (m,i)}}(t) = \sum_{n \neq m} \int_0^t \mathbb{1}_{\{V_{(n,j) \rightarrow i}^M(s) = m\}} \overline{N_{n,j}}(ds).$$

Analogously to (3.17), we have

$$\begin{aligned} d_{TV}(\overline{A_{j \rightarrow (m,i)}}(t), \tilde{A}_{j \rightarrow i}(t)) &\leq \\ &\left(1 \wedge \frac{0.74}{\sqrt{\mathbf{E}[\overline{N_{1,j}}([0, T])]}}\right) \frac{1}{M-1} \mathbf{E} \left[\left| \sum_{n \neq m} (\mathbf{E}[\overline{N_{n,j}}([0, T])] - \overline{N_{n,j}}([0, T])) \right| \right] \\ &+ \frac{1}{M-1} \left(1 \wedge \frac{1}{\mathbf{E}[\overline{N_{1,j}}([0, T])]}\right) \mathbf{E}[\overline{N_{1,j}}([0, T])]. \end{aligned}$$

where $\tilde{A}_{j \rightarrow i}$ are independent Poisson random variables with mean $\mathbf{E}[\overline{N_{1,j}}([0, T])]$.

As such, $\text{TLLN}((\overline{N_{m,i}}([0, t])))$ implies convergence in total variation of $(\overline{A_{j \rightarrow (m,i)}})$ to independent random variables $(\tilde{A}_{j \rightarrow i})$. We will now show that this implies convergence in total variation of $\sum_{j \neq i} \mu_{j \rightarrow i} \overline{A_{j \rightarrow (m,i)}}$.

Denote as before $\bar{N} = (\bar{N}_{n,j})_{n \neq m, j \neq i}$. Let $q \in \mathbb{N}^{(M-1)(K-1)}$. Let $B_1, B_2 \in \mathcal{B}(\mathbb{R}^+)$. Let $l_1 \neq l_2 \in \{1, \dots, K\} \setminus \{i\}$. Then, using the total probability formula, we have

$$\begin{aligned} \mathbf{P}(\bar{A}_{l_1 \rightarrow (m,i)} \in B_1, \bar{A}_{l_2 \rightarrow (m,i)} \in B_2) &= \\ \sum_q \mathbf{P}(\bar{A}_{l_1 \rightarrow (m,i)} \in B_1, \bar{A}_{l_2 \rightarrow (m,i)} \in B_2 | \bar{N} = q) \mathbf{P}(\bar{N} = q). \end{aligned}$$

Using Lemma 3.2.7, by conditional independence, we have

$$\begin{aligned} \mathbf{P}(\bar{A}_{l_1 \rightarrow (m,i)} \in B_1, \bar{A}_{l_2 \rightarrow (m,i)} \in B_2) &= \\ \sum_q \mathbf{P}(\bar{A}_{l_1 \rightarrow (m,i)} \in B_1 | \bar{N} = q) \mathbf{P}(\bar{A}_{l_2 \rightarrow (m,i)} \in B_2 | \bar{N} = q) \mathbf{P}(\bar{N} = q). \end{aligned}$$

Using the same reasoning as in Lemma 3.2.14, we have that:

$$\mathbf{P}(\bar{A}_{l_1 \rightarrow (m,i)} \in B_1 | \bar{N} = q) \rightarrow \mathbf{P}(\tilde{A}_{l_1 \rightarrow i} \in B_1 | \bar{N} = q) = \mathbf{P}(\tilde{A}_{l_1 \rightarrow i} \in B_1),$$

$$\mathbf{P}(\bar{A}_{l_2 \rightarrow (m,i)} \in B_2 | \bar{N} = q) \rightarrow \mathbf{P}(\tilde{A}_{l_2 \rightarrow i} \in B_2 | \bar{N} = q) = \mathbf{P}(\tilde{A}_{l_2 \rightarrow i} \in B_2)$$

and

$$\mathbf{P}(\bar{N} = q) \rightarrow \mathbf{P}(\tilde{N} = q).$$

By dominated convergence,

$$\begin{aligned} \mathbf{P}(\bar{A}_{l_1 \rightarrow (m,i)} \in B_1, \bar{A}_{l_2 \rightarrow (m,i)} \in B_2) &\rightarrow \sum_q \mathbf{P}(\tilde{A}_{l_1 \rightarrow i} \in B_1) \mathbf{P}(\tilde{A}_{l_2 \rightarrow i} \in B_2) \mathbf{P}(\bar{N} = q) \\ &= \mathbf{P}(\tilde{A}_{l_1 \rightarrow (m,i)} \in B_1, \tilde{A}_{l_2 \rightarrow (m,i)} \in B_2). \end{aligned}$$

This implies convergence of $\sum_j \mu_{j \rightarrow i} \bar{A}_{j \rightarrow (m,i)}$. Finally, the mapping theorem implies convergence in total variation of $\lambda_{m,i}(t)$ when $M \rightarrow \infty$. All conditions of Lemma 3.2.18 are thus satisfied. Applying it completes the proof. \square

Thus, we can now state the result that we were aiming to prove:

Lemma 3.2.20. *Denote by $(N_{m,i})$ the point processes of the M -replica RMF dynamics (3.4) that are the fixed point of Φ . Then TLLN($(N_{m,i}([0, T]))$) holds.*

Proof. Let $(\overline{N}_{n,j})$ be random variables satisfying TLLN($\overline{N}_{n,j}$). Let us first write out equalities and justify them afterwards. We have

$$\begin{aligned}
 & \lim_{M \rightarrow \infty} \frac{1}{M-1} \mathbf{E} \left[\left| \sum_{n \neq m} (\mathbf{E}[N_{n,j}([0, T])] - N_{n,j}([0, T])) \right| \right] \\
 &= \lim_{M \rightarrow \infty} \frac{1}{M-1} \mathbf{E} \left[\left| \sum_{n \neq m} \left(\mathbf{E}[\lim_{l \rightarrow \infty} \Phi^l(\overline{N}_{n,j}([0, T]))] - \lim_{l \rightarrow \infty} \Phi^l(\overline{N}_{n,j}([0, T])) \right) \right| \right] \\
 &= \lim_{M \rightarrow \infty} \lim_{l \rightarrow \infty} \frac{1}{M-1} \mathbf{E} \left[\left| \sum_{n \neq m} (\mathbf{E}[\Phi^l(\overline{N}_{n,j}([0, T]))] - \Phi^l(\overline{N}_{n,j}([0, T]))) \right| \right] \\
 &= \lim_{l \rightarrow \infty} \lim_{M \rightarrow \infty} \frac{1}{M-1} \mathbf{E} \left[\left| \sum_{n \neq m} (\mathbf{E}[\Phi^l(\overline{N}_{n,j}([0, T]))] - \Phi^l(\overline{N}_{n,j}([0, T]))) \right| \right] \\
 &= 0.
 \end{aligned}$$

The first equality is given by Lemma 3.2.17. To obtain the second equality, we use the dominated convergence theorem and the fact that all moments all uniformly bounded through Lemma 3.2.8 (note that initial conditions are fixed in the definition of Φ and are chosen to verify Assumption 3.2.2). To justify the third equality, note that from (3.23), using Lemma 3.2.8 to obtain again a uniform bound of the moments, we get that the Cauchy sequence of iterates of Φ verifies the uniform Cauchy criterion and thus converges uniformly to the fixed point, which in turn allows for the exchange of limits in M and l . The last equality stems directly from Lemma 3.2.19. \square

3.2.3 Tightness and convergence on \mathbb{R}

The goal of this section is to generalize the main result of the paper. In Theorem 3.2.4, we proved weak convergence of the replica-mean-field processes on compacts of \mathbb{R}^+ . We now prove weak convergence on \mathbb{R}^+ . One motivation for doing so is that the results on the Galves-Löcherbach replica-mean-field model in the paper [8] by Baccelli and Taillefumier assumes that the Poisson Hypothesis holds at the limit in the stationary regime. As such, the following result provides the missing rigorous justification, albeit for a slightly simplified model due to Assumption 3.2.1.

Theorem 3.2.21. *Let $K, M \geq 2$. For all $m \in \{1, \dots, M\}$ and $i \in \{1, \dots, K\}$, the process $\lambda_{m,i}$ weakly converges in the Skorokhod space of càdlàg functions on \mathbb{R}^+ .*

Recall the following tightness criterion due to Aldous [2]:

Proposition 3.2.1. *Let $(\Omega, \mathcal{F}, (\mathcal{F}_t^{(n)}, t \in \mathbb{R}^+), \mathbf{P})$ be a probability space, let $X^{(n)}$ be adapted càdlàg processes. If for all $T > 0$, $(\mathcal{L}(\|X^{(n)}\|))$ is tight on $[0, T]$, and if for all $\varepsilon > 0$, for all $\varepsilon' > 0$, there exists $\delta \in (0, T]$ such that*

$$\limsup_{n \rightarrow +\infty} \sup_{\substack{(S_1, S_2 \in \mathcal{F}_t^{(n)}) \text{ such that} \\ S_1 \leq S_2 \leq (S_1 + \delta) \wedge T}} \mathbf{P}(|X_{S_1}^{(n)} - X_{S_2}^{(n)}| > \varepsilon') \leq \varepsilon,$$

then $(\mathcal{L}(X^{(n)}))$ is tight on the space of càdlàg functions on \mathbb{R}^+ .

We will also require the following inequality on martingales:

Lemma 3.2.22. *Let $(X_t)_{t \in [0, T]}$ be a nonnegative (\mathcal{F}_t) -martingale, let S_1, S_2 be two stopping times such that $S_1 \leq S_2 \leq S_1 + \delta$. Then*

$$\mathbf{E} \left[\int_{S_1}^{S_2} X_s \, ds \right] \leq 2\delta \sqrt{\mathbf{E}[X_T^2]}. \quad (3.25)$$

Proof. We have

$$\begin{aligned} \mathbf{E} \left[\int_{S_1}^{S_2} X_s \, ds \right] &\leq \delta \mathbf{E} \left[\sup_{s \in [0, T]} X_s \right] \\ &\leq \delta \sqrt{\mathbf{E} \left[\sup_{s \in [0, T]} X_s^2 \right]} \\ &\leq 2\delta \sqrt{\mathbf{E}[X_T^2]} \text{ (by the Doob inequality)} \end{aligned}$$

□

We are now ready to prove Theorem 3.2.21.

Proof. We use the Aldous criterion mentioned above to prove tightness. Let $T > 0$. Let $\varepsilon' > 0, \delta > 0, S_1, S_2$ two stopping times such that $S_1 \leq S_2 \leq (S_1 + \delta) \wedge T$.

Using the Markov inequality, the fact that $r_i - \lambda_{m,i}$ is negative, and the property of stochastic intensity, we have:

$$\begin{aligned}
 & \mathbf{P} (|\lambda_{m,i}(S_2) - \lambda_{m,i}(S_1)| > \varepsilon') = \\
 & \mathbf{P} \left(\left| \sum_{j \neq i} \mu_{j \rightarrow i} \sum_{n \neq m} \int_{S_1}^{S_2} \mathbb{I}_{\{V_{(n,j) \rightarrow i}(s)=m\}} N_{n,j}(ds) + \int_{S_1}^{S_2} (r_i - \lambda_{m,i}(s)) N_{m,i}(ds) \right| > \varepsilon' \right) \\
 & \leq \frac{1}{\varepsilon'} \mathbf{E} \left[\sum_{j \neq i} \mu_{j \rightarrow i} \sum_{n \neq m} \int_{S_1}^{S_2} \mathbb{I}_{\{V_{(n,j) \rightarrow i}(s)=m\}} N_{n,j}(ds) + \int_{S_1}^{S_2} (\lambda_{m,i}(s) - r_i) N_{m,i}(ds) \right] \\
 & \leq \frac{1}{\varepsilon'} \mathbf{E} \left[\int_{S_1}^{S_2} \sum_{n \neq m} \sum_{j \neq i} \mu_{j \rightarrow i} \mathbb{I}_{\{V_{(n,j) \rightarrow i}(s)=m\}} \lambda_{n,j}(s) + (\lambda_{m,i}^2(s) - r_i \lambda_{m,i}(s)) ds \right] \\
 & \leq \frac{1}{\varepsilon'} \mathbf{E} \left[\int_{S_1}^{S_2} \sum_{j \neq i} \mu_{j \rightarrow i} \lambda_{1,j}(s) + \lambda_{m,i}^2(s) - r_i \lambda_{m,i}(s) ds \right],
 \end{aligned}$$

Since $\lambda_{m,i}(s)$ is non-negative for all $s \in [0, T]$, we can write

$$\mathbf{P} (|\lambda_{m,i}(S_2) - \lambda_{m,i}(S_1)| > \varepsilon') \leq \frac{1}{\varepsilon'} \mathbf{E} \left[\int_{S_1}^{S_2} \sum_{j \neq i} \mu_{j \rightarrow i} \lambda_{1,j}(s) + \lambda_{m,i}^2(s) ds \right].$$

For all $(m, i) \in \{1, \dots, M\} \times \{1, \dots, K\}$, let

$$c_{m,i}(t) = \sum_{n \neq m} \sum_{j \neq i} \int_0^t \mathbb{I}_{\{V_{(n,j) \rightarrow i}(s)=m\}} \lambda_{n,j}(s) ds + \int_0^t (r_i - \lambda_{m,i}(s)) \lambda_{m,i}(s) ds,$$

and

$$\begin{aligned}
 d_{m,i}(t) &= \sum_{n \neq m} \sum_{j \neq i} \int_0^t \mathbb{I}_{\{V_{(n,j) \rightarrow i}(s)=m\}} (2\lambda_{n,j}(s) + 1) \lambda_{n,j}(s) ds \\
 &\quad + \int_0^t (r_i^2 - \lambda_{m,i}^2(s)) \lambda_{m,i}(s) ds.
 \end{aligned}$$

By the property of stochastic intensity, $\rho_{m,i}(t) = \lambda_{m,i}(t) - c_{m,i}(t)$ and $\nu_{m,i}(t) = \lambda_{m,i}^2(t) - d_{m,i}(t)$ are (\mathcal{F}_t) -martingales.

Therefore, we can write

$$\begin{aligned}
 & \mathbf{P} (|\lambda_{m,i}(S_2) - \lambda_{m,i}(S_1)| > \varepsilon') \leq \\
 & \frac{1}{\varepsilon'} \mathbf{E} \left[\int_{S_1}^{S_2} \sum_{j \neq i} \mu_{j \rightarrow i} \rho_{1,j}(s) + \nu_{m,i}(s) + \sum_{j \neq i} \mu_{j \rightarrow i} c_{1,j}(s) + d_{m,i}(s) ds \right]. \tag{3.26}
 \end{aligned}$$

We now bound separately the martingale part and the rest of the right-hand-side expression in (3.26).

Using Lemma 3.2.22, we have that

$$\mathbf{E} \left[\int_{S_1}^{S_2} \left(\sum_{j \neq i} \mu_{j \rightarrow i} \rho_{1,j}(s) + \nu_{m,i}(s) \right) ds \right] \leq 2\delta \sqrt{\mathbf{E}[(\sum_{j \neq i} \mu_{j \rightarrow i} \rho_{1,j}(T) + \nu_{m,i}(T))^2]}.$$

Therefore, using the Cauchy-Schwarz inequality and Lemma 3.2.8, there exists a nonnegative constant $Q_1(T)$ such that

$$\mathbf{E} \left[\int_{S_1}^{S_2} \left(\sum_{j \neq i} \mu_{j \rightarrow i} \rho_{1,j}(s) + \nu_{m,i}(s) \right) ds \right] \leq \delta Q_1(T). \quad (3.27)$$

Now, let us bound the rest of the right-hand side term in (3.26).

We have, using Fubini's theorem for the first equality and the fact that the term under the integral is nonnegative for the upper bound,

$$\begin{aligned} & \mathbf{E} \left[\int_{S_1}^{S_2} \left(\int_0^t \sum_{n \neq m} \sum_{j \neq i} \mu_{j \rightarrow i} \mathbb{I}_{\{V_{(n,j)} \rightarrow i(s)=m\}} \lambda_{n,j}(s) ds + (r_i - \lambda_{m,i}(s)) \lambda_{m,i}(s) ds \right) dt \right] \\ &= \mathbf{E} \left[\int_0^{S_2} (S_2 - (S_1 \vee s)) \left(\sum_{n \neq m} \sum_{j \neq i} \mu_{j \rightarrow i} \mathbb{I}_{\{V_{(n,j)} \rightarrow i(s)=m\}} \lambda_{n,j}(s) ds \right. \right. \\ & \quad \left. \left. + (r_i - \lambda_{m,i}(s)) \lambda_{m,i}(s) \right) ds \right] \\ &\leq \delta \int_0^T \mathbf{E} \left[\sum_{j \neq i} \mu_{j \rightarrow i} \lambda_{1,j}(s) + \lambda_{m,i}^2(s) - r_i \lambda_{m,i}(s) \right] ds. \end{aligned}$$

Therefore, using Lemma 3.2.8, we have the existence of a constant $Q_2(T) > 0$ such that for all $(n, j) \in \{1, \dots, M\} \times \{1, \dots, K\}$,

$$\mathbf{E} \left[\int_{S_1}^{S_2} c_{n,j}(s) ds \right] \leq \delta Q_2(T).$$

Similarly, we obtain the existence of a constant $Q_3(T) > 0$ such that for all $(n, j) \in \{1, \dots, M\} \times \{1, \dots, K\}$,

$$\mathbf{E} \left[\int_{S_1}^{S_2} d_{n,j}(s) ds \right] \leq \delta Q_3(T).$$

Combining the two previous bounds, we have the existence of a nonnegative constant $Q_4(T)$ such that

$$\mathbf{E} \left[\int_{S_1}^{S_2} \left(\sum_{j \neq i} \mu_{j \rightarrow i} c_{1,j}(s) + d_{m,i}(s) \right) ds \right] \leq \delta Q_4(T). \quad (3.28)$$

Finally, combining (3.27) and (3.28) in (3.26), we obtain

$$\mathbf{P}(|\lambda_{m,i}(S_2) - \lambda_{m,i}(S_1)| > \varepsilon') \leq \frac{\delta(Q_1(T) + Q_4(T))}{\varepsilon'}.$$

Therefore, we can choose δ so that

$$\limsup_{N \rightarrow +\infty} \sup_{\substack{S_1, S_2 \in \mathcal{F}_t^{(n)} \text{ s.t.} \\ S_1 \leq S_2 \leq (S_1 + \delta) \wedge T}} \mathbf{P}(|\lambda_{m,i}^N(S_1) - \lambda_{m,i}^N(S_2)| > \varepsilon') \leq \varepsilon.$$

This proves the second condition of the Aldous criterion.

