



HAL
open science

Plasticity in networks of spiking neurons in interaction

Pascal Helson

► **To cite this version:**

Pascal Helson. Plasticity in networks of spiking neurons in interaction. Neuroscience. Université Côte d'Azur, 2021. English. NNT : 2021COAZ4013 . tel-03200979v2

HAL Id: tel-03200979

<https://hal.science/tel-03200979v2>

Submitted on 8 Jun 2021

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.

THÈSE DE DOCTORAT

Étude de la plasticité pour des neurones à décharge en interaction

Pascal HELSON

Tosca, INRIA Sophia Antipolis Méditerranée

**Présentée en vue de l'obtention
du grade de docteur en Mathématiques**
d'Université Côte d'Azur
Dirigée par : Etienne TANRÉ
Co-encadrée par : Romain VELTZ
Soutenue le : 29 mars 2021

Devant le jury, composé de :
Alain Destexhe, Directeur de recherche, CNRS - UNIC EITN
Éric Luçon, Maître de conférence, Université Paris Descartes
Patricia Reynaud-Bouret, Directrice de recherche, CNRS -
Université Côte d'Azur
Philippe Robert, Directeur de recherche, INRIA
Benoîte de Saporta, Professeure, Université de Montpellier
Denis Talay, Directeur de recherche, INRIA
Etienne Tanré, Chargé de recherche, INRIA
Romain Veltz, Chargé de recherche, INRIA

Étude de la plasticité pour des neurones à décharge en interaction

Plasticity in networks of spiking neurons in interaction

Pascal HELSON

INRIA Sophia Antipolis Méditerranée

Dirigée par : Etienne Tanré

Co-encadrée par : Romain Veltz

Rapporteur·se·s

M. ROBERT, Philippe

Directeur de recherche INRIA, Paris

Mme DE SAPORTA, Benoîte

Professeure Université de Montpellier

Examineur·trice·s

M. DESTEXHE , Alain

Directeur de recherche CNRS - UNIC EITN

M. LUÇON , Éric

Maître de conférence Université Paris Descartes

Mme REYNAUD-BOURET, Patricia

Directrice de recherche CNRS - Université Côte d'Azur, France

Invité

M. TALAY, Denis

Directeur de recherche INRIA Sophia Antipolis - Méditerranée

Remerciements

Tout d'abord, un grand merci à mes deux directeurs de thèse, Etienne Tanré et Romain Veltz, qui m'ont fait découvrir le monde des modèles mathématiques en neurosciences en tant qu'enseignants avant de m'accompagner sur un sujet de thèse passionnant cette fois en tant que directeurs. Vos innombrables conseils m'ont beaucoup appris et ce dans tous les domaines. Merci pour tout le temps que vous m'avez consacré. Merci pour votre patience et votre investissement. Je vous suis infiniment reconnaissant de m'avoir donné goût aux probabilités ainsi qu'à ses diverses applications en neurosciences. Vous formez un très bon binôme d'encadrant !

Je tiens à remercier chaleureusement les membres de mon jury. Tout d'abord, merci à tous pour vos questions qui ont ouvert la voie à de nombreux nouveaux défis toujours plus passionnant ! Patricia Reynaud-Bouret, Alain Destexhe et Éric Luçon, merci d'avoir accepté de faire partie de mon jury et d'avoir consacré de votre temps sur mon travail. Benoîte de Saporta et Philippe Robert, merci infiniment d'avoir rapporté ma thèse. Benoîte, merci pour votre lecture attentive et votre rapport élogieux. Philippe, merci pour votre lecture détaillée et vos conseils sur le chapitre 4 ! Denis, merci de m'avoir accepté au sein de ton équipe TOSCA. Merci pour toutes ces anecdotes que tu as partagé avec nous et ton soutien à toutes nos initiatives aussi diverses que variées.

Le centre d'Inria Sophia-Antipolis offre un cadre de travail sans égal. Je tiens à remercier tous les membres de l'équipe de foot d'Inria qui m'ont permis de courir après le ballon ! En particulier, merci à Florent et Dorian de m'avoir fait confiance en me laissant les rennes de l'équipe ces dernières années. Ensuite, merci à tous les membres du bâtiment Galois avec qui on a su faire face aux diverses pannes dans la joie et la bonne humeur. Merci plus particulièrement aux membres de l'équipe TOSCA. Tout d'abord, merci Christine, tu as été une formidable assistante. Ensuite, merci aux ex(-post)-doctorants avec qui j'ai partagé des moments extraordinaires, souvent organisés autour du café réalisé par la machine Alicia. Milica, Émilie, Kerlyns, Radu, Quentin, Lorenzo, Christophe, vous avez rendu ces années inoubliables et j'espère qu'on continuera à se voir le plus régulièrement possible, vous me manquez déjà !

Merci à tous mes amis qui m'ont soutenu et qui m'ont donné la force d'aller jusqu'au bout ! Tout d'abord, les verts : Maël, Valentin et Zizou, l'énergie que vous me donnez est à la limite du surnaturel. Ensuite, les vençois d'ici et d'ailleurs : Guillaume, Camille et Marc, quand une amitié dure si longtemps, c'est qu'elle n'aura pas de fin ! Brigitte, merci de m'avoir nourri et Cédric, merci de m'avoir fait découvrir le badminton de compétition qui nous a bien fait transpirer !

Merci à toute ma famille de Toulouse à Cap d'Ail en passant par Paris, vous êtes tous formidables et avez su trouver les mots pour me donner confiance en moi ! En particulier, merci à ma mamie et mon papi, je suis maintenant toute ouïe, soyez prêt à tout me raconter ! Merci à mes parents qui m'ont apporté un soutien continu, vous m'avez donné bien plus que ce que vous ne pensez, merci pour tout ! Merci à toi grand frère de m'avoir montré la voie, tous tes conseils restent gravés à jamais en moi. Merci et courage à toi petit frère, de belles années t'attendent, j'aurai aimé te consacrer plus de temps ces dernières années, sache que je suis très fier de toi !

Enfin, merci aux parents d'Elfie de m'avoir accepté très tôt comme un membre à part entière de votre famille. Et la meilleure pour la fin, merci infiniment Elfie. Après toutes ces années passées éloignés l'un de l'autre, on s'est retrouvé tout proche mais voilà que l'obstacle de la thèse s'est immiscé entre nous. Ta simplicité, toutes tes petites attentions, tes pleurs, ton sourire, ton impatience ... tout chez toi me plaît et ne cesse de me revigorer. Le dernier obstacle étant levé, j'ai hâte qu'on puisse enfin s'envoler ensemble !

Abstract

In this thesis, we study a phenomenon that may be responsible for our memory capacity: the synaptic plasticity. It modifies the links between neurons over time. This phenomenon is stochastic: it is the result of a series of diverse and numerous chemical processes. The aim of the thesis is to propose a model of plasticity for interacting spiking neurons. The main difficulty is to find a model that satisfies the following conditions: it must be both consistent with the biological results of the field and simple enough to be studied mathematically and simulated with a large number of neurons.

In a first step, from a rather simple model of plasticity, we study the learning of external signals by a neural network as well as the forgetting time of this signal when the network is subjected to other signals (noise). The mathematical analysis allows us to control the probability to misevaluate the signal. From this, we deduce explicit bounds on the time during which a given signal is kept in memory.

Next, we propose a model based on stochastic rules of plasticity as a function of the occurrence time of the neural electrical discharges (Spike Timing Dependent Plasticity, STDP). This model is described by a Piecewise Deterministic Markov Process (PDMP). The long time behaviour of such a neural network is studied using a slow-fast analysis. In particular, sufficient conditions are established under which the process associated with synaptic weights is ergodic.

Finally, we make the link between two levels of modelling: the microscopic and the macroscopic approaches. Starting from the dynamics presented at a microscopic level (neuron model and its interaction with other neurons), we derive an asymptotic dynamics which represents the evolution of a typical neuron and its incoming synaptic weights: this is the mean field analysis of the model. We thus condense the information on the dynamics of the weights and that of the neurons into a single equation, that of a *typical* neuron.

Keywords: Piecewise Deterministic Markov Process, Multi-scale, Mean Field, Spike Timing Dependent Plasticity, Recurrent Neural Network, Synaptic Plasticity, Memory Lifetime.

Résumé

Dans cette thèse nous étudions un phénomène susceptible d'être responsable de notre capacité de mémoire : la plasticité synaptique. C'est le changement des liens entre les neurones au cours du temps. Ce phénomène est stochastique : c'est le résultat d'une suite de divers et nombreux mécanismes chimiques. Le but de la thèse est de proposer un modèle de plasticité pour des neurones à décharge en interaction. La principale difficulté consiste à trouver un modèle qui satisfait les conditions suivantes : ce modèle doit être à la fois cohérent avec les résultats biologiques dans le domaine et assez simple pour être étudié mathématiquement et simulé avec un grand nombre de neurones.

Dans un premier temps, à partir d'un modèle assez simple de plasticité, on étudie l'apprentissage de signaux extérieurs par un réseau de neurones ainsi que le temps d'oubli de ce signal lorsque le réseau est soumis à d'autres signaux (bruit). L'analyse mathématique nous permet de contrôler la probabilité d'une mauvaise évaluation du signal. On en déduit un minorant du temps de mémoire du signal en fonction des paramètres.

Ensuite, nous proposons un modèle basé sur des règles stochastiques de plasticité fonction du temps d'occurrence des décharges électriques neurales (STDP en anglais). Ce modèle est décrit par un Processus de Markov Déterministe par Morceaux (PDMP en anglais). On étudie le comportement en temps long d'un tel réseau de neurones grâce à une analyse lent-rapide. En particulier, on trouve des conditions suffisantes pour lesquelles le processus associé aux poids synaptiques est ergodique.

Enfin, nous faisons le lien entre deux niveaux de modélisation : l'approche microscopique et celle macroscopique. À partir des dynamiques présentées d'un point de vue microscopique (modèle du neurone et son interaction avec les autres neurones), on détermine une dynamique limite qui représente l'évolution d'un neurone typique et de ses poids synaptiques entrant : c'est l'analyse champ moyen du modèle. On condense ainsi l'information sur la dynamique des poids et celle des neurones dans une seule équation, celle d'un neurone *typique*.

Mots clefs : Processus de Markov Déterministes par Morceaux, Multi-échelle, Champ Moyen, STDP, Réseaux de Neurones Récurrents, Plasticité Synaptique, Temps de Mémoire.

Contents

1	Introduction	10
1.1	Some Biology	10
1.1.1	A single neuron	10
1.1.2	Synaptic Plasticity	12
1.2	Modelling Neural Networks with Synaptic Plasticity	15
1.2.1	Notations	15
1.2.2	Modelling STDP	15
1.2.3	Implementing Plasticity within Spiking Neural Network Models	17
1.2.4	Mathematical methods: slow fast analysis and mean field approximation	18
1.3	A new model of STDP within a neural network (Chapters 3 and 4)	20
1.3.1	Motivations for a new model and its description	20
1.3.2	Slow-Fast Analysis	22
1.3.3	Mean Field Analysis	24
1.4	Plasticity and memory (Chapter 2)	25
1.4.1	Network Models of Memory	25
1.4.2	Model and Results	27
2	A Mathematical Analysis of Memory Lifetime in a simple Network Model of Memory	31
2.1	Introduction	31
2.2	The model and the estimator	33
2.2.1	The neural network and the protocol	33
2.2.2	Presentation of the estimator	34
2.2.3	First illustrations	38
2.3	Results	39

2.3.1	Spectrum and SNR analysis	39
2.3.2	Binomial Mixture	41
2.3.3	Main results	44
2.4	Proofs	47
2.4.1	Proofs linked with the Binomial Mixture behaviour	47
2.4.2	Proofs of SNR and Spectrum results	50
2.4.3	Proofs of Theorems on Memory Lifetime	54
2.5	Simulations	57
2.6	Discussion	61
3	Slow-Fast Analysis of a New Stochastic STDP model	65
3.1	Introduction	66
3.2	Presentation of the model and notations	67
3.3	Slow-fast analysis	70
3.3.1	The fast process invariant measure	70
3.3.2	The slow process limit dynamics	73
3.4	Long time behaviour: conditions for recurrence and transience	82
3.4.1	General condition on the rates and parameters	82
3.4.2	Condition using the Laplace transform	90
3.5	Simulations	98
3.5.1	Application of our results	98
3.5.2	Beyond our results	101
3.6	Discussion	102
4	Mean Field Analysis of the Stochastic STDP model	105
4.1	Introduction	106
4.2	Presentation of the model	106
4.2.1	The microscopic model	107
4.2.2	Towards the macroscopic model	109
4.3	Study of the empirical measure and its limit	110

4.3.1	The empirical measure	111
4.3.2	Conjecture on the limit process dynamics	113
4.3.3	The case without plasticity	116
4.3.4	The case with plasticity	126
4.4	Numerical comparison with the neural network	132
4.5	Discussion	134
5	Conclusion and Perspectives	137
A	Appendix	140
A.1	Definitions	140
A.1.1	Convergence in law (or weakly)	140
A.1.2	Variation of signed measures	140
A.1.3	Around Poisson	141
A.1.4	Fréchet differentiation	142
A.1.5	Semigroup Theory and Probability Theory	142
A.1.6	Markov chains	143
A.2	Kronecker Product for Memory Lifetime	144
A.3	Invariant measure of the fast process in dimension 2	146
A.4	Code of Chapter 4	149

Main notations

General notations

$\mathcal{B}(E)$ The Borel sets of E for a metric space (E, d) .