For the first condition, for $t \in [0, T]$, let

$$\begin{aligned} G(t) &= \sum_{n \neq m} \sum_{j \neq i} \mu_{j \rightarrow i} \int_0^t \mathbb{I}_{\{V_{(n,j) \rightarrow i}(s) = m\}} (N_{n,j}(ds) - \lambda_{n,j}(s) ds) \\ &\quad + \int_0^t (\lambda_{m,i}(s) - r_i) (N_{m,i}(ds) - \lambda_{m,i}(s) ds) \end{aligned}$$

and

$$H(t) = \sum_{n \neq m} \sum_{j \neq i} \mu_{j \rightarrow i} \int_0^t \mathbb{I}_{\{V_{(n,j) \rightarrow i}(s) = m\}} \lambda_{n,j}(s) ds + \int_0^t (\lambda_{m,i}(s) - r_i) \lambda_{m,i}(s) ds.$$

As G is a martingale, by Doob's inequality and by Lemma 3.2.8, there exists a constant $K_1(T)$ such that

$$\mathbf{E} \left[\sup_{t \in [0, T]} |G(t)| \right] \leq 4 \mathbf{E} [|G(T)|^2] \leq K_1(T).$$

Moreover, since all terms under the integrals in $H(t)$ are non-negative, $H(t)$ is non-decreasing in t , so by Lemma 3.2.8, there exists a constant $K_2(T)$ such that

$$\mathbf{E} \left[\sup_{t \in [0, T]} |H(t)| \right] \leq \mathbf{E}[H(T)] \leq K_2(T).$$

By the triangular inequality,

$$\begin{aligned} \mathbf{E} \left[\sup_{[0, T]} |\lambda_{m,i}(t)| \right] &\leq \mathbf{E} \left[\sup_{[0, T]} |G(t)| \right] + \mathbf{E} \left[\sup_{[0, T]} |H(t)| \right] \\ &\leq K_1(T) + K_2(T). \end{aligned}$$

Thus, for all $M \geq 2$, if $\kappa > 0$, by the Markov inequality,

$$\begin{aligned} \mathbf{P}(\|\lambda_{m,i}^M\|_\infty > \kappa) &\leq \frac{1}{\kappa} \mathbf{E}[\|\lambda_{m,i}^M\|_\infty] \\ &\leq \frac{K_1(T) + K_2(T)}{\kappa}, \end{aligned}$$

so for all $\varepsilon > 0$, there exists $\kappa > 0$ such that for all $M \geq 2$,

$$\mathbf{P}(\|\lambda_{m,i}^M\|_\infty > \kappa) < \varepsilon,$$

which proves the first condition of the Aldous criterion.

Thus, both conditions of the Aldous criterion are verified, and the set of processes $(\lambda_{m,i}^M)_{(m,i) \in \{1, \dots, M\} \times \{1, \dots, K\}}$ is tight in the space of càdlàg functions on \mathbb{R}^+ . Combining it with statement 4 of Theorem 3.2.4 yields the result. \square

3.3 Replica-mean-field limit of continuous-time fragmentation-interaction-aggregation processes

3.3.1 Introduction

Motivation

Many phenomena of interest, in the natural sciences or elsewhere, can be modeled as punctuate interactions between agents on an underlying network. Whether it be neural computations [35] [8], opinion dynamics [4], epidemics propagation [65] or wireless communications [71], a natural way to model the evolution in time of a population of agents or nodes is to consider the times at which interactions happen at each node as the realization of a point process on the real line. The resulting models describe the phenomena through a system of differential equations verified by the stochastic intensities of the point processes. From the point of view of a single node, its state evolves in the following fashion: it aggregates arrivals from its neighbors (these arrivals can be signed, for example, to model excitatory or inhibitory inputs in neural models), and in the event of a departure, its state is updated accordingly (it can for example decrease by one if we are interested in queueing models, be reset to a resting state to mimic spiking in neural networks, be divided by two for cellular division-type phenomena). In addition, we allow for the presence of a continuous drift, for example to model a refractory period after spiking in neuroscience models.

This flexibility allows for an accurate description of the phenomenon, but this accuracy comes at a price, namely, tractability, as neither the equations themselves, nor their associated functionals such as moment generating functions, admit closed forms except for some very particular cases.

As such, a common approach is to simplify the model by neglecting certain characteristics of the phenomenon, such as considering agents to be independent, and/or considering a particular scaling of the system, typically removing finite-size effects. One classical simplification is the mean-field regime, obtained by letting the number of agents go to infinity and scaling the interactions accordingly, usually inversely proportionally to the number of agents, thus averaging interactions across the system [34] [56]. The resulting equation, common to all agents and usually of McKean-Vlasov type, often allows for closed forms to be obtained, at the cost of losing correlations between particles and the geometry of the underlying network.

In recent years, different approaches have been developed to circumvent these limitations. To incorporate heterogeneity, the properties of graphons (large dense graphs) have been used to derive new limit equations [1] [43]. In this setting, the limit object is an infinite system of ODEs. Another approach circumvents mean-field models altogether, relying instead on conditional independence properties and local weak limits to obtain local convergence [55].

Another approach to obtain closed forms is called the Poisson Hypothesis. First formulated by Kleinrock for large queueing systems [49], it states that the flow of arrivals to a given node can be approximated by a Poisson flow with rate equal to the average rate of the original flow of arrivals. In agent-based models, the flow of arrivals corresponds to the effect of interactions on a given node. Under the Poisson Hypothesis, the behavior of each agent is still described by a stochastic differential equation, but the agents are considered independent and interaction times are replaced by Poisson process clocks, which in certain models allows for tractability. This regime has been studied for queueing models by Rybko, Shlosman and others [78] and by Baccelli and Taillefumier for intensity-based models from computational neuroscience [8].

A point of interest is the construction of physical models that, when properly scaled, converge to the Poisson Hypothesis regime, analogously to the classical mean-field construction. The replica-mean-field construction has been shown to be a successful answer to this question in various settings. This approach consists in building a new physical system comprising randomly interacting copies of the original network, and then letting the number of copies, or replicas, to go to infinity. From the point of view of a single node inside one of the replicas, its state evolves in a similar fashion to that of the original model, except that arrivals are now aggregated across neighbors in all the replicas. When there is a departure from the node, for each of its neighbors, an independent routing gives the index of the

replica in which the neighbor will aggregate the arrival to its state.

As the probability of two nodes interacting scales inversely proportionally to the number of replicas, the replicas become asymptotically independent when their number goes to infinity. Additionally, the aggregated arrivals to a given node can be seen as a (random) sum of rare events, which heuristically gives rise to a Poisson process at the limit in the number of replicas. These observations give an informal idea of how it has been shown in the particular case of intensity-based neural dynamics that when the number of replicas goes to infinity, the dynamics of a typical replica converge to those of under the Poisson Hypothesis. This has been proved in both continuous and discrete time for excitatory neural dynamics [5][27]. In the discrete time framework, a class of discrete-time processes, called fragmentation-aggregation-interaction processes, or FIAPs for short, for which the same limit theorem holds, has been introduced [5]. A natural question, and the aim of this work, is to introduce an analogous class of processes in continuous time and to extend the convergence result for replica-mean-field versions of such processes obtained in the specific case of excitatory neural dynamics.

The rest of the work is organized as follows: hereafter, we introduce the class of continuous-time fragmentation-interaction-aggregation processes and state the main result. Section 2 pertains to the proof of the result. Section 3 establishes a link between continuous-time and discrete-time FIAPs in the particular case of excitatory neural dynamics. Finally, Section 4 discusses some natural extensions of the result.

Continuous-time fragmentation-interaction-aggregation processes

First, recall the definition of the stochastic intensity of a point process: we introduce the network history $(\mathcal{F}_t)_{t \in \mathbb{R}}$ as an increasing collection of σ -fields such that

$$\mathcal{F}_t^{\mathbf{N}} = \{\sigma(N_1(B_1), \dots, N_K(B_K)) | B_i \in \mathbf{B}(\mathbb{R}), B_i \subset (-\infty, t]\} \subset \mathcal{F}_t,$$

where $\mathcal{F}_t^{\mathbf{N}}$ is the internal history of the process \mathbf{N} .

Then, the \mathcal{F}_t -stochastic intensity $\{\lambda_i(t)\}_{t \in \mathbb{R}}$ of the associated point process N_i is the \mathcal{F}_t -predictable random process satisfying for all $s < t \in \mathbb{R}$:

$$\mathbf{E}[N_i(s, t) | \mathcal{F}_s] = \mathbf{E}\left[\int_s^t \lambda_i(u) du | \mathcal{F}_s\right], \quad (3.29)$$

where \mathcal{F}_t is the network history. We will hereafter refer to (3.29) as the stochastic intensity property. See [17] for more details on point processes admitting stochastic intensities.

This in turn allows us to define fragmentation-aggregation-interaction processes, or FIAPs for short, as processes with specific stochastic intensities:

Definition 3.3.1. Let $K \geq 2$. We define a continuous-time fragmentation-interaction-aggregation process, hereafter referred to as a cFIAP, as a collection of point processes $(N_i)_{1 \leq i \leq K}$ admitting stochastic intensities (λ_i) with regards to the network history such that for any $t \in \mathbb{R}^+$, there exist

- point processes $(\hat{N}_{j \rightarrow i})_{1 \leq i \leq K, j \neq i}$ admitting stochastic intensities $(\hat{\lambda}_{j \rightarrow i})$ hereafter referred to as interaction processes;
- functions $(h_{j \rightarrow i})_{1 \leq i \leq K, j \neq i} : \mathbb{R} \mapsto \mathbb{R}$ hereafter referred to as interaction functions such that there exists $H > 0$ verifying for all i, j and all $t \in \mathbb{R}$, $|h_{j \rightarrow i}(t)| \leq H$;
- functions $(g_i)_{1 \leq i \leq K} : \mathbb{R} \mapsto \mathbb{R}^+$ and $(\sigma_i)_{1 \leq i \leq K} : \mathbb{R} \mapsto \mathbb{R}^+$ hereafter referred to as autonomous evolution functions;
- a Lipschitz function $f : \mathbb{R} \mapsto \mathbb{R}^+$ such that $f(0) = 0$;

such that for all $1 \leq i \leq K$ and all $t \in \mathbb{R}^+$,

$$\begin{aligned} \lambda_i(t) &= \lambda_i(0) + f \left(\sum_{j \neq i} \int_0^t h_{j \rightarrow i}(s) \hat{N}_{j \rightarrow i}(ds) \right) + \int_0^t (g_i(s, \lambda_i(s)) - \lambda_i(s)) N_i(ds) \\ &\quad + \int_0^t (\sigma_i(s, \lambda_i(s)) - \lambda_i(s)) ds. \end{aligned} \tag{3.30}$$

We now formalize the Poisson Hypothesis.

Definition 3.3.2. We say that a cFIAP satisfies the Poisson Hypothesis if all interaction times are given by independent Poisson processes. We denote with tildes all state processes in this regime. Namely, for all $i, j \in 1 \leq i \leq K$ with $i \neq j$, $\tilde{N}_{j \rightarrow i}$ are independent Poisson processes with intensities $s \rightarrow \mathbf{E}[\tilde{\lambda}_j(s)]$ and for all $t \in \mathbb{R}^+$,

$$\begin{aligned} \tilde{\lambda}_i(t) &= \tilde{\lambda}_i(0) + f \left(\sum_{j \neq i} \int_0^t h_{j \rightarrow i}(s) \tilde{N}_{j \rightarrow i}(ds) \right) + \int_0^t (g_i(s, \tilde{\lambda}_i(s)) - \tilde{\lambda}_i(s)) \tilde{N}_i(ds) \\ &\quad + \int_0^t (\sigma_i(s, \tilde{\lambda}_i(s)) - \tilde{\lambda}_i(s)) ds. \end{aligned} \tag{3.31}$$

Given a cFIAP, we now aim to define its replica-mean-field version.

Definition 3.3.3. Let $K, M \geq 2$. The M -replica-mean-field cFIAP is given by the collection of point processes $(N_{m,i}^M)_{1 \leq i \leq K, 1 \leq m \leq M}$ admitting stochastic intensities $(\lambda_{m,i}^M)$ such that for any $t \in \mathbb{R}^+$, there exist

- point processes $(\widehat{N}_{n,j \rightarrow i})_{1 \leq i \leq K, j \neq i, 1 \leq n \leq M}$ admitting stochastic intensities $(\widehat{\lambda}_{n,j,i})$ hereafter referred to as aggregation processes;
- functions $(h_{j \rightarrow i})_{1 \leq i \leq K, j \neq i} : \mathbb{R} \mapsto \mathbb{R}$ hereafter referred to as interaction functions such that there exists $H > 0$ verifying for all i, j and all $t \in \mathbb{R}$, $|h_{j \rightarrow i}(t)| \leq H$;
- functions $(g_i)_{1 \leq i \leq K} : \mathbb{R} \mapsto \mathbb{R}^+$ and $(\sigma_i)_{1 \leq i \leq K} : \mathbb{R} \mapsto \mathbb{R}^+$ hereafter referred to as autonomous evolution functions;
- a Lipschitz function $f : \mathbb{R} \mapsto \mathbb{R}^+$ such that $f(0) = 0$;
- (\mathcal{F}_t) -predictable routing processes $\{V_{(m,i) \rightarrow j}^M(t)\}_{t \in \mathbb{R}}$ for $1 \leq m \leq M, 1 \leq i, j \leq K$, such that, for each interaction time T , i.e., each point of $N_{m,i}^M$, the random variables $\{V_{(m,i) \rightarrow j}^M(T)\}_j$ are mutually independent, independent from the past, and uniformly distributed on $\{1, \dots, M\} \setminus \{m\}$

such that for all $1 \leq m \leq M, 1 \leq i \leq K$ and all $t \in \mathbb{R}^+$,

$$\begin{aligned} \lambda_{m,i}(t) &= \lambda_{m,i}(0) + f\left(\sum_{j \neq i} \sum_{n \neq m} \int_0^t h_{j \rightarrow i}(s) \mathbb{I}_{\{V_{(n,j) \rightarrow i}^M(s)=m\}} \widehat{N}_{n,j \rightarrow i}(ds)\right) \\ &\quad + \int_0^t (g_i(s, \lambda_{m,i}(s)) - \lambda_{m,i}(s)) N_{m,i}(ds) + \int_0^t (\sigma_i(s, \lambda_{m,i}(s)) - \lambda_{m,i}(s)) ds. \end{aligned} \tag{3.32}$$

Hereafter, we will always assume that $\widehat{N}_{n,j \rightarrow i} = N_{n,j}$ for all $1 \leq n \leq M, 1 \leq j \leq K$.

We will moreover always consider the following assumptions on the functions g_i and σ_i :

Assumption 3.3.4. For all $s, t \in \mathbb{R}$, for all $i \in \{1, \dots, K\}$,

$$g_i(s, t) \leq t \text{ and } \sigma_i(s, t) \leq t.$$

In particular,

$$g_i(s, \lambda_i(s)) \leq \lambda_i(s) \text{ and } \sigma_i(s, \lambda_i(s)) \leq \lambda_i(s).$$

Note that this implies that the state processes λ_i are always decreasing in between aggregations.

In a similar fashion to [27], we also require the following assumption on the initial conditions:

Assumption 3.3.5. There exists $\xi_0 > 0$ such that for all $1 \leq m \leq M, 1 \leq i \leq K$ and all $0 < \xi \leq \xi_0$, $\mathbf{E}[e^{\xi \lambda_{m,i}(0)}] < \infty$.

Examples of continuous-time FIAPs

We now give a few instances of specific FIAPs.

Example 3.3.6. Taking for $1 \leq i, j \leq K$ and all $t \in \mathbb{R}$,
 $h_{j \rightarrow i}(t) = \mu_{j \rightarrow i} \geq 0$, $f(t) = |t|$, $g_i(t, \lambda_i(t)) = r_i > 0$, $\sigma_i(t, \lambda_i(t)) = b_i > 0$, we retrieve the excitatory Galves-Löcherbach model [35] [8].

Example 3.3.7. Taking for $1 \leq i, j \leq K$ and all $t \geq 0$,
 $h_{j \rightarrow i}(t) = \mu_{j \rightarrow i} \in \mathbb{R}$, $f(t) = \max(0, |t|)$, $g_i(t, \lambda_i(t)) = r_i > 0$, $\sigma_i(t, \lambda_i(t)) = b_i > 0$, we obtain a more general Galves-Löcherbach model incorporating inhibition.

Example 3.3.8. Taking for $1 \leq i, j \leq K$ and all $t \geq 0$,
 $h_{j \rightarrow i}(t) = \mathbb{I}_{\{j=i+1 \pmod K\}}$, $f(t) = |t|$, $g_i(t, \lambda_i(t)) = \lambda_i(t) - 1$, $\sigma_i(t) = \lambda_i(t)$, we obtain a continuous-time concatenation queueing network. Note that such a network under the Poisson Hypothesis is an instance of a Gordon-Newell queueing network [50].

The main result

Recall the following definition of convergence in total variation:

Definition 3.3.9. Let P and Q be two probability measures on a probability space (Ω, \mathcal{F}) . We define the total variation distance by

$$d_{TV}(P, Q) = \sup_{A \in \mathcal{F}} |P(A) - Q(A)|.$$

When Ω is countable, an equivalent definition is

$$d_{TV}(P, Q) = \frac{1}{2} \sum_{\omega \in \Omega} |P(\omega) - Q(\omega)|.$$

We will abusively say that a sequence of random variables converges in total variation when the sequence of their distributions does.

The following theorem is the main result of this work:

Theorem 3.3.10. There exists $T \in \mathbb{R}^+$ such that for $t \in [0, T]$, if

$$A_{m,i}^M(t) = \sum_{j \neq i} \sum_{n \neq m} \int_0^t h_{j \rightarrow i}(s) \mathbb{I}_{\{V_{(n,j) \rightarrow i}^M(s) = m\}} N_{n,j}^M(ds),$$

with N^M defined in (3.32), and

$$\tilde{A}_i(t) = \sum_{j \neq i} \int_0^t h_{j \rightarrow i}(s) \hat{N}_{j \rightarrow i}(ds),$$

with $(\hat{N}_{j \rightarrow i})_j$ independent Poisson point processes with respective intensities $s \mapsto \mathbf{E}[\tilde{\lambda}_j(s)]$, then,

1. the processes $(\tilde{A}_1, \dots, \tilde{A}_K)$ are independent, as are the processes $(\tilde{\lambda}_1, \dots, \tilde{\lambda}_K)$;
2. for all $(m, i) \in \{1, \dots, M\} \times \{1, \dots, K\}$, the random variable $A_{m,i}^M(t)$ converges in total variation to $\tilde{A}_i(t)$ when $M \rightarrow \infty$;
3. for all $(m, i) \in \{1, \dots, M\} \times \{1, \dots, K\}$, the random variable $\lambda_{m,i}^M(t)$ defined by (3.32) converges in total variation to $\tilde{\lambda}_i(t)$ defined in (3.31) when $M \rightarrow \infty$;
4. let \mathcal{N} be a finite subset of \mathbb{N}^* , for all $i \in \{1, \dots, K\}$, the processes $(A_{m,i}^M(\cdot))_{m \in \mathcal{N}}$ and $(\lambda_{m,i}^M(\cdot))_{m \in \mathcal{N}}$ weakly converge in the Skorokhod space $D([0, T])^{\text{card}(\mathcal{N})}$ endowed with the product Skorokhod topology to $\text{card}(\mathcal{N})$ independent copies of the corresponding limit processes $(\tilde{A}_i(\cdot))$ and $(\tilde{\lambda}_i(\cdot))$ when $M \rightarrow \infty$.