$\mathbb{1}$ The indicator function: $\mathbb{1}_A(x) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{otherwise.} \end{cases}$

δ_x The Dirac mass at x , for all $A \in \mathcal{B}(E)$, $\delta_A(x) = \mathbb{1}_A(x)$.

\vee, \wedge $x \vee y = \max(x, y)$ and $x \wedge y = \min(x, y)$, defined on page 35.

$\lfloor x \rfloor$ The floor function $\lfloor x \rfloor$ is equal to $k \in \mathbb{Z}$ if $k \leq x < k + 1$.

$\stackrel{\mathcal{L}}{=}, \stackrel{\mathbb{E}}{=}$ $\stackrel{\mathcal{L}}{=}$ means equality in law and $\stackrel{\mathbb{E}}{=}$ means equality in expectation.

N Number of neurons of the neural network.

$V_t^N = (V_t^{i,N})_{1 \leq i \leq N}$, describes potentials of neurons $1, \dots, N$. In all chapters $V_t^{i,N} \in \{0, 1\}$.

$S_t^N = (S_t^{i,N})_{1 \leq i \leq N} \in \mathbb{R}_+^N$, the time since the last spike of neurons $1, \dots, N$.

$W_t^N = (W_t^{ij,N})_{1 \leq i \leq N}$, the synaptic weight matrix. $W_t^{ij,N}$ represents the synaptic strength of the neuron j on the neuron i . In Chapter 2, $W_t^{ij,N} \in \{0, 1\}$. In Chapter 3, $W_t^N \in E_2$ defined on page 68. In Chapter 4, $W_t^{ij,N} \in \mathbb{Z}$.

$\|\eta\|_{TV} = \eta^+(X) + \eta^-(X)$ where $\eta = \eta^+ - \eta^-$ is the Jordan decomposition of the signed measure η . Defined 140.

α_m, α_M Bounds on the spiking rates of the neurons at resting potential 0, $0 < \alpha_m < \alpha_M < \infty$.

β Jumping rate of V^i from 1 to 0 for all i . We usually assume $0 < \alpha_m < \beta < \alpha_M < \infty$.

p^\pm Functions from $\mathbb{R}_+ \times \{\text{state space of the weight matrix}\}$ to $[0, 1]$. p^+ gives the potentiation probability. p^- gives the depression probability. For applications, we usually take $p^+(s, w) = A_+ e^{\frac{-s}{\tau^+}} g^+(w)$ and $p^-(s, w) = A_- e^{\frac{-s}{\tau^-}} g^-(w)$.

Δw Size of the weight jumps. We usually take $\Delta w = 1$.

Chapter 2

f_N The coding level: probability a neuron has to be selective to an external signal, $f_N \in (0, 1)$, defined on page 33.

$q^+ \in (0, 1]$, is the potentiation probability when $(W_t^{ij,N}, V_t^{i,N}, V_t^{j,N}) = (0, 1, 1)$, defined on page 34.

$q_{01}^{-,N} \in (0, 1]$, is the depression probability when $(W_t^{ij,N}, V_t^{i,N}, V_t^{j,N}) = (1, 0, 1)$, defined on page 34.

$q_{10}^{-,N} \in (0, 1]$, is the depression probability when $(W_t^{ij,N}, V_t^{i,N}, V_t^{j,N}) = (1, 1, 0)$, defined on page 34.

r Number of presentation of the signal to learn, defined on page 34.

$h_t^N = \sum_{j=2}^{N+1} W_t^{1j,N} V_0^{j,N}$ is the synaptic current onto the neuron 1, defined on page 35.

$K = \sum_{j=2}^{N+1} V_0^{j,N}$ is the number of neurons selective to the signal to learn, defined on page 35.

$h_{t,K}^N \stackrel{\mathcal{L}}{=} (h_t^N | K)$, defined on page 35.

$h_t^{N,y} \stackrel{\mathcal{L}}{=} (h_t^N | V_0^{1,N} = y)$ for $y \in \{0, 1\}$, defined on page 35.

$h_{t,K}^{N,y} \stackrel{\mathcal{L}}{=} (h_t^N | K, V_0^{1,N} = y)$ for $y \in \{0, 1\}$, defined on page 35.

$\hat{V}_t^N : \theta \in \mathbb{R}_+ \mapsto \mathbb{1}_{h_t^N > \theta}$, threshold estimator of $V_0^{1,N}$ at time t , defined on page 35.

$p_e^{y,N}$ For $y \in \{0, 1\}$, they are type I and II errors: $p_e^{1,N} : (t, \theta) \in \mathbb{R}_+ \times \{0, \dots, N\} \mapsto \mathbb{P}(\hat{V}_t^N = 1 | V_0^{1,N} = 0)$ and $p_e^{0,N} : (t, \theta) \mapsto \mathbb{P}(\hat{V}_t^N = 0 | V_0^{1,N} = 1)$. Defined on page 35.

$t_* : (\epsilon, r, N) \in (0, 1) \times \mathbb{N}^* \times \mathbb{N} \mapsto \max_{\theta \in \llbracket 0, N \rrbracket} \left(\inf \left\{ t \geq 1, p_e^{0,N}(t, \theta) \vee p_e^{1,N}(t, \theta) \geq \epsilon \right\} \right)$ is the largest time such that both are smaller than ϵ up to time t_* , defined on page 35.

π_K^N The invariant measure of the Markov chain $(h_{t,K}^N)_{t \geq 1}$, defined on page 37.

$\pi_\infty^N = \sum_{K=0}^N \mathbb{P}(\hat{K} = K) \pi_K^N$ where $\hat{K} = \sum_{j=2}^{N+1} V_0^{j,N}$ is the limit of the law of $(h_t^N)_{t \geq 1}$, defined on page 37.

a_N, b_N Under Assumption 2.9, $q_{01}^{-,N} = a_N f_N$ and $q_{10}^{-,N} = b_N f_N$, defined on page 37.

$P_K^{y,N}$ The transition matrix of the synaptic current $(h_{t,K}^{y,N})_{-r+1 < t \leq 1}$, defined on page 40.

$\nu_{t,K}^{y,N} = [\nu_{t,K}^{y,N}(0), \nu_{t,K}^{y,N}(1), \dots, \nu_{t,K}^{y,N}(K)]$ is the distribution of $h_{t,K}^{y,N}$, defined on page 40.

Λ_y^N, λ_i^N For $y \in \{0, 1\}$, $\Lambda_0^N = 1 - f_N q_{01}^{-,N}$, $\Lambda_1^N = 1 - (1 - f_N) q_{10}^{-,N} - f_N q^+$ and

$$\forall i, \quad \lambda_i^N = (1 - f_N)(\Lambda_0^N)^i + f_N(\Lambda_1^N)^i.$$

Defined on page 40.

$\mu_t^{y,N}, \sigma_t^{y,N}$ For $y \in \{0, 1\}$, $\mu_t^{y,N} = \mathbb{E}[h_t^{y,N}]$ and $(\sigma_t^{y,N})^2 = \text{Var}[h_t^{y,N}]$, defined on page 41.

$SNR_t^N = \frac{\mu_t^{1,N} - \mu_t^{0,N}}{\sqrt{\sigma_{t,K}^{1,N} + \sigma_t^{0,N}}}$ is the SNR at time t , defined on page 41.

$\text{Bin}(n, p)$ is the Binomial law of parameters $n \in \mathbb{N}$ and $p \in [0, 1]$.

$\text{BinMix}(k, g)$ is the Binomial mixture with mixing distribution $g \in \mathcal{P}([0, 1])$ and size parameter $k \in \mathbb{N}$, defined on page 42.

- G_t^N The cumulative distribution function associated to g_t^N . Defined on page 42.
- $g_t^N \in \mathcal{P}([0, 1])$ such that $h_{t,K}^N \stackrel{\mathcal{L}}{=} \text{BinMix}(K, g_t^N)$. Defined on page 42.
- $\mathcal{R} \quad \forall \Gamma \in F_{[0,1]}, \forall u \in \mathbb{R}, \mathcal{R}(\Gamma)(u) \stackrel{\text{def.}}{=} f_N \Gamma\left(\frac{u - f_N q^+}{\Lambda_1^N}\right) + (1 - f_N) \Gamma\left(\frac{u}{\Lambda_0^N}\right)$. Defined on page 42.
- $G^{*,N}$ The unique fixed point of \mathcal{R} . Defined on page 43.
- $g^{*,N}$ The distribution associated to $G^{*,N}$. Defined on page 43.
- $[m_\infty^N, M_\infty^N]$ is the smallest interval containing the support of $g^{*,N}$
- $M_\epsilon^N = \left(\sqrt{\frac{2\text{Var}(Y^{*,N})}{\epsilon}} + \mathbb{E}(Y^{*,N}) \right) \wedge M_\infty^N$. Defined on page 54.
- $[m_t^{1,N}, M_t^{1,N}]$ is the smallest interval containing the support of $g_t^{1,N}$. Defined on page 54.
- $m_{\epsilon,N} = \frac{1}{Nf_N} (\theta_{\epsilon,N} + \sqrt{-2 \log(\epsilon) \theta_{\epsilon,N} - 2 \log(\epsilon)})$. Defined on page 54.
- $t_c^N = \inf\{t \in \mathbb{N}^*, m_t^{1,N} \leq \frac{f_N q^+}{\Lambda_0^N - \Lambda_1^N}\}$. Defined on page 55.
- $C(\epsilon, r, N) = \sqrt{\frac{M_\epsilon^N}{m_1^{1,N} \wedge M_\infty^N}} + 2 \sqrt{\frac{-\log(\frac{\epsilon}{2})}{(m_1^{1,N} \wedge M_\infty^N) N f_N}}$. Defined on page 56.

Chapter 3

- $\alpha_i : \{0, 1\}^N \times \mathbb{R}^{N^2} \rightarrow [\alpha_m, \alpha_M]$, $\alpha_i(V_t, W_t)$ is the jump rate of the neuron i at time t . Defined on 68.
- $E = E_1 \times E_2$ with $E_1 = \{0, 1\}^N \times \mathbb{R}_+^N$, $E_2 = \{w \in \mathbb{N}^{N^2} \text{ s.t. } \forall i, w^{ii} = 0, \forall i \neq j, w^{ij} > 0\}$. Defined on page 69.
- A_w^i Set of accessible states of W_t after the spike of neuron i at time t , conditionally to $W_{t-} = w$. Defined on page 69.
- $\phi_i : \mathbb{R}_+^N \times E_2 \times E_2 \rightarrow [0, 1]$ is the transition of W conditionally to the vector S_{t-} . Defined on page 69.
- \mathcal{C} The generator of the process $(V_t, S_t, W_t)_{t \geq 0}$. Defined on page 69.
- $e_i = (0, \dots, 0, \underbrace{1}_i, 0, \dots, 0)$ the elements of both $\{0, 1\}^N$ and \mathbb{R}_+^N . Defined on page 70.
- ε Slow fast parameter in $(0, 1)$. Defined on page 70.
- \mathcal{B}_w The generator of the two first components (V_t, S_t) of the process $(V_t, S_t, W_t \equiv w)_{t \geq 0}$. Defined on page 70.
- $P_{t,w}$ The transition probability kernel of the process $(V_t, S_t)_{t \geq 0}$. Defined on page 71.
- $(N_{t,w}^i)_{t \geq 0}$ The number of jumps of the process $(V_u^i, S_u^i)_{u \geq 0}$ between the times 0 and t . Defined on page 71.
- $((\zeta^i(du, dz))_{1 \leq i \leq N})_{t \geq 0}$ is a family of N Poisson random measures on \mathbb{R}_+^2 with intensity $dudz$. Defined on page 71.