3.3.2 Proof of the theorem

We will follow the general proof framework developed by the author in the previous work [27]. We will emphasize the technical points that were adapted to this more general case.

First, we remind the following Poisson embedding representation for point processes with a stochastic intensity [18], allowing us to construct all the state processes coupled through their Poisson embeddings and initial conditions.

Lemma 3.3.11. *Let N be a point process on \mathbb{R} . Let (\mathcal{F}_t) be an internal history of N . Suppose N admits a (\mathcal{F}_t) -stochastic intensity $\{\mu(t)\}_{t \in \mathbb{R}}$. Then there exists a Poisson point process \bar{N} with intensity 1 on \mathbb{R}^2 such that, for all $C \in \mathcal{B}(\mathbb{R})$,*

$$N(C) = \int_{C \times \mathbb{R}} \mathbb{I}_{[0, \mu(s)]}(u) \bar{N}(ds \times du).$$

For $m \geq 1, M \geq 1, 1 \leq i \leq K$, let $\bar{N}_{m,i}$ be i.i.d. Poisson point processes on $\mathbb{R}^+ \times \mathbb{R}^+$ with intensity 1.

Let $\Omega = (\mathbb{R}^+ \times ((\mathbb{R}^+)^2)^{\mathbb{N}^*})^{\mathbb{N}^*}$ be a probability space endowed with the probability measure $(\mu_0 \otimes P)^{\otimes \mathbb{N}^*}$, where μ_0 is the law of the initial conditions and P is the law of a Poisson process with intensity 1 on $(\mathbb{R}^+)^2$. We construct on Ω the following processes:

- The processes $(N_{m,i}^M(t)), m \geq 1, M \geq 1, 1 \leq i \leq K$, with stochastic intensities

$(\lambda_{m,i}^M(t))$ verifying

$$\begin{aligned} \lambda_{m,i}^M(t) &= \int_0^t \int_0^{+\infty} (g_i(s, \lambda_{m,i}(s)) - \lambda_{m,i}^M(s)) \mathbb{I}_{[0, \lambda_{m,i}^M(s)]}(u) \bar{N}_{m,i}(ds \times du) \\ &+ f \left(\sum_{n \neq m} \sum_{j \neq i} \int_0^t \int_0^{+\infty} h_{j \rightarrow i}(s) \mathbb{I}_{\{V_{(n,j) \rightarrow i}^M(s) = m\}} \mathbb{I}_{[0, \lambda_{n,j \rightarrow i}^M(s)]}(u) \bar{N}_{n,j \rightarrow i}(ds \times du) \right) \\ &+ \int_0^t (\sigma_i(s, \lambda_{m,i}(s)) - \lambda_{m,i}(s)) ds + \lambda_{m,i}^M(0), \end{aligned} \quad (3.33)$$

with $\lambda_{m,i}^M(0) = Z_i$ for all $m \in \mathbb{N}^*$ and where, for all M , $(V_{(n,j) \rightarrow i}^M(t))_j$ are càdlàg stochastic processes such that for each point T of $\hat{N}_{n,j}$, the random variables $(V_{(n,j) \rightarrow i}^M(T))_j$ are independent of the past, mutually independent and uniformly distributed on $\{1, \dots, M\} \setminus \{n\}$, considered as marks of the Poisson point process $\bar{N}_{n,j}$. Namely, to each point of the Poisson embedding, we attach a mark that is an element of $(\mathbb{N}^K)^{\mathbb{N}^*}$, where the M th term of the sequence corresponds to $(V_{(n,j) \rightarrow i}^M(T))_j$.

- The processes $(\tilde{N}_i(t))$, $1 \leq i \leq K$, with stochastic intensities $(\tilde{\lambda}_i(t))$ verifying

$$\begin{aligned} \tilde{\lambda}_i(t) &= \tilde{\lambda}_i(0) + f \left(\sum_{j \neq i} \int_0^t \int_0^{+\infty} h_{j \rightarrow i}(s) \mathbb{I}_{[0, \mathbf{E}[\tilde{\lambda}_j(s)]]}(u) \bar{N}_{j,i}(ds \times du) \right) \\ &+ \int_0^t \int_0^{+\infty} (g_i(s, \tilde{\lambda}_i(s)) - \tilde{\lambda}_i(s)) \mathbb{I}_{[0, \tilde{\lambda}_i(s)]}(u) \bar{N}_{i,i}(ds \times du), \end{aligned} \quad (3.34)$$

with $\tilde{\lambda}_i(0) = Z_i$.

Just as in the particular case of neural dynamics, this representation is sufficient to derive the following, which is statement 1 of Theorem 3.3.10.

Lemma 3.3.12. *The processes $(\tilde{A}_i)_{1 \leq i \leq K}$ are independent, as are the processes $(\tilde{\lambda}_1, \dots, \tilde{\lambda}_K)$.*

Proof. For all $t \in [0, T]$, we can write using the construction above

$$\tilde{A}_i(t) = \sum_{j \neq i} \int_0^t \int_0^{+\infty} h_{j \rightarrow i}(s) \mathbb{I}_{[0, \mathbf{E}[\tilde{\lambda}_j(s)]]}(u) \bar{N}_{j,i}(ds \times du).$$

Therefore, all the randomness in \tilde{A}_i is contained in the Poisson embeddings $(\hat{N}_{k,i})_{1 \leq k \leq K}$. Thus, for $i \neq j$, \tilde{A}_i and \tilde{A}_j are independent. The independence of the processes $(\tilde{\lambda}_1, \dots, \tilde{\lambda}_K)$ follows in the same manner by a mapping theorem argument. \square

Properties of the RMF and limit processes

In this section, we prove several properties of the RMF and limit dynamics that will be used throughout the proof.

In what follows, we will often omit the M superscript in the notations $N_{m,i}^M$, $A_{m,i}^M$ and $\lambda_{m,i}^M$ to increase readability.

We generalize from [27] the following representation of the arrival process $A_{m,i}(t)$. For $n \neq m$ and $j \neq i$, if $S \in \text{Supp}(N_{n,j}|_{[0,T]})$, we define $B_{S,(n,j) \rightarrow (m,i)}^M$ the random variable equal to 1 if the routing between replicas at time S caused by a departure in node j in replica n chose the replica m for the recipient i of the interaction thus produced, and 0 otherwise. As such, it is clear that we can write for all $t \in [0, T]$, $m \in \{1, \dots, M\}$ and $i \in \{1, \dots, K\}$,

$$A_{m,i}(t) = \sum_{n \neq m} \sum_{j \neq i} \sum_{k \in N_{n,j} \cap [0,t]} h_{j \rightarrow i}(k) B_{k,(n,j) \rightarrow (m,i)}^M. \quad (3.35)$$

Note that when m, n, i and j are fixed, the random variables $(B_{k,(n,j) \rightarrow (m,i)}^M)_{k \in N_{n,j}([0,T])}$ are i.i.d. Also note that when n, j, i and k are fixed, the joint distribution of $(B_{k,(n,j) \rightarrow (m,i)}^M)_m$ with $m \in \{1, \dots, M\}$ is that of Bernoulli random variables with parameter $\frac{1}{M-1}$ such that exactly one of them is equal to 1, all the others being equal to 0. Combining these two observations allows us to show that the following lemma, highlighting a key property of the replica-mean-field approach, holds:

Lemma 3.3.13. *Fix $(m, i) \in \{1, \dots, M\} \times \{1, \dots, K\}$. Keeping notation from (3.35), let*

$$N = (N_{n,j}([0, t]))_{n \neq m, j \neq i} \in \mathbb{N}^{(K-1)(M-1)}.$$

Conditionally on the event $\{N = q\}$, for $q = (q_{n,j})_{n \neq m, j \neq i} \in \mathbb{N}^{(K-1)(M-1)}$, the random variables $(B_{k,(n,j) \rightarrow (m,i)}^M)_{n \neq m, j \neq i, k \in \{1, \dots, q_{n,j}\}}$ are independent Bernoulli random variables with parameter $\frac{1}{M-1}$.

Proof. The structure of the proof is unchanged from [27]: since N is entirely determined by the Poisson embeddings $(\hat{N}_{n,j})_{j \neq i}$ and the arrivals to the nodes (n, j) from all the nodes $h \neq j$ across replicas, it is sufficient to show that these arrivals and the routing variables $(B_{k,(n,j) \rightarrow (m,i)}^M)_{k \leq \hat{N}_{n,j}([0,t] \times \mathbb{R}^+)}$ are independent. Intuitively, this holds because arrivals are aggregated across all replicas, which will erase the eventual dependencies due to the routing variables to nodes i choosing one replica rather than another.

In order to transcribe this intuition rigorously, we first show that the total number of departures from nodes i up to time t , that is, $\sum_{l=1}^M N_{l,i}([0, t])$, and the routing variables

$(B_{k,(n,j) \rightarrow (m,i)}^M)_{k \leq \hat{N}_{n,j}([0,t] \times \mathbb{R}^+)}$ are independent. Indeed, using the representation given by Lemma 3.3.11, due to the structure of the Poisson embeddings $(\hat{N}_{l,i})_{l \in \{1, \dots, M\}}$,

there is a point of $\sum_{l=1}^M N_{l,i}$ in some interval I iff there is a point of the superposition of the corresponding Poisson embeddings such that the x-coordinate is in I and the y-coordinate is under the curve of the function $t \rightarrow \sum_{l=1}^M \lambda_{l,i}(t)$. In turn, the last event does not depend on $(B_{k,(n,j) \rightarrow (m,i)}^M)_{k \leq \hat{N}_{n,j}([0,t] \times \mathbb{R}^+)}$, as the symmetry inherent to the replica structure ensures that all arrivals increment $t \rightarrow \sum_{l=1}^M \lambda_{l,i}(t)$ by the same amount, which concludes the proof of this preliminary remark.

For all (n, j) such that $n \neq m$ and $j \neq i$, let

$$A_{i \rightarrow (n,j)}(t) = \sum_{l \neq n} \sum_{k \in N_{l,i} \cap [0,t]} h_{j,i}(k) B_{k,(l,i) \rightarrow (n,j)}^M.$$

Note that $A_{i \rightarrow (n,j)}(t)$ represents the arrivals to node j in replica n from all nodes i across replicas. As such, it is clear that we can write

$$A_{i \rightarrow (n,j)}(t) = \sum_{k \in \sum_{l \neq n} N_{l,i} \cap [0,t]} h_{j,i}(k) B_{k,(i) \rightarrow (n,j)}^M,$$

where $(B_{k,(i) \rightarrow (n,j)}^M)$ are independent Bernoulli random variables with parameter $\frac{1}{M-1}$ such that they and $(B_{k,(n,j) \rightarrow (m,i)}^M)$ are independent. Then by the previous observation, $A_{i \rightarrow (n,j)}(t)$ and $(B_{k,(n,j) \rightarrow (m,i)}^M)$ are independent. Therefore, N , which is entirely determined by the Poisson embeddings $(\hat{N}_{n,j})$ and the arrivals $(A_{h \rightarrow (n,j)}(t))_{h \neq j}$, and $(B_{k,(n,j) \rightarrow (m,i)}^M)_{k \leq \hat{N}_{n,j}([0,T] \times \mathbb{R}^+)}$, are independent. Thus, conditioning on N does not break independence between the variables $(B_{k,(n,j) \rightarrow (m,i)}^M)$. \square

We will now give bounds on the moments of both the M-replica and limit processes, using the bounds on the moments of the initial conditions. The validity of this bound is the main reason for the introduction of Assumption 3.3.4, which allows to stochastically dominate the dynamics by the same dynamics without the autonomous evolution integral terms, which enables Grönwall's lemma.

Lemma 3.3.14. *Suppose the initial conditions verify Assumption 3.3.5. Then, for all $p \geq 1$, for all $(m, i) \in \{1, \dots, M\} \times \{1, \dots, K\}$, for all $t \in [0, T]$, there exists $Q_p \in \mathbb{R}_p[X]$ a polynomial of degree exactly p such that*

$$\mathbf{E}[\lambda_{m,i}^p(t)] \leq Q_p(\mathbf{E}[\lambda_{m,i}(0)]). \tag{3.36}$$

Proof. Note that, by Assumption 3.3.4 and monotonicity, the dynamics that we consider are stochastically dominated by the same dynamics without the autonomous evolution terms, by which we mean the two last integral terms in (3.32). Thus, we can restrict ourselves to this special case. We first prove the result for

$p = 1$. Let $t \in [0, T]$. Let $i(t) = \arg \max_{j \in \{1, \dots, K\}} \left| \mathbf{E}[\widehat{\lambda}_{n, j \rightarrow i}(t)] \right|$. Using the property of stochastic intensity, we have

$$\mathbf{E}[\lambda_{m, i(t)}(t)] = \mathbf{E}[\lambda_{m, i(t)}(0)] + \mathbf{E} \left[f \left(\sum_{n \neq m} \sum_{j \neq i(t)} \int_0^t \mathbb{1}_{\{V_{(n, j) \rightarrow i(t)}(s) = m\}} \widehat{\lambda}_{n, j \rightarrow i(t)}(s) \, ds \right) \right].$$

Using the assumptions on f and $h_{j \rightarrow i}$, we have

$$\mathbf{E}[\lambda_{m, i(t)}(t)] \leq \mathbf{E}[\lambda_{m, i(t)}(0)] + \frac{H}{M-1} \sum_{n \neq m} \sum_{j \neq i(t)} \mathbf{E} \left[\int_0^t \lambda_{n, j}(s) \, ds \right].$$

By the definition of $i(t)$ and exchangeability of the replicas, we have

$$\mathbf{E}[\lambda_{m, i(t)}(t)] \leq \mathbf{E}[\lambda_{m, i(t)}(0)] + H \sum_{j \neq i(t)} \int_0^t \mathbf{E}[\lambda_{m, i(s)}(s)] \, ds.$$

This gives by Grönwall's lemma the desired result:

$$\mathbf{E}[\lambda_{m, i(t)}(t)] \leq \mathbf{E}[\lambda_{m, i(t)}(0)] e^{(K-1)HT} =: Q_1(\mathbf{E}[\lambda_{m, i(t)}(0)]). \quad (3.37)$$

This reasoning can be extended by induction to all $p \geq 2$. \square

Finally, note that the exact same reasoning can be applied to obtain an equivalent result for the limit process, which we will only state:

Lemma 3.3.15. *For all $p \geq 1$, for all $i \in \{1, \dots, K\}$, for all $t \in [0, T]$, there exists $\tilde{Q}_p \in \mathbb{R}_p[X]$ a polynomial of degree exactly p such that*

$$\mathbf{E}[\tilde{\lambda}_i^p(t)] \leq \tilde{Q}_p[\mathbf{E}[\tilde{\lambda}_i(0)]]. \quad (3.38)$$

Lemma 3.3.14 allows us to prove the following result, which states that Assumption 3.3.5 can be propagated to any time t less than some fixed T .

Lemma 3.3.16. *There exists $\xi_0 > 0$ and $T > 0$ such that for $\xi \leq \xi_0$ and all $t \leq T$,*

$$\mathbf{E}[e^{\xi \lambda_{m, i(t)}}] < \infty \text{ and } \mathbf{E}[e^{\xi \tilde{\lambda}_{m, i(t)}}] < \infty. \quad (3.39)$$

Proof. To prove this result, we once again consider exchangeable dynamics without resets, using the same observation as previously, namely that nonexchangeable dynamics with resets are stochastically dominated by exchangeable dynamics without resets, to generalize the result. Note in addition that exchangeable dynamics

with interaction functions $h_{j \rightarrow i}$ are dominated by exchangeable dynamics with interaction functions $|h_{j \rightarrow i}|$. Let ξ_0 as in Assumption 3.3.5. Let $t \in [0, T]$. Let us write out the equation verified by $e^{\xi \lambda_{m,i}(t)}$:

$$e^{\xi \lambda_{m,i}(t)} = e^{\xi \lambda_{m,i}(0)} + f \left(\sum_{j \neq i} \sum_{n \neq m} \int_0^t \mathbb{I}_{\{V_{(n,j) \rightarrow i}^M(s) = m\}} e^{\xi \lambda_{m,i}(s)} (e^{\xi h_{j \rightarrow i}(s)} - 1) N_{n,j}(ds) \right).$$

Taking the expectation, using the stochastic intensity property and the conditions on f and $(h_{j \rightarrow i})$, we get

$$\mathbf{E}[e^{\xi \lambda_{m,i}(t)}] \leq \mathbf{E}[e^{\xi \lambda_{m,i}(0)}] + \frac{1}{M-1} \sum_{j \neq i} \sum_{n \neq m} \int_0^t \mathbf{E}[e^{\xi \lambda_{m,i}(s)} (e^{\xi H} - 1) \lambda_{n,j}(s)] ds.$$

Using exchangeability between replicas, this boils down to

$$\mathbf{E}[e^{\xi \lambda_{m,i}(t)}] = \mathbf{E}[e^{\xi \lambda_{m,i}(0)}] + \sum_{j \neq i} \int_0^t \mathbf{E}[e^{\xi \lambda_{m,i}(s)} (e^{\xi H} - 1) \lambda_{m,j}(s)] ds.$$

Since we are looking at dynamics without resets and with only nonnegative interactions, $\lambda_{m,i}(s)$ and $\lambda_{m,j}(s)$ are positively correlated. Therefore, we have

$$\mathbf{E}[e^{\xi \lambda_{m,i}(t)}] \leq \mathbf{E}[e^{\xi \lambda_{m,i}(0)}] + (e^{\xi \mu} - 1) \sum_{j \neq i} \int_0^t \mathbf{E}[e^{\xi \lambda_{m,i}(s)}] \mathbf{E}[\lambda_{m,j}(s)] ds.$$

By Lemma 3.3.14 and Assumption 3.3.5, we have the existence of a constant $B > 0$ such that

$$\mathbf{E}[e^{\xi \lambda_{m,i}(t)}] \leq \mathbf{E}[e^{\xi \lambda_{m,i}(0)}] + (e^{\xi \mu} - 1)(K-1)B \int_0^t \mathbf{E}[e^{\xi \lambda_{m,i}(s)}] ds.$$

The desired result follows from Grönwall's lemma. The equivalent result for $\tilde{\lambda}_{m,i}(t)$ is obtained in the same way. \square

Poisson approximation bound

The goal of this section is to extend the bound obtained using the Chen-Stein method [23] in [27] to obtain a bound in total variation distance between the arrivals term (3.35) and the limit sum of Poisson random variables. Recall that (3.35) states that for all $t \in [0, T]$, $m \in \{1, \dots, M\}$, $i \in \{1, \dots, K\}$,

$$A_{m,i}(t) = \sum_{n \neq m} \sum_{j \neq i} \sum_{k \in \tilde{N}_{n,j} \cap [0,t]} h_{j \rightarrow i}(k) B_{k,(n,j) \rightarrow (m,i)}^M.$$

We will write

$$A_{m,i}(t) = \sum_{j \neq i} A_{j \rightarrow (m,i)}(t),$$

where for all $j \neq i$,

$$A_{j \rightarrow (m,i)}(t) = \sum_{n \neq m} \sum_{k \in \hat{N}_{n,j} \cap [0,t]} h_{j \rightarrow i}(k) B_{k,(n,j) \rightarrow (m,i)}^M.$$

We will hereafter consider that the functions $h_{j \rightarrow i}$ are simple, that is, that they are finite linear combinations of indicator functions. This is only a temporary assumption to obtain a bound in total variation distance using the Chen-Stein method, and we will relax it later by using a density argument.