- π_w The invariant measure of the process with generator \mathcal{B}_w . Defined on page 72.
- $\phi_i^\varepsilon : \mathbb{R}_+^N \times E_2 \times E_2 \rightarrow [0, 1]$ is the transition probability of the weights W_{t-}^ε conditionally to S_{t-}^ε . Defined on page 73.
- $R_i^\varepsilon(s, w) = 1 - \phi_i^\varepsilon(s, w, w)$ for all $(s, w) \in \mathbb{R}_+^N \times E_2$ Defined on page 74.
- $\varphi_i, K_i^\varepsilon(s, \tilde{w}, w) = \phi_i^\varepsilon(s, \tilde{w}, w) - \varepsilon \varphi_i(s, \tilde{w}, w)$ for all $(s, \tilde{w} \neq w, w) \in \mathbb{R}_+^N \times E_2 \times E_2$. Defined on page 74.
- $E_{ij} = (\delta_{ik} \delta_{jl})_{1 \leq k, l \leq N}$, defined on page 74.
- $\varphi^M \in \mathbb{R}^+$ such that for all $i, \sum_{\tilde{w} \neq w} \varphi_i(s, \tilde{w}, w) \leq \varphi^M$. Defined on page 75.
- \mathcal{C}^ε is the same as \mathcal{C} with ϕ_i^ε instead of ϕ_i . Defined on page 75.
- \mathcal{B} . $\mathcal{B}_{net}, \mathcal{B}_{syn}, \mathcal{B}_r^\varepsilon$ and $\mathcal{B}_\uparrow^\varepsilon$ are operators defined in equation (3.13) on page 75.
- $(V_t^\varepsilon, S_t^\varepsilon, W_t^\varepsilon) = (V_{t/\varepsilon}, S_{t/\varepsilon}, W_{t/\varepsilon})$. Defined on page 75.
- $\mathcal{B}_{net}^1, \mathcal{B}_{syn}^2$ are operators defined on page 75.
- $\mathcal{L}_m(X)$ is the space of measures η on $\mathbb{R}_+ \times X$ such that for all $t \geq 0$, we have $\eta([0, t] \times X) = t$, with X a complete separable metric space. Defined on page 76.
- $D_X[0, +\infty)$ The space of càdlàg functions (right continuous with left limits) from $[0, +\infty)$ to X . Defined on page 76.
- Γ^ε The occupation measure associated to the process $(V_t^\varepsilon, S_t^\varepsilon)_{t \geq 0}$: for all $t \geq 0, A \in \mathcal{B}(E_1)$,
- $$\Gamma^\varepsilon([0, t], A) \stackrel{def.}{=} \int_0^t \mathbb{1}_A(V_u^\varepsilon, S_u^\varepsilon) du.$$
- Defined on page 76.
- $(\Gamma, \overline{W}_t)_{t \geq 0}$ $(\Gamma^\varepsilon, W_t^\varepsilon)_{t \geq 0}$ converges in law to $(\Gamma, \overline{W}_t)_{t \geq 0}$ in $\mathcal{L}_m(E_1) \times D_{E_2}[0, +\infty)$. Defined on page 77.
- $\mathcal{C}_{avg}(w) = \int_{E_1} \mathcal{B}_{syn}^2 g(v, s, w) \pi_w(dv, ds)$. for all $g \in D(\mathcal{C}_{avg})$ is the generator of $W_t^\varepsilon)_{t \geq 0}$. Defined on page 77.
- $r : E_2^2 \rightarrow \mathbb{R}$ is the Q -matrix of $(\overline{W}_t)_{t \geq 0}$, $r(w, w) = -\sum_{\tilde{w} \neq w} r(\tilde{w}, w)$ and
- $$\forall \tilde{w} \neq w, \quad r(\tilde{w}, w) \stackrel{def.}{=} \int_{E_1} \sum_k \delta_0(v^k) \alpha_k(v, w) \varphi_k(s, \tilde{w}, w) \pi_w(dv, ds).$$
- Defined on page 82.
- $r_{ij}^\pm \quad \forall i \neq j, \forall w \in E_2, \quad r_{ij}^+(w) = \sum_{\substack{\tilde{w} \in A_w^i \\ \tilde{w}^{ij} = w^{ij+1}}} r(\tilde{w}, w), \quad r_{ij}^-(w) = \sum_{\substack{\tilde{w} \in A_w^j \\ \tilde{w}^{ij} = w^{ij-1}}} r(\tilde{w}, w).$
- Defined on page 82.
- R^\pm N -by- N matrices, such that $R^{\pm ii} = 0$ and $R^+ = (r_{ij}^+)_{1 \leq i \neq j \leq N}$, $R^- = (r_{ij}^-)_{1 \leq i \neq j \leq N}$. Defined on page 82.
- $A \cdot B = \sum_{1 \leq i, j \leq N} A^{ij} B^{ij}$ for any square matrices A, B . Defined on page 82.

$(V_w, S_w) \stackrel{\mathcal{L}}{=} \pi_w$. Defined on page 82.

$\mathcal{B}_\alpha f(\hat{v}, \hat{s}) = \partial_{\hat{s}} f(\hat{v}, \hat{s}) + \beta \delta_1(\hat{v}) [f(0, \hat{s}) - f(1, \hat{s})] + \alpha \delta_0(\hat{v}) [f(1, 0) - f(0, \hat{s})]$ for $\alpha \in \{\alpha_m, \alpha_M\}$ and $\forall f \in D(\mathcal{B}_\alpha), \forall (\hat{v}, \hat{s}) \in \{0, 1\} \times \mathbb{R}^+$. Defined on page 85.

$(\bar{V}_{t,j}, \bar{S}_{t,j})_{t \geq 0}$ and $(\underline{V}_{t,j}, \underline{S}_{t,j})_{t \geq 0}$ are processes in E_1 defined on page 85.

μ_α The invariant probability measure associated to the generator \mathcal{B}_α . Defined on page 85.

$\underline{\pi}_j, \bar{\pi}_j$ Respectively the invariant measure of $(\underline{V}_{t,j}, \underline{S}_{t,j})_{t \geq 0}$ and $(\bar{V}_{t,j}, \bar{S}_{t,j})_{t \geq 0}$. Defined on page 85.

$(\underline{V}_k, \underline{S}_k) \stackrel{\mathcal{L}}{=} \underline{\pi}_k$ and $(\bar{V}_k, \bar{S}_k) \stackrel{\mathcal{L}}{=} \bar{\pi}_k$. Defined on page 87.

$\mathcal{L}\{\rho\}$ is the Laplace transform of the measure ρ on $\mathbb{R}_+^N: \forall \lambda = (\lambda^1, \dots, \lambda^N) \in \mathbb{C}_+^N$

$$\mathcal{L}\{\rho\}(\lambda) = \int_{\mathbb{R}_+^N} e^{-\lambda \cdot s} \rho(ds),$$

where $\lambda \cdot s = \sum_{k=1}^N \lambda^k s^k$. Defined on page 90.

ν_w^v, π_w^v Respectively the marginal in v of the distribution π_w the law of π_w conditionally to v . Defined on page 90.

Υ_{π_w} is the vector of the Laplace transforms of $\nu_w^v \pi_w^v$. Defined on page 91.

(v_1, \dots, v_{2N}) is an enumeration of $\{0, 1\}^N$ such that $k \geq l \Rightarrow \sum_{i=1}^N v_k^i \geq \sum_{i=1}^N v_l^i$. Defined on page 94.

$\hat{\lambda}_{l_1, \dots, l_d} \forall i \in \{l_1, \dots, l_d\}, \hat{\lambda}_{l_1, \dots, l_d}^i = 0$ and $\forall i \in \llbracket 1, N \rrbracket \setminus \{l_1, \dots, l_d\}, \hat{\lambda}_{l_1, \dots, l_d}^i = \lambda^i$. Defined on page 94.

$\check{\lambda}_l = \hat{\lambda}_{1, \dots, l-1, l+1, \dots, N} = (0, \dots, 0, \lambda^l, 0, \dots, 0)$. Defined on page 94.

Chapter 4

α $\alpha : \mathbb{R} \rightarrow [\alpha_m, \alpha_M]$ is a strictly positive and bounded function, defined on page 107.

$d\theta$ $\prod_{k=1}^N d\theta_k^+ \prod_{k=1}^N d\theta_k^-$. Defined on page 107.

$E^{(N)} = \mathbb{R}^+ \times [0, 1]^{2N}$. Defined on page 107.

$Y_t^N = (Y_t^{1,N}, \dots, Y_t^{N,N})$ where $Y_t^{i,N} = (V_t^{i,N}, S_t^{i,N}, (W_t^{ij,N})_{1 \leq j \leq N})$. Defined on page 107.

κ_0^N, ρ_0 $Y_t^{1,N}, \dots, Y_0^{N,N}$ are *i.i.d.* with law κ_0^N . The law ρ_0 of $S_0^{i,N}$ is absolutely continuous with respect to the Lebesgue measure λ . The density is assumed to be bounded. Defined on page 108.

T, τ Respectively the sequence of the jumping times of $(\mu_t^N)_{t \geq 0}$ and the sequence of the spiking times. Defined on page 109.

$\mathcal{P}_N(G)$ The set of atomic probability measures on G with N distinct atoms of weight $\frac{1}{N}$. Defined on page 109.

$\xi_t^{i,N}$ The empirical measure $\forall i = 1, \dots, N$, $\xi_t^{i,N} \stackrel{def.}{=} \frac{1}{N} \sum_j \delta_{(V_t^{j,N}, S_t^{j,N}, W_t^{ij,N})}$. Defined on page 109.

$E_m = \{0, 1\} \times \mathbb{R}^+ \times \mathbb{Z}$. Defined on page 109.

$I : \mathcal{P}(E_m) \rightarrow \mathbb{R}$ such that $\forall \xi \in \mathcal{P}(E_m)$, $I(\xi) \stackrel{def.}{=} \mathbb{E}^\xi(WV) = \int_{E_m} wv \xi(dv, ds, dw)$. Defined on page 109.

$X_t^{i,N}$ The neurons are described by the following triplets

$$\forall i = 1, \dots, N, \quad X_t^{i,N} \stackrel{def.}{=} (V_t^{i,N}, S_t^{i,N}, \xi_t^{i,N}) \in E.$$

Defined on page 110.

μ_t^N The empirical measure $\mu_t^N \stackrel{def.}{=} \frac{1}{N} \sum_i \delta_{X_t^{i,N}} = \frac{1}{N} \sum_i \delta_{(V_t^{i,N}, S_t^{i,N}, \xi_t^{i,N})}$. Defined on page 111.

$E = \{0, 1\} \times \mathbb{R}^+ \times \mathcal{P}(E_m)$. Defined on page 111.

$q = \{q_1, q_2, \dots\}$ is the sequence of the spiking neuron numbers. Defined on page 111.

$\lambda^J : \mathcal{P}_N(E) \rightarrow [\alpha_m N, \alpha_M N]$ is the jump rate of μ_t^N such that for all $t \geq 0$,

$$\lambda^J(\mu_t^N) \stackrel{def.}{=} \sum_i \alpha(I_t^{i,N}) \mathbb{1}_{\{V_t^{i,N}=0\}} + \beta \mathbb{1}_{\{V_t^{i,N}=1\}}.$$

Defined on page 111.

$\hat{\xi}_{t,j_0}^{j,N} = \xi_{t^-}^{j,N} + \frac{1}{N} \sum_{(\tilde{V}, \tilde{S}, \tilde{W}) \in \text{supp}(\xi_{t^-}^{j,N})} \mathbb{1}_{\{\tilde{S}=S_{t^-}^{j_0,N}\}} \left(\delta_{(0, \tilde{S}, \tilde{W})} - \delta_{(1, \tilde{S}, \tilde{W})} \right)$. Defined on page 111.

$\check{\xi}_{t,j_0}^{j,N} = \xi_{t^-}^{j,N} + \frac{1}{N} \sum_{(\tilde{V}, \tilde{S}, \tilde{W}) \in \text{supp}(\xi_{t^-}^{j,N})} \mathbb{1}_{\{\tilde{S}=S_{t^-}^{j_0,N}\}} \left(\delta_{(1,0, \tilde{W} - \mathbb{1}_{\{U_{\tilde{V}, \tilde{S}}^{j,k} \leq p^-(S_{t^-}^{j,N}, \tilde{W})\}})} - \delta_{(\tilde{V}, \tilde{S}, \tilde{W})} \right)$.

Defined on page 111.

μ_t^* Candidate to be the deterministic limit of μ_t^N as N tends to infinity. Defined on page 114.

$X_t^* = (V_t^*, S_t^*, \xi_t^*)$ with law μ_t^* . Defined on page 114.

ξ^y ξ in $\mathcal{P}(E_m)$ can be split into the following form

$$\forall (v, A, w) \in \{0, 1\} \times \mathcal{B}(\mathbb{R}^+) \times \mathbb{Z}, \quad \xi(\{v\}, A, \{w\}) = \delta_0(v) \xi^0(A, \{w\}) + \delta_1(v) \xi^1(A, \{w\}).$$

Defined on page 114.

$\delta_x \otimes \xi^y$ For all $x, y \in \{0, 1\}$, $\delta_x \otimes \xi^y$ is the measure such that

$$\forall (v, A, w) \in \{0, 1\} \times \mathcal{B}(\mathbb{R}^+) \times \mathbb{Z}, \quad \delta_x \otimes \xi^y(\{v\}, A, \{w\}) \stackrel{def.}{=} \delta_x(v) \xi^y(A, \{w\}).$$

Defined on page 114.

$p_t^{*v}(\cdot, w)$ The densities of the laws $\xi_t^{*v}(\cdot, \{w\})$. Defined on page 115.

$\mathcal{M}(E_m)$ Space of the signed measures on E_m . Defined on page 117.