Since $h_{j \rightarrow i}$ is simple, there exist $p \geq 0$, $a_1, \dots, a_p \in \mathbb{R}$ and A_1, \dots, A_p measurable subsets of \mathbb{R} such that for all $t \in \mathbb{R}$,

$$h_{j \rightarrow i}(t) = \sum_{l=0}^p a_l \mathbb{I}_{A_l}(t)$$

. Therefore, we can write

$$A_{j \rightarrow (m,i)}(t) = \sum_{l=0}^p a_l \sum_{n \neq m} \sum_{k \in \hat{N}_{n,j} \cap ([0,t] \cap A_l)} B_{k,(n,j) \rightarrow (m,i)}^M. \quad (3.40)$$

By the independence property of the Poisson point process, we see that without loss of generality, we can assume that $h_{j \rightarrow i}$ is a constant.

This allows us to now simply reuse the Lemma proved by the author in [27], which we recall below, using notation consistent with (3.35), with the addition of a multiplicative constant C in the bound to account for the function $h_{j \rightarrow i}$. Since what follows is done with $t \in [0, T]$ fixed, we will additionally denote $N_{n,j}([0, t])$ by $N_{n,j}$, continuing to omit the M superscript to simplify notation.

Lemma 3.3.17. *Let $M > 1$. Let $(m, i) \in \{1, \dots, M\} \times \{1, \dots, K\}$. For $j \in \{1, \dots, K\} \setminus \{i\}$, let $A_{j \rightarrow (m,i)} = \sum_{n \neq m} \sum_{k \leq N_{n,j}} h_{j \rightarrow i}(k) B_{k,(n,j) \rightarrow (m,i)}$ with $h_{j \rightarrow i}$ simple, and let $\tilde{A}_{j \rightarrow i}$ be independent Poisson random variables with means $\mathbf{E}[N_{1,j}]$. Then, there exists $C > 0$ such that*

$$\begin{aligned} d_{TV}(A_{j \rightarrow (m,i)}, \tilde{A}_{j \rightarrow i}) &\leq C \left(\left(1 \wedge \frac{0.74}{\sqrt{\mathbf{E}[N_{1,j}]}} \right) \frac{1}{M-1} \mathbf{E} \left[\left| \sum_{n \neq m} (\mathbf{E}[N_{n,j}] - N_{n,j}) \right| \right] \right. \\ &\quad \left. + \frac{1}{M-1} \left(1 \wedge \frac{1}{\mathbf{E}[N_{1,j}]} \right) \mathbf{E}[N_{1,j}] \right). \end{aligned} \quad (3.41)$$

We refer to [27] for the proof and comments on this result, which relies on the Chen-Stein method for Poisson approximation, using the conditional independence property proved in Lemma 3.3.13 to condition on the random amount of aggregations.

Note that if the bound on the right-hand side goes to 0 when $M \rightarrow \infty$, as we will endeavor to prove in the next section, then we will obtain convergence in total variation for all measurable functions $h_{j \rightarrow i}$, as any such function can be represented as a uniform limit of simple functions, and the uniform limit commutes with total variation convergence.

Decoupling arrivals and outputs: a fixed point scheme approach

As we have seen in the previous lemma, for the Poisson approximation to hold, it is sufficient to prove a law of large numbers-type result on the random variables $(N_{n,j})_{n \neq m}$. However, since these random variables themselves depend on the random variables $(A_{j \rightarrow (m,i)})_{m \in \{1, \dots, M\}}$, a direct proof seems difficult to obtain.

As such, we follow the approach of [27] and consider Equation (3.32) as the fixed point equation of some function on the space of probability laws on the space of càdlàg trajectories. This fixed point exists and is necessarily unique due to the fact that Equation (3.32) admits a unique solution. The main idea goes as follows: if we endow this space with a metric that makes it complete, in order to prove that the law of large numbers holds at the fixed point, it is sufficient to show that, on one hand, if this law of large numbers holds for a given probability law, it also holds for its image by the function; and that on the other hand, the function's iterates form a Cauchy sequence. This approach is similar to the one developed in [5], where propagation of chaos is proven in discrete time by showing that the one-step transition of the discrete dynamics preserves a triangular law of large numbers.

Our goal in this section is to prove the two aforementioned points. We start by introducing the metric space we will be considering and defining the function on it.

Fix $T \in \mathbb{R}$, and let D_T be the space of càdlàg functions on $[0, T]$ endowed with the Billingsley metric [12]: for $x, y \in D_T$, let

$$d_{D_T}(x, y) = \inf_{\theta \in \Theta} \max(\|\theta\|, \|x - y \circ \theta\|),$$

where

$$\Theta = \{\theta : [0, T] \rightarrow [0, T], \text{ s.t. } \theta(0) = 0, \theta(T) = T, \text{ and } \|\theta\| < \infty\},$$

where

$$\|\theta\| = \sup_{s \neq t \in [0, T]} \left| \log \left(\frac{\theta(t) - \theta(s)}{t - s} \right) \right|.$$

Intuitively, Θ represents all possible "reasonable" time shifts allowing one to minimize the effect of the jumps between the two functions x and y , where "reasonable" means that all slopes of θ are close to 1.

We denote by $d_{D_T,U}$ the uniform metric on D_T : for $x, y \in D_T$,

$$d_{D_T,U}(x, y) = \|x - y\|.$$

Note that we have for all $x, y \in D_T$, $d_{D_T}(x, y) \leq d_{D_T,U}(x, y)$, since the uniform metric corresponds precisely to the case where θ is the identity function.

Let $\mathcal{P}(D_T)$ be the space of probability measures on D_T . We endow it with the Kantorovitch metric [46] (also known as the Wasserstein distance or the earth mover's distance): for $\mu, \nu \in \mathcal{P}(D_T)$, let

$$K_T(\mu, \nu) = \inf_{\Pi \in D_T \times D_T} \mathbf{E}[d_{D_T}(x, y)],$$

where Π is a coupling s.t. $x \stackrel{\mathcal{L}}{=} \mu$ and $y \stackrel{\mathcal{L}}{=} \nu$.

Finally, we fix $K, M \in \mathbb{N}$ and consider the space $(\mathcal{P}(D_T))^{MK}$ endowed with the 1-norm metric: for $\mu, \nu \in \mathcal{P}(D_T)$, let

$$K_T^{MK}(\mu, \nu) = \sum_{m=1}^M \sum_{i=1}^K K_T(\mu_{m,i}, \nu_{m,i}).$$

It is known that (D_T, d_{D_T}) is a complete separable metric space, see [12], and thus that $(\mathcal{P}(D_T), K_T)$ and $(\mathcal{P}(D_T))^{MK}, K_T^{MK}$ are as well, see [14].

We will also need to consider $\mathcal{P}(D_T)$ endowed with a Kantorovitch metric based on the uniform metric: we introduce for $\mu, \nu \in \mathcal{P}(D_T)$,

$$K_{T,U}(\mu, \nu) = \inf_{\Pi \in D_T \times D_T} \mathbf{E}[d_{D_T,U}(x, y)],$$

where Π is a coupling s.t. $x \stackrel{\mathcal{L}}{=} \mu$ and $y \stackrel{\mathcal{L}}{=} \nu$. We also introduce its product version $K_{T,U}^{MK}$ defined analogously to above. Note that even though $(D_T, d_{D_T,U})$ is a complete metric space, it is not separable, therefore $(\mathcal{P}(D_T), K_{T,U})$ is not a priori a complete metric space.

We now define the following mapping:

$$\begin{aligned} \Phi: (\mathcal{P}(D_T))^{MK} &\rightarrow (\mathcal{P}(D_T))^{MK} \\ \mathcal{L}(M) &\mapsto \Phi(\mathcal{L}(M)), \end{aligned}$$

where for all $(m, i) \in \{1, \dots, M\} \times \{1, \dots, K\}$, $\Phi(\mathcal{L}(M))$ is the law of the stochastic

intensity $\lambda_{m,i}^\Phi$ of a point process $N_{m,i}^\Phi$ such that $\lambda_{m,i}^\Phi$ is the solution of the stochastic differential equation

$$\begin{aligned} \lambda_{m,i}^\Phi(t) &= \lambda_{m,i}^\Phi(0) + f\left(\sum_{j \neq i} \sum_{n \neq m} \int_0^t h_{j \rightarrow i}(s) \mathbb{I}_{\{V_{(n,j) \rightarrow i}^M(s)=m\}} M_{n,j \rightarrow i}(ds)\right) \\ &\quad + \int_0^t (g_i(s, \lambda_{m,i}^\Phi(s)) - \lambda_{m,i}^\Phi(s)) N_{m,i}^\Phi(ds) + \int_0^t (\sigma_i(s, \lambda_{m,i}^\Phi(s)) - \lambda_{m,i}^\Phi(s)) ds, \end{aligned} \quad (3.42)$$

where $(\lambda_{m,i}^\Phi(0))$ are random variables verifying Assumption 3.3.5. This is well-defined for the same reason (3.32) is.

We formalize the law of large numbers we aim to prove as follows:

Definition 3.3.18. *Let $M \in \mathbb{N}$. Let $(X_n^M)_{1 \leq n \leq M}$ be M -exchangeable random variables with finite expectation. We say they satisfy an \mathcal{L}^1 triangular law of large numbers, which we denote $\text{TLLN}(X_n^M)$, if when $M \rightarrow \infty$,*

$$\mathbf{E} \left[\left| \frac{1}{M-1} \sum_{n=1}^M (X_n^M - \mathbf{E}[X_n^M]) \right| \right] \rightarrow 0 \quad (3.43)$$

and

$$X_n^M \Rightarrow \tilde{X}, \quad (3.44)$$

where the convergence takes place in distribution.

From (3.41), we know that if the triangular law of large numbers holds for the fixed point of Φ , it allows for convergence in total variation of arrivals across replicas from a given neuron j to a given neuron i to a Poisson random variable. As such, our aim here is twofold:

1. Show that for all $(m, i) \in \{1, \dots, M\} \times \{1, \dots, K\}$, $\text{TLLN}(\overline{N_{m,i}}([0, t]))$ implies $\text{TLLN}(\Phi(\overline{N_{m,i}}([0, t])))$;
2. Show that $(\Phi^l)_{l \in \mathbb{N}^*}$ is a Cauchy sequence that converges to the fixed point.

Since we can choose $\overline{N_{m,i}}([0, t])$ to be i.i.d. to ensure that there exist inputs for which TLLN holds, this will allow us to propagate the property and show that TLLN holds at the fixed point as well.

We will start by proving a lemma that will be key for the second point. The adaptation of this lemma from [27] to the more general framework is the main challenge in this part, but assumptions made on f and Assumption 3.3.4 allow this extension.

Lemma 3.3.19. *There exists $T > 0$ such that for $\rho, \nu \in (\mathcal{P}(D_T))^{MK}$, there exists a constant $C_T > 0$ such that*

$$K_{T,U}^{MK}(\Phi(\rho), \Phi(\nu)) \leq C_T \int_0^T K_{t,U}^{MK}(\rho, \nu) dt. \quad (3.45)$$

Proof. Let $T > 0$. Let $t \in [0, T]$. Fix $(m, i) \in \{1, \dots, M\} \times \{1, \dots, K\}$. Let N^ρ (resp. $N^\nu, N^{\Phi(\rho)}, N^{\Phi(\nu)}$) be a point process admitting ρ (resp. $\nu, \Phi(\rho), \Phi(\nu)$) as a stochastic intensity. We have

$$\begin{aligned} & \Phi(\rho)_{m,i}(t) - \Phi(\nu)_{m,i}(t) = \\ & f \left(\sum_{j \neq i} \sum_{n \neq m} \left(\int_0^t h_{j \rightarrow i}(s) \mathbb{1}_{\{V_{(n,j)}^M \rightarrow i}(s) = m\}} (N_{n,j}^\rho(ds) - N_{n,j}^\nu(ds)) \right) \right) \\ & + \int_0^t (g_i(s, \Phi(\rho)_{m,i}(s)) - \Phi(\rho)_{m,i}(s)) N_{m,i}^{\Phi(\rho)}(ds) + \int_0^t (\sigma_i(s, \Phi(\rho)(s)) - \Phi(\rho)_{m,i}(s)) ds \\ & - \int_0^t (g_i(s, \Phi(\nu)_{m,i}(s)) - \Phi(\nu)_{m,i}(s)) N_{m,i}^{\Phi(\nu)}(ds) - \int_0^t (\sigma_i(s, \Phi(\nu)(s)) - \Phi(\nu)_{m,i}(s)) ds. \end{aligned}$$

Let $(\hat{N}_{m,i})_{(m,i) \in \{1, \dots, M\} \times \{1, \dots, K\}}$ be independent Poisson point processes with intensity 1 on $[0, T] \times \mathbb{R}^+$. Using the Poisson embedding construction, we can write

$$\begin{aligned} & \Phi(\rho)_{m,i}(t) - \Phi(\nu)_{m,i}(t) = \\ & f \left(\sum_{j \neq i} \sum_{n \neq m} \int_0^t \int_0^{+\infty} h_{j \rightarrow i}(s) \mathbb{1}_{\{V_{(n,j)}^M \rightarrow i}(s) = m\}} (\mathbb{1}_{\{u \leq \rho_{n,j}(s)\}} - \mathbb{1}_{\{u \leq \nu_{n,j}(s)\}}) \hat{N}_{n,j}(ds du) \right) \\ & + \int_0^t \int_0^{+\infty} g_i(s, \Phi(\rho)_{m,i}(s)) \mathbb{1}_{\{u \leq \Phi(\rho)_{m,i}(s)\}} - g_i(s, \Phi(\nu)_{m,i}(s)) \mathbb{1}_{\{u \leq \Phi(\nu)_{m,i}(s)\}} \hat{N}_{m,i}(ds du) \\ & + \int_0^t \int_0^{+\infty} \Phi(\nu)_{m,i}(s) \mathbb{1}_{\{u \leq \Phi(\nu)_{m,i}(s)\}} - \Phi(\rho)_{m,i}(s) \mathbb{1}_{\{u \leq \Phi(\rho)_{m,i}(s)\}} \hat{N}_{m,i}(ds du) \\ & + \int_0^t (\Phi(\nu)_{m,i}(s) - \Phi(\rho)_{m,i}(s)) ds + \int_0^t (\sigma_i(s, \Phi(\rho)_{m,i}(s)) - \sigma_i(s, \Phi(\nu)_{m,i}(s))) ds. \end{aligned}$$

Therefore, using the fact that f is Lipschitz and Assumption 3.3.4, we have

the existence of a constant $D > 0$ such that

$$\begin{aligned}
& |\Phi(\rho)_{m,i}(t) - \Phi(\nu)_{m,i}(t)| \leq \\
& D \sum_{j \neq i} \sum_{n \neq m} \int_0^t \int_0^{+\infty} h_{j \rightarrow i}(s) \mathbb{1}_{\{V_{(n,j)}^M \rightarrow i(s) = m\}} \mathbb{1}_{\{u \leq \sup_{z \in [0,s]} |\rho_{n,j}(z) - \nu_{n,j}(z)|\}} \hat{N}_{n,j}(ds du) \\
& + 2 \int_0^t \int_0^{+\infty} \sup_{z \in [0,s]} |\Phi(\rho)_{m,i}(z) - \Phi(\nu)_{m,i}(z)| \mathbb{1}_{\{u \leq \Phi(\rho)_{m,i}(s) \wedge \Phi(\nu)_{m,i}(s)\}} \hat{N}_{m,i}(ds du) \\
& + 2 \int_0^t \int_0^{+\infty} |\Phi(\rho)_{m,i}(s) \vee \Phi(\nu)_{m,i}(s)| \mathbb{1}_{\{u \leq \sup_{z \in [0,s]} |\Phi(\rho)_{m,i}(z) - \Phi(\nu)_{m,i}(z)|\}} \hat{N}_{m,i}(ds du) \\
& + 2 \int_0^t |\Phi(\nu)_{m,i}(s) - \Phi(\rho)_{m,i}(s)| ds.
\end{aligned}$$

Taking the expectation, using the property of stochastic intensity and proceeding as before to obtain the $\frac{1}{M-1}$ from the routing indicators, we get

$$\begin{aligned}
& \mathbf{E} \left[\sup_{t \in [0,T]} |\Phi(\rho)_{m,i}(t) - \Phi(\nu)_{m,i}(t)| \right] \leq \\
& \frac{DH}{M-1} \sum_{j \neq i} \sum_{n \neq m} \int_0^T \mathbf{E} \left[\sup_{z \in [0,s]} |\rho_{n,j}(z) - \nu_{n,j}(z)| \right] ds \\
& + 2 \int_0^T \mathbf{E} \left[\sup_{z \in [0,s]} |\Phi(\rho)_{m,i}(z) - \Phi(\nu)_{m,i}(z)| (\Phi(\rho)_{m,i}(s) \wedge \Phi(\nu)_{m,i}(s)) \right] ds \\
& + 2 \int_0^T \mathbf{E} \left[\sup_{z \in [0,s]} |\Phi(\rho)_{m,i}(z) - \Phi(\nu)_{m,i}(z)| (\Phi(\rho)_{m,i}(s) \vee \Phi(\nu)_{m,i}(s)) \right] ds \\
& + 2 \int_0^T \mathbf{E} \left[\sup_{z \in [0,s]} |\Phi(\rho)_{m,i}(z) - \Phi(\nu)_{m,i}(z)| \right] ds.
\end{aligned}$$

We then have

$$\begin{aligned}
& \frac{DH}{M-1} \sum_{j \neq i} \mu_{j \rightarrow i} \sum_{n \neq m} \int_0^T \mathbf{E} \left[\sup_{z \in [0,s]} |\rho_{n,j}(z) - \nu_{n,j}(z)| \right] ds \leq \\
& DH \sum_{j=1}^K \sum_{n=1}^M \int_0^T d_{D_s, U}(\rho_{n,j}, \nu_{n,j}) ds,
\end{aligned}$$

from which we immediately get by definition of $K_{T,U}^{MK}$

$$\frac{DH}{M-1} \sum_{j \neq i} \sum_{n \neq m} \int_0^T \mathbf{E} \left[\sup_{z \in [0,s]} |\rho_{n,j}(z) - \nu_{n,j}(z)| \right] ds \leq DH \int_0^T K_{s,U}^{MK}(\rho, \nu) ds. \quad (3.46)$$

Let $C > 0$. As before, let $A_C([0, T]) = \{(\omega, t) \in \Omega \times [0, T], \Phi(\rho)_{m,i}(t) \vee \Phi(\nu)_{m,i}(t) > C\}$. Using the exact same reasoning as in [27], we have the existence of a constant $K_T > 0$ such that

$$\begin{aligned} & \int_0^t \mathbf{E} \left[\sup_{z \in [0, s]} |\Phi(\rho)_{m,i}(z) - \Phi(\nu)_{m,i}(z)| (\Phi(\rho)_{m,i}(s) \wedge \Phi(\nu)_{m,i}(s)) \right] ds \\ & \leq C \int_0^t \mathbf{E} \left[\sup_{z \in [0, s]} |\Phi(\rho)_{m,i}(z) - \Phi(\nu)_{m,i}(z)| \right] ds + K_T e^{-3CT}. \end{aligned}$$

Plugging in (3.46) and applying the same reasoning as above to the last integral term, we get the existence of a constant $K'_T > 0$ such that

$$\begin{aligned} & \mathbf{E} \left[\sup_{t \in [0, T]} |\Phi(\rho)_{m,i}(t) - \Phi(\nu)_{m,i}(t)| \right] \leq DH \int_0^T K_{s,U}^{MK}(\rho, \nu) ds \\ & + (2(1+C)) \int_0^T \mathbf{E} \left[\sup_{z \in [0, s]} |\Phi(\rho)_{m,i}(z) - \Phi(\nu)_{m,i}(z)| \right] ds \\ & + (K_T + K'_T) e^{-3CT}. \end{aligned}$$

Applying Grönwall's lemma, we get

$$\begin{aligned} & \mathbf{E} \left[\sup_{t \in [0, T]} |\Phi(\rho)_{m,i}(t) - \Phi(\nu)_{m,i}(t)| \right] \leq \\ & \left(DH \int_0^T K_{s,U}^{MK}(\rho, \nu) ds + (K_T + K'_T) e^{-3CT} \right) e^{2(1+C)T}. \end{aligned}$$

For any $\varepsilon > 0$, we can choose $C > 0$ such that

$$\mathbf{E} \left[\sup_{t \in [0, T]} |\Phi(\rho)_{m,i}(t) - \Phi(\nu)_{m,i}(t)| \right] \leq \left(DH \int_0^T K_{s,U}^{MK}(\rho, \nu) ds \right) e^{2(1+C)T} + \varepsilon.$$

Letting ε go to 0 and taking the sum over all coordinates and the infimum across all couplings, we get the result. \square

As previously mentioned, we need to prove convergence of the sequence of iterates of Φ to the fixed point of Φ to prove the triangular law of large numbers. We will now derive this from Lemma 3.3.19.