$C_b^{1,1}(E)$ The space of functions from E to \mathbb{R} that are bounded, continuously differentiable with respect to their second variable, Fréchet differentiable (see Definition A.6) with respect to their third variable and finally, with both these derivatives bounded. Defined on page 117.

$\partial_\xi \Psi$ $\forall \Psi \in C_b^{1,1}(E)$, $\forall (v, s, \xi) \in E$ and $\forall h \in \mathcal{M}(E_m)$, we denote by $\partial_\xi \Psi(v, s, \xi) \cdot h$ the Fréchet derivative of Ψ at (v, s, ξ) in the direction h (see Definition A.6). Defined on page 117.

ν^N, η^N $\forall (s, \xi) \in \mathbb{R}^+ \times \mathcal{P}_N(E_m)$,

$$\eta^N(s, \xi) \stackrel{def.}{=} \frac{1}{N} \sum_{(1, \tilde{s}, \tilde{w}) \in \text{supp}(\xi)} \mathbb{1}_{\{\tilde{s}=s\}} \left[\delta_{(0, \tilde{s}, \tilde{w})} - \delta_{(1, \tilde{s}, \tilde{w})} \right]$$

$$\nu^N(s, \xi) \stackrel{def.}{=} \frac{1}{N} \sum_{(0, \tilde{s}, \tilde{w}) \in \text{supp}(\xi)} \mathbb{1}_{\{\tilde{s}=s\}} \left[\delta_{(1, 0, \tilde{w})} - \delta_{(0, \tilde{s}, \tilde{w})} \right].$$

Defined on page 117.

$\langle -\varepsilon', \phi \rangle = \langle \varepsilon, \phi' \rangle \forall \phi \in C_b^1(\mathbb{R}^+)$ with bounded derivative. Defined on page 117.

$\xi \oplus t$ Let $\xi \in \mathcal{P}(E_m)$, $t \geq 0$, then $\forall (v, A, w) \in \mathcal{B}(\{0, 1\}, [t, +\infty[, \mathbb{Z})$,

$$(\xi \oplus t)(\{v\}, A, \{w\}) = \xi(\{v\}, A - t, \{w\}).$$

Defined on page 117.

$$\textcircled{1} = \int_0^t \sum_{i,j} \mathbb{1}_{\{V_{u^-}^{i,N}=1\}} \frac{\beta}{N} \left[\Psi \left(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u^-}^{j,N} + \eta^N(S_{u^-}^{i,N}, \xi_{u^-}^{j,N}) \right) - \Psi \left(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u^-}^{j,N} \right) \right] du.$$

Defined on page 120.

$$\textcircled{2} = \int_0^t \sum_{i,j} \mathbb{1}_{\{V_{u^-}^{i,N}=0\}} \frac{\alpha(I_{u^-}^{i,N})}{N} \left[\Psi \left(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u^-}^{j,N} + \nu^N(S_{u^-}^{i,N}, \xi_{u^-}^{j,N}) \right) - \Psi \left(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u^-}^{j,N} \right) \right] du.$$

Defined on page 120.

$$\gamma_\xi \quad \text{Let } \xi \in \mathcal{P}(E_m), \text{ then } \forall (s, w) \in \mathbb{R}^+ \times \mathbb{Z}, \gamma_\xi(s, w) = \begin{cases} \lim_{\epsilon \rightarrow 0} \frac{\xi^0([s, s+\epsilon], \{w\})}{\xi^0([s, s+\epsilon], \mathbb{Z})} & \text{if } s \in \text{supp}(\xi) \\ 0 & \text{otherwise.} \end{cases}$$

Defined on page 122.

ν^1, ν^0 Let $\xi \in \mathcal{P}(E_m)$ and $\forall \mu \in \mathcal{P}(E)$, then $\forall (A, w) \in \mathcal{B}(\mathbb{R}^+) \times \mathbb{Z}$,

$$\nu^0(\xi, \mu)(A, \{w\}) \stackrel{def.}{=} \int_{\mathcal{P}(E_m)} \int_A \alpha(I(\tilde{\xi})) \gamma_\xi(s, w) \mu(\{0\}, ds, d\tilde{\xi}),$$

$$\nu^1(\xi, \mu)(A, \{w\}) \stackrel{def.}{=} \mathbb{1}_{\{0 \in A\}} \nu^0(\xi, \mu)(\mathbb{R}^+, \{w\}).$$

Defined on page 123.

J_i^N The possible increments of the weights when the neuron i spikes

$$J_i^N = \{A = (a^{kl})_{1 \leq k, l \leq N} : a^{ii} \in \{-1, 0, 1\}, \forall k, l \neq i, a^{il} \in \{0, 1\} \text{ and } a^{ki} \in \{0, -1\}\}.$$

Defined on page 127.

$P_{t,i}^\Delta$ For all $\Delta \in J_i^N$, we denote by $P_{t,i}^\Delta$ the probability that the weight matrix be incremented by Δ conditionally to X_t^N . Defined on page 127.

$\xi_{t,\Delta}^{i,N}, \xi_{t,\Delta,i}^{j,N}$ After the spike of the neuron i at time t and assuming the weights are incremented of $\Delta \in J_i^N$ at this time, $\xi_{t^-}^{i,N}$ jumps to $\xi_{t,\Delta}^{i,N}$ and for all j , $\xi_{t^-}^{j,N}$ jumps to $\xi_{t,\Delta,i}^{j,N}$. Defined on page 127.

$$\textcircled{3} = \sum_i \mathbb{1}_{\{V_{u^-}^{i,N}=0\}} \frac{\alpha(I_{u^-}^{i,N})}{N} \sum_{\Delta \in J_i^N} P_{u,i}^\Delta \left[\Psi(1, 0, \xi_{u,\Delta}^{i,N}) - \Psi(0, S_{u^-}^{i,N}, \xi_{u,\Delta,i}^{i,N}) \right]. \text{ Defined on page 128.}$$

$$\textcircled{4} = \sum_i \mathbb{1}_{\{V_{u^-}^{i,N}=0\}} \frac{\alpha(I_{u^-}^{i,N})}{N} \sum_{\Delta \in J_i^N} P_{u,i}^\Delta \sum_j \left[\Psi(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u,\Delta,i}^{j,N}) - \Psi(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u^-}^{j,N}) \right]. \text{ Defined on page 128.}$$