Lemma 3.3.20. *The sequence $(\Phi^l)_{l \in \mathbb{N}^*}$ of iterates of the function Φ is a Cauchy sequence. Moreover, it converges to the unique fixed point of Φ .*

Proof. Let $\rho \in (\mathcal{P}(D_T))^{MK}$. By immediate induction, from (3.45), we have, for all $l \in \mathbb{N}^*$,

$$K_{T,U}^{MK}(\Phi^{l+1}(\rho), \Phi^l(\rho)) \leq C_T \frac{T^l}{l!} K_{T,U}^{MK}(\Phi(\rho), \rho).$$

This in turn implies that for any $p < q \in \mathbb{N}^*$,

$$K_{T,U}^{MK}(\Phi^p(\rho), \Phi^q(\rho)) \leq \sum_{l=p}^{q-1} \frac{C_T T^l}{l!} K_{T,U}^{MK}(\Phi(\rho), \rho). \quad (3.47)$$

Since the series on the right hand side is converging, it proves that the sequence $(\Phi^l)_{l \in \mathbb{N}^*}$ is a Cauchy sequence for the $K_{T,U}^{MK}$ metric. The space $(\mathcal{P}(D_T))^{MK}, K_{T,U}^{MK}$ is not complete. However, since for any $\mu, \nu \in D_T, d_{D_T}(\mu, \nu) \leq d_{D_T,U}(\mu, \nu)$, it follows that $(\Phi^l)_{l \in \mathbb{N}^*}$ is a Cauchy sequence for the K_T^{MK} metric as well. By completeness of $(\mathcal{P}(D_T))^{MK}, K_T^{MK}$, $(\Phi^l)_{l \in \mathbb{N}^*}$ converges to the unique fixed point of Φ . \square

All that remains is proving that the triangular law of large numbers is carried over by the function Φ , namely, that if we have some input X that verifies $\text{TLLN}(X)$, then we have $\text{TLLN}(\Phi(X))$.

To do so, the key lemma will be the following law of large numbers.

Lemma 3.3.21. *Let $M \in \mathbb{N}^*$. Let (X_1^M, \dots, X_M^M) be M -exchangeable centered random variables with finite exponential moments. Suppose that for any $N \in \mathbb{N}^*$, $(X_1^M, \dots, X_N^M) \xrightarrow{\mathcal{L}} (\tilde{X}_1, \dots, \tilde{X}_N)$ when $M \rightarrow \infty$, where $(\tilde{X}_i)_{i \in \mathbb{N}^*}$ are i.i.d. random variables and the convergence takes place in distribution. Then*

$$\mathbf{E} \left[\left| \frac{1}{M} \sum_{n=1}^M X_n^M \right| \right] \rightarrow 0 \quad (3.48)$$

when $M \rightarrow \infty$.

We refer to [27] for the proof of this result.

The following lemma is the last step needed to prove the main theorem:

Lemma 3.3.22. *Let $(\overline{N_{m,i}})$ be point processes on $[0, T]$ with finite exponential moments. Let $t \in [0, T]$. Suppose $\text{TLLN}(\overline{N_{m,i}}([0, t]))$ holds. Then, $\text{TLLN}(\Phi(\overline{N_{m,i}}([0, t])))$ holds as well.*

We once again refer to [27] for the proof of this result, as it is done in an analogous matter.

Thus, we can now state the result that we were aiming to prove:

Lemma 3.3.23. *Denote by $(N_{m,i})$ the point processes of the M -replica RMF dynamics (3.32) that are the fixed point of Φ . Then $\text{TLLN}((N_{m,i}([0, T])))$ holds.*

Proof. The proof is identical to [27], but we reproduce it here for completeness' sake. Let $(\overline{N}_{n,j})$ be random variables satisfying $\text{TLLN}(\overline{N}_{n,j})$. Let us first write out equalities and justify them afterwards. We have

$$\begin{aligned}
 & \lim_{M \rightarrow \infty} \frac{1}{M-1} \mathbf{E} \left[\left| \sum_{n \neq m} (\mathbf{E}[N_{n,j}([0, T])] - N_{n,j}([0, T])) \right| \right] \\
 &= \lim_{M \rightarrow \infty} \frac{1}{M-1} \mathbf{E} \left[\left| \sum_{n \neq m} \left(\mathbf{E}[\lim_{l \rightarrow \infty} \Phi^l(\overline{N}_{n,j}([0, T]))] - \lim_{l \rightarrow \infty} \Phi^l(\overline{N}_{n,j}([0, T])) \right) \right| \right] \\
 &= \lim_{M \rightarrow \infty} \lim_{l \rightarrow \infty} \frac{1}{M-1} \mathbf{E} \left[\left| \sum_{n \neq m} (\mathbf{E}[\Phi^l(\overline{N}_{n,j}([0, T]))] - \Phi^l(\overline{N}_{n,j}([0, T]))) \right| \right] \\
 &= \lim_{l \rightarrow \infty} \lim_{M \rightarrow \infty} \frac{1}{M-1} \mathbf{E} \left[\left| \sum_{n \neq m} (\mathbf{E}[\Phi^l(\overline{N}_{n,j}([0, T]))] - \Phi^l(\overline{N}_{n,j}([0, T]))) \right| \right] \\
 &= 0.
 \end{aligned}$$

The first equality is given by Lemma 3.3.20. To obtain the second equality, we use the dominated convergence theorem and the fact that all moments are uniformly bounded through Lemma 3.3.14 (note that initial conditions are fixed in the definition of Φ and are chosen to verify Assumption 3.3.5). To justify the third equality, note that from (3.47), using Lemma 3.3.14 to obtain again a uniform bound of the moments, we get that the Cauchy sequence of iterates of Φ verifies the uniform Cauchy criterion and thus converges uniformly to the fixed point, which in turn allows for the exchange of limits in M and l . The last equality stems directly from Lemma 3.3.22. \square

3.3.3 Link between RMF FIAPs in discrete and continuous times : the example of the excitatory Galves-Löcherbach model

FIAPs, and replica-mean-field versions of FIAPs, were originally introduced in discrete time in [5]. One natural question is to explore the links between these original FIAPs and the cFIAPS introduced in this work. The goal of this section is to show a link between replica-mean-field versions of continuous-time and discrete-time FIAPs for a specific instance of FIAP : the excitatory Galves-Löcherbach model. In this particular case, the goal is to prove the existence of the horizontal

equivalences in the following diagram:

$$\begin{array}{ccc}
 (\text{infinite-replica FIAP with } \delta\text{-time step})_{\delta>0} & \longleftrightarrow & \text{infinite-replica GL cFIAP} \\
 \uparrow M \rightarrow \infty & & \uparrow M \rightarrow \infty \\
 (M\text{-replica FIAP with } \delta\text{-time step})_{\delta>0} & \longleftrightarrow & M\text{-replica excitatory GL cFIAP}
 \end{array}$$

The left up arrow corresponds to the proof of the Poisson Hypothesis for a collection of discrete-time FIAPs as introduced in [5] with time-step δ , for all $\delta > 0$. The right up arrow corresponds to the proof of the Poisson Hypothesis for the cFIAP excitatory GL model, which was presented in [27]. In this section we complete the diagram by showing that it is possible to construct the discrete-time RMF FIAPs given the continuous-time dynamics, and vice-versa.

We recall the definition of discrete-time FIAPs originally introduced in [5]:

Definition 3.3.24. *An instance of the class \mathcal{C} of discrete fragmentation-interaction-aggregation processes is determined by:*

- *An integer K representing the number of nodes;*
- *A collection of initial conditions for the integer-valued state variables at step zero, which we denote by $\{X_i\}$, where $i \in \{1, \dots, K\}$;*
- *A collection of fragmentation random variables $\{U_i\}$, which are i.i.d. uniform in $[0, 1]$ and independent from $\{X_i\}$, where $i \in \{1, \dots, K\}$;*
- *A collection of fragmentation functions $\{g_{1,i} : \mathbb{N} \rightarrow \mathbb{N}\}_{i \in \{1, \dots, K\}}$ and $\{g_{2,i} : \mathbb{N} \rightarrow \mathbb{N}\}_{i \in \{1, \dots, K\}}$;*
- *A collection of bounded interaction functions $\{h_{j \rightarrow i} : \mathbb{N} \rightarrow \mathbb{N}\}_{i,j \in \{1, \dots, K\}}$;*
- *A collection of activation probabilities $\{\sigma_i(0), \sigma_i(1), \dots\}_{i \in \{1, \dots, K\}}$ verifying the conditions $\sigma_i(0) = 0$, and $0 < \sigma_i(1) \leq \sigma_i(2) \leq \dots \leq 1$ for all i .*

The associated dynamics take as input the initial integer-valued state variables $\{X_i\}$ and define the state variables at the next step as

$$Y_i = g_{1,i}(X_i) \mathbb{1}_{\{U_i < \sigma_i(X_i)\}} + g_{2,i}(X_i) \mathbb{1}_{\{U_i > \sigma_i(X_i)\}} + A_i, \quad \forall i = 1, \dots, K, \quad (3.49)$$

with arrival processes

$$A_i = \sum_{j \neq i} h_{j \rightarrow i}(X_j) \mathbb{1}_{\{U_j < \sigma_j(X_j)\}}, \quad \forall i = 1, \dots, K. \quad (3.50)$$

Given a FIAP, we now consider its replica-mean-field model. We once again recall the precise definition from [5]:

Definition 3.3.25. For any process in \mathcal{C} , the associated M -replica dynamics is entirely specified by

- A collection of initial conditions for the integer-valued state variables at step zero, which we denote by $\{X_{n,i}^M\}$, where $n \in \{1, \dots, M\}$ and $i \in \{1, \dots, K\}$, such that for all M, n and i , $X_{n,i}^M = X_i$;
- A collection of fragmentation random variables $\{U_{n,i}\}$, which are i.i.d. uniform in $[0, 1]$ and independent from $\{X_{n,i}^M\}$, where $n \in \{1, \dots, M\}$ and $i \in \{1, \dots, K\}$;
- A collection of i.i.d. routing random variables $\{R_{(n,j) \rightarrow i}^M\}$ independent from $\{X_{n,i}^M\}$ and $\{U_{n,i}\}$, uniformly distributed on $\{1, \dots, M\} \setminus \{m\}$ for all $i, j \in \{1, \dots, K\}$ and $m \in \{1, \dots, M\}$. In other words, if $R_{(n,j) \rightarrow i}^M = n$, then an eventual activation of node j in replica m at step 0 induces an arrival of size $h_{j \rightarrow i}(X_{m,j}^M)$ in node i of replica n , and n is chosen uniformly among replicas and independently from the state variables. Note that these variables are defined regardless of the fact that an activation actually occurs. Also note that for $i' \neq i$, the activation in question will induce an arrival in node i' of replica n' , with n' sampled in the same way but independently of n .

Then, the integer-valued state variables at step one, denoted by $\{Y_{n,i}^M\}$, are given by the M -RMF equations

$$Y_{n,i}^M = g_{1,i}(X_{n,i}^M) \mathbb{I}_{\{U_{n,i} < \sigma_i(X_{n,i}^M)\}} + g_{2,i}(X_{n,i}^M) \mathbb{I}_{\{U_{n,i} > \sigma_i(X_{n,i}^M)\}} + A_{n,i}^M, \quad (3.51)$$

where $g_{1,i}$, $g_{2,i}$ denotes fragmentation functions, σ_i denotes activation probabilities, and where

$$A_{n,i}^M = \sum_{m \neq n} \sum_{j \neq i} h_{j \rightarrow i}(X_{m,j}^M) \mathbb{I}_{\{U_{m,j} < \sigma_i(X_{m,j}^M)\}} \mathbb{I}_{\{R_{(n,j) \rightarrow i}^M = n\}} \quad (3.52)$$

is the number of arrivals to node i of replica n via the interaction functions $h_{j \rightarrow i}$.

We will focus here on the case of the excitatory Galves-Löcherbach model (we will omit excitatory hereafter), in both discrete and continuous time settings, which we will now recall.

In the continuous time setting, the M -replica-mean-field of the Galves-Löcherbach model is defined as follows:

$$\begin{aligned} \lambda_{m,i}^M(t) &= \lambda_{m,i}^M(0) + \frac{1}{\tau_i} \int_0^t (b_i - \lambda_{m,i}^M(s)) ds \\ &+ \sum_{j \neq i} \mu_{j \rightarrow i} \sum_{n \neq m} \int_0^t \mathbb{I}_{\{V_{(n,j) \rightarrow i}^M(s) = m\}} N_{n,j}^M(ds) + \int_0^t (r_i - \lambda_{m,i}^M(s)) N_{m,i}^M(ds). \end{aligned} \quad (3.53)$$

In the discrete time setting with time step length δ , the one-step transition of the RMF GL FIAP is given by

$$Y_{m,i}^M = \mathbb{1}_{\{U_{m,i} > \sigma_\delta(X_{m,i}^M)\}} X_{m,i}^M + \mathbb{1}_{\{U_{m,i} > \sigma_\delta(X_{m,i}^M)\}} r_i + A_{m,i}^M, \quad (3.54)$$

where

$$A_{m,i}^M = \sum_{n \neq m} \sum_{j \neq i} \mu_{j \rightarrow i} \mathbb{1}_{\{U_{n,j} < \sigma_\delta(X_{n,j}^M)\}} \mathbb{1}_{\{R_{n,j,i}^M = m\}} \quad (3.55)$$

is the number of arrivals to neuron i of replica m , $X_{m,i}^M$ is the state of neuron i in replica m at time 0 and $Y_{m,i}^M$ is its state at time 1.

In FIAP models, the arrivals to a given neuron at a given time are conditionally independent from the spiking activity of that neuron given the states at that time. Since in a continuous-time model all events are asynchronous, such a property is no longer verified. Thus, in order to map the continuous-time model to a FIAP, we must “separate” the arrivals and the spikes. In order to do that, we introduce a $\delta > 0$ unit of time. We then show that one can construct a discrete-time Markov chain that is similar to the embedded Markov chain of the continuous time model but which belongs to the RMF FIAP class.

Since in FIAPs, all the states taken by the state variables are discrete, we must make the following simplifying assumption:

Assumption 3.3.26.

- For all $i \in \{1, \dots, K\}$, $\tau_i = \infty$ (no exponential decay) and $r_i \in \mathbb{N}^*$;
- For all $i, j \in \{1, \dots, K\}$, $\mu_{j \rightarrow i} \in \mathbb{N}$.

Under Assumption 3.3.26, it is known (see [8]) that the generator of the M -replica dynamics is given by

$$\mathcal{A}[f](\lambda) = \sum_{i=1}^K \sum_{n=1}^M \frac{1}{|\mathcal{V}_{m,i}|} \sum_{v \in \mathcal{V}_{m,i}} (f(\lambda + \mu_{m,i,v}(\lambda)) - f(\lambda)) \lambda_{m,i},$$

where the update due to the spiking of neuron (m, i) is defined by

$$[\mu_{m,i,v}(\lambda)]_{n,j} = \begin{cases} \mu_{j \rightarrow i} & \text{if } j \neq i, n = v_j \\ r_i - \lambda_{m,i} & \text{if } j = i, n = m \\ 0 & \text{otherwise .} \end{cases}$$

We now consider the embedded discrete-time Markov chain of the RMF GL dynamics, where all updates happen at the spiking times of the dynamics. Since

the spiking times are all distinct, all the transitions of the embedded Markov chain are given by

$$\begin{array}{c} (\lambda_{1,1}, \dots, \lambda_{1,K}, \lambda_{2,1}, \dots, \lambda_{M,K}) \\ \downarrow \\ (\lambda_{1,1}, \dots, \lambda_{m_1,1} + \mu_{i,1}, \dots, \lambda_{m,i-1}, r_i, \lambda_{m,i+1}, \dots, \lambda_{M,K}) \end{array}$$

with associated transition probabilities $p_{m,i}^M \left(\frac{1}{M-1}\right)^{K-1}$, where $p_{m,i}^M$ is the probability that neuron (m, i) spikes conditioned on the event that a spike happens. The main complexity with this Markov chain is that the transition times correspond to the spiking times of the RMF GL network, which are not tractable.

Therefore, we now define a new discrete time Markov chain on \mathbb{N}^{MK} with steps in time of fixed length δ . The informal idea is to reset all neurons that spike in the RMF GL dynamics during such a step in time and update all the states with the potential due to these spikes. Since a single neuron could very well spike multiple times in a δ unit of time, we only consider the updates due to the first spike of a given neuron.

Note that the informal idea given above is the motivation behind the definition of the following discrete time dynamics, which is defined *ad hoc*. We characterize the dynamics of this chain through its transitions, and we will show that the chain we define belongs to the class of discrete time RMF FIAPs.

In order to simplify notation and facilitate understanding, let us define the following “half-step” fragmentation and aggregation transitions $\mathcal{P}^A(\delta)$ and $\mathcal{P}^F(\delta)$. We define $\mathcal{P}^F(\delta)$ as the transitions

$$\begin{array}{c} \mathcal{P}^F(\delta) : (\lambda_{1,1}, \dots, \lambda_{1,K}, \lambda_{2,1}, \dots, \dots, \lambda_{M,K}) \\ \downarrow \\ (\lambda_{1,1}, \dots, r_{i_1}, \dots, r_{i_2}, \dots, r_{i_L}, \dots, \lambda_{M,K}) \end{array}$$

where the transition probability is given by

$$\prod_{(m,i) \in \mathcal{L}} p_{m,i}^M(\delta) \prod_{(m,i) \notin \mathcal{L}} (1 - p_{m,i}^M(\delta)),$$

where $p_{m,i}^M(\delta)$ is the probability that neuron (m, i) spikes, i.e., is set to the value r_i , in a δ -unit of time and $\mathcal{L} = \{(m_1, i_1), \dots, (m_L, i_L)\}$.

Informally, this transition corresponds to a *fragmentation* of the state: we reset to their base rate all the neurons that spike during the δ step of time. Note that $p_{m,i}^M(\delta) = 1 - e^{-\sigma(\lambda_{m,i})\delta}$, where for all k , $\sigma(k)$ is the probability that a neuron of the RMF GL network in state k spikes in a unit of time in length 1.

We then define the *aggregation* transitions $\mathcal{P}^A(\delta)$ in the following fashion : for all $1 \leq k \leq L$, each neuron (m_k, i_k) which has spiked and been reset to its base rate r_{i_k} in the previous step, for each $j \neq i_k$, we randomly, uniformly and independently from each other and from the rest of the dynamics, choose an index $n_k \neq m_k$ and increment neuron (n_k, j) by $\mu_{i_k, j}$. Thus, all transitions are of the form

$$\begin{aligned} \mathcal{P}^A(\delta) : (\mathcal{P}^F(\delta) (\lambda_{1,1}, \dots, \lambda_{M,K})) \\ \downarrow \\ (\lambda_{1,1}, \dots, \lambda_{n_{i_1}, 1} + \mu_{i_1, 1}, \dots, \lambda_{n_{i_k}, j} + \mu_{i_k, j}, \dots, \lambda_{M,K}) \end{aligned}$$

with transition probability $(\frac{1}{M-1})^{(K-1)L}$ (conditioned on the transition probability of $\mathcal{P}^F(\delta)$). Note that (n_{i_k}) are not necessarily distinct for $k' \neq k$. Note that all the routings are done independently from one another and that updates to a single neuron are independent of whether that particular neuron has spiked or not.

We can then define the full transitions of our new discrete time Markov chain as

$$(\lambda_{1,1}, \dots, \lambda_{M,K}) \mapsto \mathcal{P}^A(\delta) \circ \mathcal{P}^F(\delta) (\lambda_{1,1}, \dots, \lambda_{M,K}). \quad (3.56)$$

By the total probability formula, the transition probability is given by

$$\sum_{l=1}^{MK} \sum_{\substack{J \subset \{1, \dots, M\} \times \{1, \dots, K\} \\ |J|=l}} \prod_{(n,j) \in J} p_{n,j}^M(\delta) \prod_{(n,j) \notin J} (1 - p_{n,j}^M(\delta)) \left(\frac{1}{M-1} \right)^l. \quad (3.57)$$

Our goal is now to show that the Markov chain defined this way belongs to the class of RMF FIAPs. In order to do that, we will require the following lemma, which gives the transition probability of a single coordinate of the above Markov chain.

Lemma 3.3.27. *Let $k, l \in \mathbb{N}$. Let $P_{k \rightarrow l}^{M, m, i}(\delta)$ be the probability that neuron (m, i) following the RMF dynamics given above in (3.56) transitions from state k to state l in a single step of time of length δ . Then*

$$P_{k \rightarrow l}^{M, m, i} = p_{m,i}^M(\delta) \mathbf{P}(A_{m,i}^M = l - r_i) + (1 - p_{m,i}^M(\delta)) \mathbf{P}(A_{m,i}^M = l - k), \quad (3.58)$$

where for all $l \in \mathbb{N}$,

$$\begin{aligned} \mathbf{P}(A_{m,i}^M = l) = & \sum_{p=l}^{(K-1)(M-1)} \sum_{\substack{J \subset \{1, \dots, M\} \setminus \{m\} \times \{1, \dots, K\} \setminus \{i\} \\ \sum_{(n,j) \in J} \mu_{j \rightarrow i} = p}} \prod_{(n,j) \in J} p_{n,j}^M(\delta) \\ & \prod_{(n,j) \notin J} (1 - p_{n,j}^M(\delta)) \left(\frac{1}{M-1} \right)^l \left(1 - \frac{1}{M-1} \right)^{p-l}. \end{aligned} \quad (3.59)$$

Here $A_{m,i}^M$ are the updates due to arrivals to neuron (m, i) in a δ -unit of time.

Proof. Equation (3.58) is due to the independence between the spiking of (m, i) and the arrivals to (m, i) . Let $S_{m,i}^M$ be the arrivals caused by spikes in the system in a δ -unit of time, discounting the spikes in replica m and the spikes in neuron i across replicas. Informally, $S_{m,i}^M$ is the quantity from spikes that could potentially reach neuron i in replica m if the routing variables allow it. Then

$$\mathbf{P}(S_{m,i}^M = p) = \sum_{\substack{J \subset \{1, \dots, M\} \setminus \{m\} \times \{1, \dots, K\} \setminus \{i\} \\ \sum_{(n,j) \in J} \mu_{j \rightarrow i} = p}} \prod_{(n,j) \in J} p_{n,j}^M(\delta) \prod_{(n,j) \notin J} (1 - p_{n,j}^M(\delta)). \quad (3.60)$$

By the total probabilities formula,

$$\mathbf{P}(A_{m,i}^M = l) = \sum_{p=l}^{(M-1)(K-1)} \mathbf{P}(S_{m,i}^M = p) \mathbf{P}(A_{m,i}^M = l | S_{m,i}^M = p). \quad (3.61)$$

Since the routings are independent of the rest of the process, we have

$$\mathbf{P}(A_{m,i}^M = l | S_{m,i}^M = p) = \left(\frac{1}{M-1} \right)^l \left(1 - \frac{1}{M-1} \right)^{p-l}. \quad (3.62)$$

Putting together (3.60), (3.61) and (3.62), we get (3.59). \square

Now we prove the following lemma:

Lemma 3.3.28. *The discrete-time Markov chain with transitions defined by (3.56) is a RMF FIAP.*

Proof. In order to achieve that, we compute the transition probabilities of RMF FIAPs and we show that the transition probabilities of the Markov chain defined above are of that type.

Consider the RMF GL FIAP model following dynamics given by (3.54) and (3.55). Let $X = \{X_{m,i}^M\}$ be the state variables at step 0, let $\sigma_\delta : \mathbb{N} \rightarrow [0, 1]$ be the spiking probabilities of the neuron satisfying the conditions given in Definition 3.3.24.

We now give the transition probabilities of the Markov RMF FIAP dynamics. Let $k, l \in \mathbb{N}$. Let $Q_{k \rightarrow l}^{M,m,i}$ be the probability that neuron (m, i) in the RMF FIAP dynamics is in state l at time δ given that it is in state k at time 0. In other words, $Q_{k \rightarrow l}^{M,m,i} = \mathbf{P}(Y_{m,i}^M = l | X_{m,i}^M = k)$. Then

$$Q_{k \rightarrow l}^{M,m,i} = \sigma_\delta(X_{m,i}^M) \mathbf{P}(A_{m,i}^M = l - r_i) + (1 - \sigma_\delta(X_{m,i}^M)) \mathbf{P}(A_{m,i}^M = l - k), \quad (3.63)$$

where

$$\begin{aligned} \mathbf{P}(A_{m,i}^M = l) &= \sum_{p=l}^{(K-1)(M-1)} \sum_{\substack{J \subset \{1, \dots, M\} \setminus \{m\} \times \{1, \dots, K\} \setminus \{i\} \\ \sum_{(n,j) \in J} \mu_{j \rightarrow i} = p}} \prod_{(n,j) \in J} \sigma_\delta(X_{n,j}^M) \\ &\quad \prod_{(n,j) \notin J} (1 - \sigma_\delta(X_{n,j}^M)) \left(\frac{1}{M-1}\right)^l \left(1 - \frac{1}{M-1}\right)^{p-l}. \end{aligned} \quad (3.64)$$

Here $A_{m,i}^M$ are the updates due to arrivals to neuron (m, i) in a δ -unit of time.

Note that in the case where $\mu_{j \rightarrow i} = 1$ for all i, j , $A_{m,i}^M$ represents the number of arrivals to neuron (m, i) in a δ -unit of time.

Equation (3.63) is simply due to the independence of arrivals and spikes in the RMF FIAP model. We now proceed identically to the proof of the particular case above. Let $S_{m,i}^M$ be the arrivals caused by spikes in the system in a δ -unit of time, discounting the spikes in replica m and the spikes in neuron i across replicas. Then

$$\mathbf{P}(S_{m,i}^M = p) = \sum_{\substack{J \subset \{1, \dots, M\} \setminus \{m\} \times \{1, \dots, K\} \setminus \{i\} \\ \sum_{(n,j) \in J} \mu_{j \rightarrow i} = p}} \prod_{(n,j) \in J} \sigma(X_{n,j}^M) \prod_{(n,j) \notin J} (1 - \sigma(X_{n,j}^M)). \quad (3.65)$$

By the total probabilities formula,

$$\mathbf{P}(A_{m,i}^M = l) = \sum_{p=l}^{(M-1)(K-1)} \mathbf{P}(S_{m,i}^M = p) \mathbf{P}(A_{m,i}^M = l | S_{m,i}^M = p). \quad (3.66)$$

Since the routing variables are independent of the rest of the process, we have

$$\mathbf{P}(A_{m,i}^M = l | S_{m,i}^M = p) = \left(\frac{1}{M-1}\right)^l \left(1 - \frac{1}{M-1}\right)^{p-l}. \quad (3.67)$$

Putting together (3.65), (3.66) and (3.67), we get (3.64).

As such, we see that the transition probabilities in our Markov chain model given by (3.58) and (3.59) are a particular case of these general FIAP transition probabilities (3.63) and (3.64). This concludes the proof. \square

In this way, we have shown that given a RMF GL continuous-time model, given the initial conditions, we can uniquely define a collection of RMF FIAP discrete-time models with varying time step lengths associated with it.

Note that a reverse construction is also possible in the following sense: given RMF FIAP dynamics of the type defined above for all $\delta > 0$, since for all k , $\sigma_\delta(k) = \delta\sigma(k) + o(\delta)$, we can reconstruct the infinitesimal generator of the continuous-time dynamics by considering the transition operator $\frac{1}{\delta}(P_\delta - Id)$, where Id is the identity operator and P_δ is the transition operator of the RMF FIAP dynamics with time steps of length δ , and letting δ go to 0.

Chapter 4

Replica-mean-field limit of a migration-contagion process on a network

4.1 Introduction

This chapter presents elements of a work in progress on an epidemiological model. We introduce a replica-mean-field model of epidemics propagation on a network. Unlike previous chapters, we are interested here not in the proof of the convergence of the replica-mean-field model to limit dynamics under the Poisson Hypothesis, but rather in the numerical study of the conjectured limit. More specifically, the main goal of this chapter is to study the relationship between the geometry of the underlying network and the propagation of the epidemics under the Poisson Hypothesis.

4.2 Replica-mean-field limit of a migration-contagion process on a network

4.2.1 Introduction

In a migration-contagion process, individuals, or customers, travel in between stations, called *reactors* when considering them as single open stations. Each customer is marked with their state, infected (I) or susceptible (S). Inside each reactor, the customer states change according to mechanisms depending on the migration-contagion process chosen. In [6], several reactors were introduced in order to study the steady state of the queues. The main $M/M/\infty$ SIS reactor is

defined as follows: consider a $M/M/\infty$ queue with input rate λ and service rate μ . Let $X(t)$ be the number of susceptible customers at time t , and $Y(t)$ be the number of infected customers at time t . Let $X(0) = x, Y(0) = y$. Input customers are infected with probability p , and susceptible with probability $q = 1 - p$. Inside the queue, any susceptible customer becomes infected with rate $\alpha Y(t)$, and any infected customer becomes susceptible with rate β .

The SIS-DOCS (Departure On Change of State) migration-contagion reactor is defined as follows: if a customer changes states, they immediately leave the reactor. In this setting, the PDE verified by the stationary generating function can be solved explicitly. Moreover, the stationary distribution of Y is Poisson with parameter $\frac{\lambda p}{\mu + \beta}$. Similarly to the SIS case, it is possible to consider the thermodynamic limit of N SIS-DOCS stations. The rigorous derivation of this limit is left as a conjecture.

Here, we study a network version of the SIS-DOCS dynamics. Namely, we consider a network where each node is a SIS-DOCS reactor. Upon exiting a reactor, each customer chooses a neighboring node and travels to it.

For $1 \leq i \leq K$, denote μ_i the service rate of reactor i and λ_i its arrival rate. Similarly, denote α_i and β_i its parameters for the rates of change of state, from susceptible to infected and from infected to susceptible respectively.

Formally, we consider N individuals traveling on a network with K nodes. The geometry of the network is encoded in the routing matrix $P = (p_{j \rightarrow i})_{1 \leq i, j \leq K}$. To transcribe the routings in between nodes, we introduce routing processes $\{W_j^{x,1}(t)\}_{t \in \mathbb{R}}, \{W_j^{y,1}(t)\}_{t \in \mathbb{R}}, \{W_j^{x,2}(t)\}_{t \in \mathbb{R}}$ and $\{W_j^{y,2}(t)\}_{t \in \mathbb{R}}$ for $1 \leq j \leq K$, such that, for each interaction time T , the random variables $\{W_j(T)\}_j$ are mutually independent, independent from the past, and distributed on $\mathcal{V}(j) = \{i \neq j, p_{j \rightarrow i} > 0\}$ with weights $(p_{j \rightarrow i})$.

Then, for $t \geq 0$, the state equations of node $i \in \{1, \dots, K\}$ are given by

$$\begin{aligned}
 X_i(t) = & X_i(0) + \sum_{j \neq i} \int_0^t \int_0^{+\infty} \mathbb{1}_{\{W_j^{x,1}(s)=i\}} \mathbb{1}_{\{u \leq \mu_j X_j(s)\}} \widehat{N}_j^{x,1}(ds du) \\
 & + \int_0^t \int_0^{+\infty} \mathbb{1}_{\{W_j^{x,2}(s)=i\}} \mathbb{1}_{\{u \leq \beta_j Y_j(s)\}} \widehat{N}_j^{y,2}(ds du) \\
 & - \int_0^t \int_0^{+\infty} (\mathbb{1}_{\{u \leq \alpha_i X_i(s) Y_i(s)\}} \widehat{N}_i^{x,2}(ds du) + \mathbb{1}_{\{u \leq \mu_i X_i(s)\}} \widehat{N}_i^{x,1}(ds du))
 \end{aligned} \tag{4.1}$$

and

$$\begin{aligned}
 Y_i(t) = & Y_i(0) + \sum_{j \neq i} \int_0^t \int_0^{+\infty} \mathbb{I}_{\{W_j^{y,1}(s)=i\}} \mathbb{I}_{\{u \leq \mu_j Y_j(s)\}} \widehat{N}_j^{y,1}(ds du) \\
 & + \int_0^t \int_0^{+\infty} \mathbb{I}_{\{W_j^{y,2}(s)=i\}} \mathbb{I}_{\{u \leq \alpha_j X_j(s) Y_j(s)\}} \widehat{N}_j^{x,2}(ds du) \\
 & - \int_0^t \int_0^{+\infty} (\mathbb{I}_{\{u \leq \beta_i Y_i(s)\}} \widehat{N}_i^{y,2}(ds du) + \mathbb{I}_{\{u \leq \mu_i Y_i(s)\}} \widehat{N}_i^{y,1}(ds du)),
 \end{aligned} \tag{4.2}$$

where for $1 \leq i \leq K$ and $1 \leq m \leq M$, $\widehat{N}_i^{x,1}$, $\widehat{N}_i^{x,2}$, $\widehat{N}_i^{y,1}$ and $\widehat{N}_i^{y,2}$ are independent Poisson point processes with intensity 1 on $(\mathbb{R}^+)^2$ and $X_i(t)$ (resp. $Y_i(t)$) is the number of susceptible (resp. infected) individuals in reactor i at time t .

As the dynamics above are not tractable, we are interested in studying them under the Poisson Hypothesis, which consists in assuming that the reactors are independent and that arrival processes are Poisson processes. We pursue two topics in this work: show that there exists a physical system constructed from the network SIS-DOCS model that, when properly scaled, converges to the Poisson Hypothesis regime; and study the relationship between the structure of the network and the epidemics propagation under the Poisson Hypothesis.

For the first topic, we introduce the replica-mean-field version of the network SIS-DOCS model. This physical system consists of randomly interacting copies, or replicas, of the network dynamics of interest. It has been shown in previous work that for a large class of dynamics in both discrete and continuous time, at the infinite-replica limit, the Poisson Hypothesis arises. We conjecture that the result still holds for this model and discuss why the previously known results do not directly apply to it.

For the second topic, we derive under the Poisson Hypothesis a system of nonlinear equations verified by the mean numbers of infected and susceptible individuals at each reactor.

4.2.2 The replica-mean-field limit of the SIS-DOCS network

The replica-mean-field dynamics is defined as follows: given an instance of network SIS-DOCS dynamics, take M copies, or replicas, of it, coupling them on the initial conditions. When a departure occurs from a reactor in one of the replicas, a replica index is sampled uniformly and independently and a neighboring reactor is uniformly chosen in that replica.

The dynamics can be formally be described by the following equations:
for $t \geq 0, 1 \leq i \leq K$ and $1 \leq m \leq M$,

$$\begin{aligned}
 X_{m,i}(t) &= \sum_{j \neq i} \sum_{n \neq m} \int_0^t \int_0^{+\infty} \mathbb{I}_{\{W_j^{x,1}(s)=i\}} \mathbb{I}_{V_{(n,j)}^{x,1}(s)=m} \mathbb{I}_{\{u \leq \mu_j X_{n,j}(s)\}} \widehat{N}_{n,j}^{x,1}(ds du) \\
 &+ X_{m,i}(0) + \int_0^t \int_0^{+\infty} \mathbb{I}_{\{W_j^{x,2}(s)=i\}} \mathbb{I}_{V_{(n,j)}^{x,2}(s)=m} \mathbb{I}_{\{u \leq \beta_j Y_{n,j}(s)\}} \widehat{N}_{n,j}^{y,2}(ds du) \\
 &- \int_0^t \int_0^{+\infty} (\mathbb{I}_{\{u \leq \alpha_i X_{m,i}(s) Y_{m,i}(s)\}} \widehat{N}_{m,i}^{x,2}(ds du) + \mathbb{I}_{\{u \leq \mu_i X_{m,i}(s)\}} \widehat{N}_{m,i}^{x,1}(ds du))
 \end{aligned} \tag{4.3}$$

and

$$\begin{aligned}
 Y_{m,i}(t) &= \sum_{j \neq i} \sum_{n \neq m} \int_0^t \int_0^{+\infty} \mathbb{I}_{\{W_j^{y,1}(s)=i\}} \mathbb{I}_{V_{(n,j)}^{y,1}(s)=m} \mathbb{I}_{\{u \leq \mu_j Y_{n,j}(s)\}} \widehat{N}_{n,j}^{y,1}(ds du) \\
 &+ Y_{m,i}(0) + \int_0^t \int_0^{+\infty} \mathbb{I}_{\{W_j^{y,2}(s)=i\}} \mathbb{I}_{V_{(n,j)}^{y,2}(s)=m} \mathbb{I}_{\{u \leq \alpha_j X_{n,j}(s) Y_{n,j}(s)\}} \widehat{N}_{n,j}^{x,2}(ds du) \\
 &- \int_0^t \int_0^{+\infty} (\mathbb{I}_{\{u \leq \beta_i Y_{m,i}(s)\}} \widehat{N}_{m,i}^{y,1}(ds du) + \mathbb{I}_{\{u \leq \mu_i Y_{m,i}(s)\}} \widehat{N}_{m,i}^{y,2}(ds du)),
 \end{aligned} \tag{4.4}$$

where for $1 \leq i \leq K$ and $1 \leq m \leq M$, $\widehat{N}_{m,i}$, $\widehat{N}_{m,i}^1$ and $\widehat{N}_{m,i}^2$ are independent Poisson point processes with intensity 1 on $(\mathbb{R}^+)^2$.

We now define the dynamics under the Poisson Hypothesis. We denote all state processes with tildes for clarity.

For $t \geq 0$ and $1 \leq i \leq K$,

$$\begin{aligned}
 \widetilde{X}_i(t) &= \widetilde{X}_i(0) + \sum_{j \neq i} \int_0^t \int_0^{+\infty} \mathbb{I}_{\{W_j^{x,1}(s)=i\}} \mathbb{I}_{\{u \leq \mu_j \mathbf{E}[\widetilde{X}_j(s)]\}} \widehat{N}_j^{x,1}(ds du) \\
 &+ \int_0^t \int_0^{+\infty} \mathbb{I}_{\{W_j^{x,2}(s)=i\}} \mathbb{I}_{\{u \leq \beta_j \mathbf{E}[\widetilde{Y}_j(s)]\}} \widehat{N}_{i,j}^{x,2}(ds du) \\
 &- \int_0^t \int_0^{+\infty} (\mathbb{I}_{\{u \leq \alpha_i \widetilde{X}_i(s) \widetilde{Y}_i(s)\}} \widehat{N}_i^{x,1}(ds du) + \mathbb{I}_{\{u \leq \mu_i \widetilde{X}_i(s)\}} \widehat{N}_i^{x,2}(ds du))
 \end{aligned} \tag{4.5}$$

and

$$\begin{aligned}
 \widetilde{Y}_i(t) &= \widetilde{Y}_i(0) + \sum_{j \neq i} \int_0^t \int_0^{+\infty} \mathbb{I}_{\{W_j^{y,1}(s)=i\}} \mathbb{I}_{\{u \leq \mu_j \mathbf{E}[\widetilde{Y}_j(s)]\}} \widehat{N}_{i,j}^1(ds du) \\
 &+ \int_0^t \int_0^{+\infty} \mathbb{I}_{\{W_j^{y,2}(s)=i\}} \mathbb{I}_{\{u \leq \alpha_j \mathbf{E}[\widetilde{X}_j(s) \widetilde{Y}_j(s)]\}} \widehat{N}_{i,j}^2(ds du) \\
 &- \int_0^t \int_0^{+\infty} (\mathbb{I}_{\{u \leq \beta_i \widetilde{Y}_i(s)\}} \widehat{N}_i^{y,1}(ds du) + \mathbb{I}_{\{u \leq \mu_i \widetilde{Y}_i(s)\}} \widehat{N}_i^{y,2}(ds du)),
 \end{aligned} \tag{4.6}$$

where for $1 \leq i \leq K$ and $1 \leq m \leq M$, \widehat{N}_i , \widehat{N}_i^1 and \widehat{N}_i^2 are independent Poisson point processes with intensity 1 on $(\mathbb{R}^+)^2$.

We require the following condition on the initial conditions, in a similar fashion to [27]:

Assumption 4.2.1. *There exists $\xi_0 > 0$ such that for all $\xi \leq \xi_0$, for all $1 \leq m \leq M$, $1 \leq i \leq K$, $\mathbf{E}[e^{\xi X_{m,i}(0)}] < \infty$ and $\mathbf{E}[e^{\xi Y_{m,i}(0)}] > \infty$.*

Under this assumption, we make the following conjecture:

Conjecture 4.2.1. *1. for all $(m, i) \in \{1, \dots, M\} \times \{1, \dots, K\}$, the random variables $X_{m,i}(t)$ and $Y_{m,i}(t)$ defined by (4.3) and (4.4) respectively converge in total variation to $\widetilde{X}_i(t)$ and $\widetilde{Y}_i(t)$ defined in (4.5) and (4.6) when $M \rightarrow \infty$;*

2. let \mathcal{N} be a finite subset of \mathbb{N}^ , for all $i \in \{1, \dots, K\}$, the processes $(X_{m,i}(\cdot))_{m \in \mathcal{N}}$ and $(Y_{m,i}(\cdot))_{m \in \mathcal{N}}$ weakly converge in the Skorokhod space of càdlàg functions on $[0, T]$ to $\text{card}(\mathcal{N})$ independent copies of the corresponding limit processes $(\widetilde{X}_i(\cdot))$ and $(\widetilde{Y}_i(\cdot))$ when $M \rightarrow \infty$;*

3. the previous convergence remains true in the Skorokhod space of càdlàg functions on \mathbb{R}^+ .

First, note that although the SIS-DOCS network dynamics do not directly belong to the class of continuous-time fragmentation-interaction-aggregation processes introduced in [28] for which the conjecture holds, the differences seem minor enough for the result to be true in some capacity. The three main differences are as follows: first, note that there are two types of arrivals to each node in the network, but as the arrivals of susceptible and infected individuals are independent, this is not a significant issue. Second, note that there are nonlinear terms in the state equations (4.3) and (4.4) of the form $\alpha_i X_i(t) Y_i(t)$. However, by considering a fixed threshold and separating the integrals in which the term appears in parts above and below the threshold, it is possible to control this term in fashion similar to [28]. The third, and most significant issue, is that the random routings W_i break the conditional independence property which is Lemma 3.3.13 from [28], on which the proof of propagation of chaos crucially relies. This explains the fact that the convergence result is currently an open question.

4.2.3 The relationship between network geometry and epidemics propagation under the Poisson Hypothesis

In the open SIS-DOCS reactor presented in [6], it was possible to obtain a closed form for the mean numbers of susceptible and infected customers at stationarity.

This relied on the interpretation of the infected process as an $M/M/\infty$ queue and on rate conservation principles. Our goal here is to extend this result to the case of a heterogenous network by deriving a system of equations for the mean numbers of susceptible and infected customers under the Poisson Hypothesis at stationarity.

For $1 \leq i \leq K$ let Φ_i be the stationary generating function of the i th SIS-DOCS reactor: for all $0 \leq x, y \leq 1$, $\Phi_i(x, y) = \mathbf{E}[x^{\tilde{X}_i} y^{\tilde{Y}_i}]$.

Then for $0 \leq x, y \leq 1$, by classical arguments, Φ_i satisfies the PDE

$$\begin{aligned} & \left(\sum_{j \neq i} \mu_j p_{j \rightarrow i} \mathbf{E}[\tilde{X}_j] + \sum_{j \neq i} \beta_j p_{j \rightarrow i} \mathbf{E}[\tilde{Y}_j] \right) (1-x) \\ & + \left(\sum_{j \neq i} \mu_j p_{j \rightarrow i} \mathbf{E}[\tilde{Y}_j] + \sum_{j \neq i} \alpha_j p_{j \rightarrow i} \mathbf{E}[\tilde{X}_j \tilde{Y}_j] \right) (1-y) \Phi_i(x, y) \\ & = \mu_i (1-x) \Phi_{i,x}(x, y) + \alpha_i y (1-x) \Phi_{i,xy}(x, y) + (\mu_i + \beta_i) (1-y) \Phi_{i,y}(x, y), \end{aligned} \quad (4.7)$$

where for all $1 \leq i \leq K$, $V(j) = \text{card}(\mathcal{V}(j))$, where $\mathcal{V}(j) = \{l \neq j, p_{l,j} \neq 0\}$.

Differentiating (4.7) w.r.t. x and then setting $x = 1$, we obtain

$$\begin{aligned} & - \left(\sum_{j \neq i} \mu_j p_{j \rightarrow i} \mathbf{E}[\tilde{X}_j] + \sum_{j \neq i} \beta_j p_{j \rightarrow i} \mathbf{E}[\tilde{Y}_j] \right) \Phi_i(1, y) \\ & + \left(\sum_{j \neq i} \mu_j p_{j \rightarrow i} \mathbf{E}[\tilde{Y}_j] + \sum_{j \neq i} \alpha_j p_{j \rightarrow i} \mathbf{E}[\tilde{X}_j \tilde{Y}_j] \right) (1-y) \Phi_{i,x}(1, y) \\ & = -\mu_i \Phi_{i,x}(1, y) - \alpha_i y \Phi_{i,xy}(1, y) + (\mu_i + \beta_i) (1-y) \Phi_{i,xy}(1, y). \end{aligned}$$

Let $\Psi_i(y) = \Phi_{i,x}(1, y)$. Then we can rewrite the previous equation as an ODE verified by Ψ_i :

$$\begin{aligned} & ((\mu_i + \beta_i + \alpha_i)(1-y) - \alpha_i) \Psi_i'(y) = - \left(\sum_{j \neq i} \mu_j p_{j \rightarrow i} \mathbf{E}[\tilde{X}_j] + \sum_{j \neq i} \beta_j p_{j \rightarrow i} \mathbf{E}[\tilde{Y}_j] \right) \Phi_i(1, y) \\ & + \left(\mu_i + \left(\sum_{j \neq i} \mu_j p_{j \rightarrow i} \mathbf{E}[\tilde{Y}_j] + \sum_{j \neq i} \alpha_j p_{j \rightarrow i} \mathbf{E}[\tilde{X}_j \tilde{Y}_j] \right) (1-y) \right) \Psi_i(y). \end{aligned} \quad (4.8)$$

Since the \tilde{Y}_i process can be seen an autonomous $M/M/\infty$ queue with input rate $\sum_{j \neq i} \mu_j p_{j \rightarrow i} \mathbf{E}[\tilde{Y}_j] + \sum_{j \neq i} \alpha_j p_{j \rightarrow i} \mathbf{E}[\tilde{X}_j \tilde{Y}_j]$ and service rate $\mu_i + \beta_i$, we have that

$$\Phi_i(1, y) = e^{-\theta_i(1-y)},$$

where

$$\theta_i = \frac{\sum_{j \neq i} \mu_j p_{j \rightarrow i} \mathbf{E}[\tilde{Y}_j] + \sum_{j \neq i} \alpha_j p_{j \rightarrow i} \mathbf{E}[\tilde{X}_j \tilde{Y}_j]}{\mu_i + \beta_i}.$$

Therefore, we can rewrite (4.8) as

$$\Psi'_i(y) = h_i(y)\Psi_i(y) - g_i(y), \quad (4.9)$$

where

$$h_i(y) = \frac{(\mu_i + (\sum_{j \neq i} \mu_j p_{j \rightarrow i} \mathbf{E}[\tilde{Y}_j] + \sum_{j \neq i} \alpha_j p_{j \rightarrow i} \mathbf{E}[\tilde{X}_j \tilde{Y}_j]))(1 - y)}{(\mu_i + \beta_i + \alpha_i)(1 - y) - \alpha_i}$$

and

$$g_i(y) = \frac{(\sum_{j \neq i} \mu_j p_{j \rightarrow i} \mathbf{E}[\tilde{X}_j] + \sum_{j \neq i} \beta_j p_{j \rightarrow i} \mathbf{E}[\tilde{Y}_j])e^{-\theta_i(1-y)}}{(\mu_i + \beta_i + \alpha_i)(1 - y) - \alpha_i}.$$

Note that this ODE admits a singularity at

$$y_i^* = \frac{\mu_i + \beta_i}{\mu_i + \beta_i + \alpha_i}.$$

However, for y in a neighborhood of 1 not including y_i^* , it can be checked that

$$\Psi_i(y) = e^{\int_1^y h_i(z) dz} (\Psi_i(1) - \int_1^y g_i(u) e^{-\int_1^u h_i(z) dz} du). \quad (4.10)$$

Using the same arguments as in [6], it can be shown that since for u in a right neighborhood of y_i^* , $e^{\int_1^y h_i(z) dz} \sim K(u - y_i^*)^b$ with K a constant and an explicit value of b , the integral

$$\int_1^{y_i^*} g_i(u) e^{-\int_1^u h_i(z) dz} du$$

is well-defined and finite, and must therefore be equal to $\Psi_i(1) = \mathbf{E}[\tilde{X}_i]$.

Using classical arguments and a change of variables, in a fashion similar to [6], we obtain

$$\mathbf{E}[\tilde{X}_i] = \frac{(\sum_{j \neq i} \mu_j p_{j \rightarrow i} \mathbf{E}[\tilde{X}_j] + \sum_{j \neq i} \beta_j p_{j \rightarrow i} \mathbf{E}[\tilde{Y}_j])}{\mu_i + \beta_i + \alpha_i} \int_0^1 e^{-\xi_i^1(t)} t^{\xi_i^2(t)-1} dt, \quad (4.11)$$

where

$$\xi_i^1(t) = \frac{(\sum_{j \neq i} \mu_j p_{j \rightarrow i} \mathbf{E}[\tilde{Y}_j] + \sum_{j \neq i} \alpha_j p_{j \rightarrow i} \mathbf{E}[\tilde{X}_j \tilde{Y}_j])\alpha_i^2(1 - t)}{(\mu_i + \beta_i)(\mu_i + \beta_i + \alpha_i)^2}$$

and

$$\xi_i^2(t) = \frac{\mu_i}{\mu_i + \beta_i + \alpha_i} + \frac{(\sum_{j \neq i} \mu_j p_{j \rightarrow i} \mathbf{E}[\tilde{Y}_j] + \sum_{j \neq i} \alpha_j p_{j \rightarrow i} \mathbf{E}[\tilde{X}_j \tilde{Y}_j])\alpha_i}{(\mu_i + \beta_i + \alpha_i)^2}.$$

Now, note the following result, which is a first-order rate conservation principle:

Lemma 4.2.2. For all $1 \leq i \leq K$,

$$\sum_{j \neq i} \mu_j p_{j \rightarrow i} \mathbf{E}[\tilde{X}_j] + \sum_{j \neq i} \beta_j p_{j \rightarrow i} \mathbf{E}[\tilde{Y}_j] = \mu_i \mathbf{E}[\tilde{X}_i] + \alpha_i \mathbf{E}[\tilde{X}_i \tilde{Y}_i] \quad (4.12)$$

and

$$\sum_{j \neq i} \mu_j p_{j \rightarrow i} \mathbf{E}[\tilde{Y}_j] + \sum_{j \neq i} \alpha_j p_{j \rightarrow i} \mathbf{E}[\tilde{X}_j \tilde{Y}_j] = (\mu_i + \beta_i) \mathbf{E}[\tilde{Y}_i]. \quad (4.13)$$

Proof. Derived from the PDE (4.7), by differentiating it w.r.t. y and taking $x = y = 1$. \square

Using the rate conservation principle, we can rewrite (4.11) as

$$\mathbf{E}[\tilde{X}_i] = \frac{(\sum_{j \neq i} \mu_j p_{j \rightarrow i} \mathbf{E}[\tilde{X}_j] + \sum_{j \neq i} \beta_j p_{j \rightarrow i} \mathbf{E}[\tilde{Y}_j])}{\mu_i + \beta_i + \alpha_i} \int_0^1 e^{-\frac{(\mu_i + \beta_i) \mathbf{E}[\tilde{Y}_i] \alpha_i^2 (1-t)}{(\mu_i + \beta_i)(\mu_i + \beta_i + \alpha_i)^2} t} \left(\frac{\mu_i}{\mu_i + \beta_i + \alpha_i} + \frac{(\mu_i + \beta_i) \mathbf{E}[\tilde{Y}_i] \alpha_i}{(\mu_i + \beta_i + \alpha_i)^2} - 1 \right) dt. \quad (4.14)$$

Finally, combining (4.12) and (4.13), we get

$$\begin{aligned} \mathbf{E}[\tilde{Y}_i] &= \frac{\sum_{j \neq i} p_{j \rightarrow i} (\sum_{l \neq j} \mu_l p_{l \rightarrow j} \mathbf{E}[\tilde{X}_l] + \sum_{l \neq j} \beta_l p_{l \rightarrow j} \mathbf{E}[\tilde{Y}_l] - \mu_j \mathbf{E}[\tilde{X}_j])}{\mu_i + \beta_i} \\ &+ \frac{\sum_{j \neq i} \mu_j p_{j \rightarrow i} \mathbf{E}[\tilde{Y}_j]}{\mu_i + \beta_i}. \end{aligned} \quad (4.15)$$

Moreover, let $N = \sum_{i=1}^K X_{1,i} + Y_{1,i}$ be the total population in a given replica in the M -replica model. Then, by exchangeability,

$$\frac{1}{M} \sum_{m=1}^M \sum_{i=1}^K X_{m,i} + Y_{m,i} = N.$$

This gives, letting M go to infinity and assuming the conjecture 4.2.1 holds,

$$\sum_{i=1}^K \mathbf{E}[\tilde{X}_i] + \mathbf{E}[\tilde{Y}_i] = N. \quad (4.16)$$

The equations (4.14) (4.15) and (4.16) can be regrouped as a system of nonlinear equations verified by $\mathbf{E}[\tilde{X}_i]$ and $\mathbf{E}[\tilde{Y}_i]$:

$$\left\{ \begin{array}{l}
 \mathbf{E}[\tilde{X}_i] = \frac{(\sum_{j \neq i} \mu_j p_{j \rightarrow i} \mathbf{E}[\tilde{X}_j] + \sum_{j \neq i} \beta_j p_{j \rightarrow i} \mathbf{E}[\tilde{Y}_j])}{\mu_i + \beta_i + \alpha_i} \\
 \int_0^1 e^{-\frac{(\mu_i + \beta_i) \mathbf{E}[\tilde{Y}_i] \alpha_i^2 (1-t)}{(\mu_i + \beta_i)(\mu_i + \beta_i + \alpha_i)^2} t} \left(\frac{\mu_i}{\mu_i + \beta_i + \alpha_i} + \frac{(\mu_i + \beta_i) \mathbf{E}[\tilde{Y}_i] \alpha_i}{(\mu_i + \beta_i + \alpha_i)^2} - 1 \right) dt \\
 \mathbf{E}[\tilde{Y}_i] = \frac{\sum_{j \neq i} \mu_j p_{j \rightarrow i} \mathbf{E}[\tilde{Y}_j] + \sum_{j \neq i} p_{j \rightarrow i} (\sum_{l \neq j} \mu_l p_{l \rightarrow j} \mathbf{E}[\tilde{X}_l] + \sum_{l \neq j} \beta_l p_{l \rightarrow j} \mathbf{E}[\tilde{Y}_l] - \mu_j \mathbf{E}[\tilde{X}_j])}{\mu_i + \beta_i} \\
 \sum_{i=1}^K \mathbf{E}[\tilde{X}_i] + \mathbf{E}[\tilde{Y}_i] = N.
 \end{array} \right. \tag{4.17}$$

Current work in progress involves studying this system numerically.

Bibliography

- [1] Z. Agathe-Nerine. Multivariate hawkes processes on inhomogeneous random graphs. *Stochastic Processes and their Applications*, 152:86–148, 2022.
- [2] D. Aldous. Stopping Times and Tightness. *The Annals of Probability*, 6(2):335 – 340, 1978.
- [3] S. Allmeier and N. Gast. Mean field and refined mean field approximations for heterogeneous systems: It works! *Association for Computing Machinery*, 50(1):103–104, jun 2022.
- [4] F. Amblard and G. Deffuant. The role of network topology on extremism propagation with the relative agreement opinion dynamics. *Physica A: Statistical Mechanics and its Applications*, 343:725–738, 11 2004.
- [5] F. Baccelli, M. Davydov, and T. Taillefumier. Replica-mean-field limits of fragmentation-interaction-aggregation processes. *Journal of Applied Probability*, 59:1–22, 01 2022.
- [6] F. Baccelli, S. Foss, and V. Shneer. Migration-contagion processes. *Advances in Applied Probability*, Mar. 2023.
- [7] F. Baccelli, D. R. McDonald, and J. Reynier. A Mean-Field Model for Multiple TCP Connections through a Buffer Implementing RED. *Performance Evaluation*, pages 77–97, 2002.
- [8] F. Baccelli and T. Taillefumier. Replica-mean-field limits for intensity-based neural networks. *SIAM Journal on Applied Dynamical Systems*, 18(4):1756–1797, 1 2019.
- [9] F. Baccelli and T. Taillefumier. The pair-replica-mean-field limit for intensity-based neural networks. <https://arxiv.org/abs/2004.06246>, 2020.
- [10] A. Barbour, L. Holst, and S. Janson. *Poisson approximation*. Oxford Studies in Probability. Oxford University Press, 1992.

- [11] M. Benaïm and J.-Y. Le Boudec. A class of mean field interaction models for computer and communication systems. *Perform Evalalution*, 65:589–590, 05 2008.
- [12] P. Billingsley. *Convergence of Probability Measures*. Wiley Series in Probability and Mathematical Statistics. Wiley, 1968.
- [13] B. Blaszczyzyn. Lecture Notes on Random Geometric Models — Random Graphs, Point Processes and Stochastic Geometry. Lecture, 12 2017.
- [14] V. I. Bogachev and A. V. Kolesnikov. The Monge-Kantorovich problem: achievements, connections, and perspectives. *Russian Mathematical Surveys*, 67(5):785–890, 10 2012.
- [15] M. Boguñá and R. Pastor-Satorras. Epidemic spreading in correlated complex networks. *Phys. Rev. E*, 66:047104, Oct 2002.
- [16] K. A. Borovkov. Propagation of chaos for queueing networks. *Theory of Probability & Its Applications*, 42(3):385–394, 1998.
- [17] P. Brémaud. *Point Process Calculus in Time and Space: An Introduction with Applications*. Probability Theory and Stochastic Modelling. Springer International Publishing, 2020.
- [18] P. Brémaud and L. Massoulié. Stability of nonlinear hawkes processes. *Annals of Probability*, 24:1563–1588, 1996.
- [19] T. Castellani and A. Cavagna. Spin-glass theory for pedestrians. *Journal of Statistical Mechanics: Theory and Experiment*, 2005(05):5–12, 5 2005.
- [20] P. Cattiaux, L. Colombani, and M. Costa. Limit theorems for hawkes processes including inhibition. *Stochastic Processes and their Applications*, 149:404–426, 2022.
- [21] B. Cessac. A discrete time neural network model with spiking neurons. *Journal of Mathematical Biology*, 56(3):311–345, 9 2007.
- [22] D. Chakrabarti, Y. Wang, C. Wang, J. Leskovec, and C. Faloutsos. Epidemic thresholds in real networks. *ACM Trans. Inf. Syst. Secur.*, 10(4), jan 2008.
- [23] L. H. Y. Chen. Poisson Approximation for Dependent Trials. *The Annals of Probability*, 3(3):534 – 545, 1975.
- [24] L. H. Y. CHEN and A. XIA. Poisson process approximation for dependent superposition of point processes. *Bernoulli*, 17(2):530–544, 2011.

- [25] J. Chevallier. Approximation par champ-moyen : le couplage à la Sznitman pour les nuls. Ces notes se veulent être abordables par un large public et être un complément à l'exposé " Méthode d'approximation de champ-moyen : couplage à la Sznitman sur un exemple jouet. " présenté le 27 Janvier 2017 dans le cadre de la 5ème journée Young Statisticians and Probabilists., Jan. 2017.
- [26] C. F. Coletti and I. Papageorgiou. Asymptotic analysis of the elephant random walk. *Journal of Statistical Mechanics: Theory and Experiment*, 2021(1):013205, jan 2021.
- [27] M. Davydov. Propagation of chaos and poisson hypothesis for replica mean-field models of intensity-based neural networks. 2211.11490, 2022.
- [28] M. Davydov. Replica-mean-field limit of continuous-time fragmentation-interaction-aggregation processes, 2023.
- [29] S. Delattre, N. Fournier, and M. Hoffmann. Hawkes processes on large networks. *Annals of Applied Probability*, 2016.
- [30] S. Ditlevsen and E. Löcherbach. Multi-class oscillating systems of interacting neurons. *Stochastic Processes and their Applications*, 127(6):1840–1869, 2017.
- [31] R. L. Dobrushin. Vlasov equations. *Functional Analysis and Its Applications*, 13(2):115–123, Apr 1979.
- [32] X. Erny, E. Löcherbach, and D. Loukianova. Conditional propagation of chaos for mean field systems of interacting neurons. *Electronic Journal of Probability*, 26:1 – 25, 2021.
- [33] R. Fitzhugh. Impulses and Physiological States in Theoretical Models of Nerve Membrane. *Biophysical Journal*, 1(6):445–466, July 1961.
- [34] N. Fournier and E. Löcherbach. On a toy model of interacting neurons. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 52(4):1844 – 1876, 2016.
- [35] A. Galves and E. Löcherbach. Infinite systems of interacting chains with memory of variable length—a stochastic model for biological neural nets. *Journal of Statistical Physics*, 151(5):896–921, 3 2013.
- [36] N. Gast, D. Latella, and M. Massink. A refined mean field approximation of synchronous discrete-time population models. *Performance Evaluation*, 126:1–21, 10 2018.

- [37] W. Gerstner and W. M. Kistler. *Spiking Neuron Models: Single Neurons, Populations, Plasticity*. Cambridge University Press, 2002.
- [38] Grazieschi, Paolo, Leocata, Marta, Mascart, Cyrille, Chevallier, Julien, Delarue, François, and Tanré, Etienne. Network of interacting neurons with random synaptic weights. *ESAIM: ProcS*, 65:445–475, 2019.
- [39] B. Grigelionis. On the convergence of sums of random step processes to a poisson process. *Theory of Probability & Its Applications*, 8(2):177–182, 1963.
- [40] A. Gupta, C. Briat, and M. Khammash. A scalable computational framework for establishing long-term behavior of stochastic reaction networks. *PLoS Computational Biology*, 10(6):e1003669, 6 2014.
- [41] A. L. Hodgkin and A. F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology*, 117(4):500–544, 1952.
- [42] E. M. Izhikevich. *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting*. The MIT Press, 07 2006.
- [43] P.-E. Jabin, D. Poyato, and J. Soler. Mean-field limit of non-exchangeable systems. 2112.15406, 2022.
- [44] M. Kac. *FOUNDATIONS OF KINETIC THEORY*, pages 173–200. University of California Press, Berkeley, 1956.
- [45] O. Kallenberg. *Random Measures, Theory and Applications*, volume 77. Springer, 01 2017.
- [46] L. Kantorovich. On the translocation of masses. *Journal of Mathematical Sciences*, 133, 03 2006.
- [47] R. Karim, R. Laeven, and M. Mandjes. Compound multivariate hawkes processes: Large deviations and rare event simulation, 2022.
- [48] F. Karpelevich and A. Rybko. Asymptotic behavior of the thermodynamical limit for symmetric closed queueing networks. *Problems of Information Transmission*, 36, 01 2000.
- [49] L. Kleinrock. *Queueing Systems*, volume I: Theory. Wiley Interscience, 1975.
- [50] L. Kleinrock. *Queueing Systems*, volume II: Computer Applications. Wiley Interscience, 1977.

- [51] L. Kleinrock. *Communication nets: Stochastic message flow and delay*. Courier Corporation, 2007.
- [52] C. Kuehn. Network dynamics on graphops. *New Journal of Physics*, 22(5):053030, may 2020.
- [53] T. G. Kurtz. Limit theorems for sequences of jump markov processes approximating ordinary differential processes. *Journal of Applied Probability*, 8(2):344–356, 1971.
- [54] T. G. Kurtz. *Limit theorems and diffusion approximations for density dependent Markov chains*, pages 67–78. Springer Berlin Heidelberg, Berlin, Heidelberg, 1976.
- [55] D. Lacker, K. Ramanan, and R. Wu. Local weak convergence for sparse networks of interacting processes. *Annals of Applied Probability*, to appear, 2020.
- [56] J.-Y. Le Boudec, D. McDonald, and J. Munding. A generic mean field convergence result for systems of interacting objects. In *Proceedings of the Fourth International Conference on Quantitative Evaluation of Systems, QEST '07*, page 3–18, USA, 2007. IEEE Computer Society.
- [57] L. Le Cam. *Asymptotic Methods in Statistical Decision Theory*. Springer Series in Statistics. Springer, 1986.
- [58] T. Lindvall. *Lectures on the Coupling Method*. Dover Books on Mathematics Series. Dover Publications, Incorporated, 2002.
- [59] A. Masi, A. Galves, E. Löcherbach, and E. Presutti. Hydrodynamic limit for interacting neurons. *Journal of Statistical Physics*, 158:866–902, 2015.
- [60] H. P. McKean. A class of markov processes associated with nonlinear parabolic equations. *Proceedings of the National Academy of Sciences*, 56(6):1907–1911, 1966.
- [61] P. V. Mieghem, J. Omic, and R. E. Kooij. Virus spread in networks. *IEEE/ACM Transactions on Networking*, 17:1–14, 2009.
- [62] C. Morris and H. Lecar. Voltage oscillations in the barnacle giant muscle fiber. *Biophysical Journal*, 35(1):193–213, 1981.
- [63] J. Nagumo, S. Arimoto, and S. Yoshizawa. An active pulse transmission line simulating nerve axon. *Proceedings of the IRE*, 50(10):2061–2070, 1962.

- [64] K. Pakdaman, M. Thieullen, and G. Wainrib. Fluid limit theorems for stochastic hybrid systems with application to neuron models. *Advances in Applied Probability*, 42(3):761–794, 2010.
- [65] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani. Epidemic processes in complex networks. *Reviews of Modern Physics*, 87(3):925–979, 8 2015.
- [66] R. Pemantle and A. Stacey. The branching random walk and contact process on galton-watson and nonhomogeneous trees. *The Annals of Probability*, 29(4):1563–1590, 2001.
- [67] T. C. Phi, E. Löcherbach, and P. Reynaud-Bouret. Kalikow decomposition for counting processes with stochastic intensity and application to simulation algorithms. *Journal of Applied Probability*, page 1–32, 2023.
- [68] J. W. Pillow, J. Shlens, L. Paninski, A. Sher, A. M. Litke, E. J. Chichilnisky, and E. P. Simoncelli. Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature*, 454(7207):995–999, Aug 2008.
- [69] J. Rissanen. A universal data compression system. *IEEE Trans. Inf. Theory*, 29:656–663, 1983.
- [70] P. Robert and J. Touboul. On the dynamics of random neuronal networks. *Journal of Statistical Physics*, 165(3):545–584, Nov 2016.
- [71] A. Sankararaman, F. Baccelli, and S. Foss. Interference queueing networks on grids. *The Annals of Applied Probability*, 2017.
- [72] Y. Seol. Limit theorems for discrete hawkes processes. *Statistics & Probability Letters*, 99:223–229, 2015.
- [73] O. Shriki, J. Alstott, F. Carver, T. Holroyd, R. N. Henson, M. L. Smith, R. Coppola, E. Bullmore, and D. Plenz. Neuronal avalanches in the resting meg of the human brain. *Journal of Neuroscience*, 33(16):7079–7090, 2013.
- [74] S. M. Stigler. Stigler’s law of eponymy*. *Transactions of the New York Academy of Sciences*, 39(1 Series II):147–157, 1980.
- [75] A.-S. Sznitman. Topics in propagation of chaos. *Ecole d’Ete de Probabilites de Saint-Flour XIX, vol.1464*, pages 165–251, 1989.
- [76] K. Teerapabolarn. Poisson approximation for random sums of poisson random variables. *IJPAM*, 95(4):543–546, 2014.

- [77] W. Truccolo, U. T. Eden, M. R. Fellows, J. P. Donoghue, and E. N. Brown. A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects. *Journal of Neurophysiology*, 93:1074–1089, 2005.
- [78] A. A. Vladimirov, S. A. Pirogov, A. N. Rybko, and S. B. Shlosman. Propagation of chaos and Poisson hypothesis. *Problems of Information Transmission*, 54(3):290–299, 7 2018.
- [79] N. Vvedenskaya, R. Dobrushin, and F. Karpelevich. Queueing system with selection of the shortest of two queues: An asymptotic approach. *Probl. Peredachi Inf.*, 32:20–34, 1996.

RÉSUMÉ

Dans cette thèse, nous nous intéressons à des modèles mathématiques de phénomènes pouvant être interprétés comme des dynamiques sur des réseaux. Cela inclut par exemple des modèles de populations de neurones interagissant à des instants aléatoires avec leurs voisins ou de propagation d'épidémies où les individus infectés ou susceptibles de l'être se déplacent de ville en ville. La description mathématique de tels phénomènes repose en général sur un compromis entre niveau de détail recherché et tractabilité mathématique. La majeure partie des travaux de cette thèse concerne l'élaboration de preuves mathématiques pour justifier l'introduction de modèles permettant de prendre en compte de la géométrie du réseau sous-jacent dans ces phénomènes tout en restant tractables. L'outil mathématique central pour cela est le champ moyen à répliques, qui consiste en des copies du réseau étudié entre lesquelles les interactions sont mélangées aléatoirement. Les résultats principaux de cette thèse concernent le comportement d'un tel système dynamique lorsque le nombre de répliques tend vers l'infini. Dans de multiples cadres, nous montrons qu'il converge vers une dynamique sous hypothèse poissonnienne, c'est-à-dire dans laquelle les temps d'interactions sont remplacés par des processus de Poisson indépendants, ce qui permet d'effectuer des calculs explicites pour certains modèles. Le chapitre 2 de cette thèse est consacré à l'établissement de ce résultat pour une classe de processus en temps discret, les processus d'interaction-agrégation-fragmentation. A l'échelle d'un noeud du réseau, ces processus modélisent l'état du noeud en agréant à son évolution autonome les effets des interactions de ses voisins. Le chapitre 3 étend ces résultats au cas du temps continu, où les instants des interactions sont vus comme des réalisations de processus ponctuels, en mettant en exergue le cas du modèle de Galves-Löcherbach utilisé en neurosciences computationnelles. Enfin, le chapitre 4 s'intéresse à l'étude d'un modèle de propagation d'épidémie sous hypothèse poissonnienne: le processus de migration-contagion, qui consiste en un réseau fermé de files d'attente entre lesquelles migrent des individus infectés ou susceptibles de l'être. Plus précisément, nous établissons un système d'équations non-linéaires vérifié par les nombres moyens d'individus infectés et susceptibles dans le but de l'étudier ensuite numériquement.

MOTS CLÉS

Propagation du chaos, champ moyen à répliques, approximation poissonnienne, processus ponctuel, modèles d'intensité, méthode de Chen-Stein, processus de migration-contagion, modèles compartimentaux, champ moyen

ABSTRACT

In this thesis, we are interested in mathematical models of phenomena that can be interpreted as network dynamics. This includes for example neuron population models in which neurons interact at random times with their neighbors or epidemics propagation where infected or susceptible individuals move from town to town. The mathematical description of such phenomena generally requires a compromise between physical or biological relevance and mathematical tractability. The main focus of this work is the elaboration of mathematical proofs to justify the introduction of models taking into account the geometry of the underlying networks whilst preserving tractability. The main mathematical tool for that purpose is the replica-mean-field, which consists in copies of the studied network between which interactions are routed randomly. The main results of this thesis concern the behavior of such a dynamical system when the number of replicas goes to infinity. In various settings, we show that it converges to dynamics under the Poisson Hypothesis, that is, interaction times are replaced by independent Poisson processes, which allows to obtain closed forms in certain models. In chapter 2 of the thesis, we prove this result for a class of discrete-time dynamics: fragmentation-interaction-aggregation processes. At the scale of a given node, these processes model its state by an autonomous evolution to which are aggregated the effects of the interactions with its neighbors. Chapter 3 extends these results to the continuous-time framework, where interaction times are seen as realizations of point processes, highlighting the case of Galves-Löcherbach model used in computational neuroscience. Finally, chapter 4 focuses on the study of a model of epidemics propagation under the Poisson Hypothesis: the migration-contagion process, consisting in a closed network of queues in between which infected and susceptible customers migrate. More precisely, we establish a system of nonlinear equations verified by the mean numbers of infected and susceptible individuals in the objective of studying it numerically.

KEYWORDS

Propagation of chaos, replica-mean-field, Poisson approximation, point process, intensity-based models, Chen-Stein method, migration-contagion process, compartmental models, mean-field