$\nu^{-,1}$ Let $(s, \xi, \mu) \in \mathbb{R}^+ \times \mathcal{P}(E_m) \times \mathcal{P}(E)$, then $\forall (A, \{w\}) \in \mathcal{B}(\mathbb{R}^+) \times \mathbb{Z}$,

$$\begin{aligned} \nu^{-,1}(s, \xi, \mu)(A, \{w\}) &\stackrel{def.}{=} \mathbb{1}_{\{0 \in A\}} \int_{\mathcal{P}(E_m)} \alpha(I(\tilde{\xi})) p^-(s, w+1) \gamma_\xi(\tilde{s}, w+1) \mu(\{0\}, d\tilde{s}, d\tilde{\xi}) \\ &\quad + \mathbb{1}_{\{0 \in A\}} \int_{\mathcal{P}(E_m)} \alpha(I(\tilde{\xi})) (1 - p^-(s, w)) \gamma_\xi(\tilde{s}, w) \mu(\{0\}, d\tilde{s}, d\tilde{\xi}), \end{aligned}$$

Defined on page 130.

$$\nu^+(\xi) = \int_A p^+(s, w-1) \xi(\{v\}, ds, \{w-1\}) + \int_A (1 - p^+(s, w)) \xi(\{v\}, ds, \{w\}) \text{ for all } \xi \in \mathcal{P}(E_m). \text{ Defined on page 131.}$$

Chapter 1

Introduction

In this Introduction, we first present a brief overview of the biology of the brain based on Section I.2 of the book [73]. Then, we describe some models of synaptic plasticity before presenting the one we developed in Chapters 3 and 4. We finally introduce the network model of memory that we consider in Chapter 2. We have opted to end the introduction with a description of this chapter because it contains a very simplified model quite far from the biology that we present first.

1.1 Some Biology

The two main types of cells in the nervous system are the nerve cells (neurons) and the glial cells (glia). Understanding the building blocks of the brain is the first step toward its global comprehension. In most of the brain modelling field, these blocks are the neurons. Thereby, we do not consider glia in this thesis but rather models of networks composed of interacting neurons. In this section, we briefly describe how a plastic neural network biologically works: neurons convey information through action potentials reaching other neurons via synapses whose efficiencies change over time.

1.1.1 A single neuron

A neuron is an electrically excitable cell with a soma (cell body), an axon and dendrites, see Figure 1.1a. The soma is composed of a nucleus (containing the genes) and an endoplasmic reticulum (where the protein synthesis occurs). The cell body is the metabolic centre of the neuron giving rise to short dendrites and one long tubular axon. The dendrites receive the incoming signals from the other cells when the axon carries electrical signals to other neurons situating from 0.1 mm to 2 m. These electrical signals are called action potentials. They are initiated near the origin of the axon and propagate along the axon without failure or distortion and with a constant amplitude around 100 mV. It is an all-or-none impulse regularly regenerated along the axon. This is the (main) way by which the brain receives, analyses and conveys information. One striking fact is that the action potentials are very similar over the brain even though they can be initiated by very different events: “the information conveyed by an action potential is determined not by the form of the signal but by the pathway the signal travels in the brain” [73]. Then, our brain creates our sensations from this pathway.

Neural signalling is mainly governed by the electrical properties of the neuron membranes. The membrane potential of a neuron is defined as the potential difference between inside and outside the cell which stems from the respective ions concentrations. The majority of these ions are either potassium ions (K^+) or sodium ions (Na^+). Their flows in and out the neurons are regulated by

membrane proteins called ion pumps (active) and ion channels (passive). This potential usually ranges from around -100 mV to 60 mV (it depends on the neuron considered). Its dynamics is very complex, ranging from oscillations under the action potential threshold to bursting, see [72, Sec 7] for more details. In this thesis, we only use the generic dynamics : a neuron is either at rest, with a potential fluctuating in the interval -80 to -40 mV, until it emits an action potential (also called spike) defined as a brief (1 ms) increase of around 100 mV of the potential and then a brief (1 ms) decrease under the resting potential followed by a return to this resting state, see Figure 1.1b. The last part of a spike is called the refractory period. In this manuscript, the following expressions are synonymous: to emit an action potential, to fire a spike, to fire, to spike. Spikes are

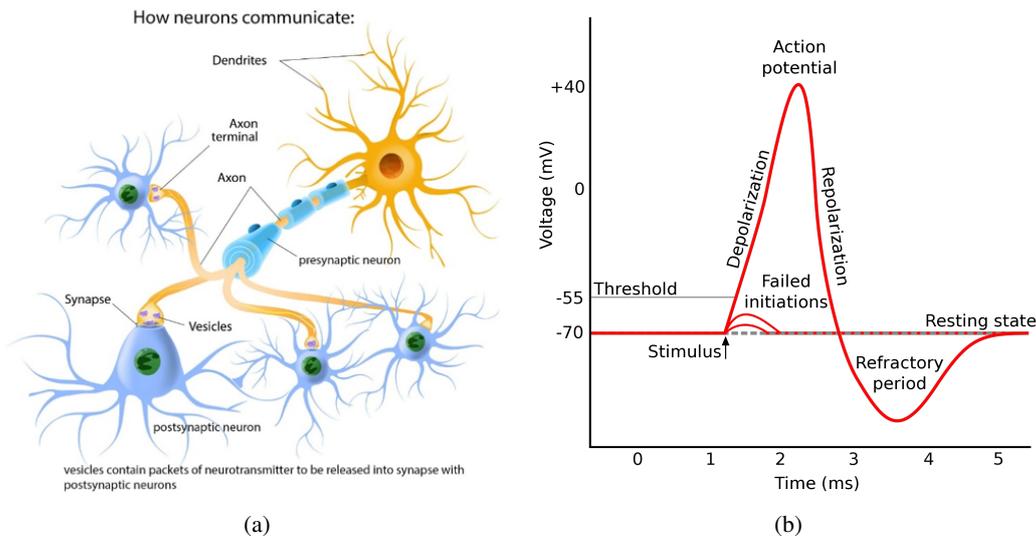


Figure 1.1: (1.1a): Scheme presenting the main actors of neuron communication. (1.1b): Scheme of a typical action potential of a neuron. (1.1a) is taken from the website [1] and (1.1b) is taken from wikipedia [131].

the principal way of conveying information between neurons. Signals from the other neurons are gathered by the dendrites of a neuron. The soma processes these signals and generates an action potential or not, see the failed initiations in Figure 1.1b. This spike is carried through the axon until it reaches the axon terminal where the synapses transfer the information to the other neurons. This information can be noticed by small variations in the postsynaptic membrane potential of the neurons concerned. If the spiking neuron is excitatory (resp. inhibitory), then the membrane is depolarised (resp. hyperpolarised) and it is called an excitatory postsynaptic potential (EPSP) (resp. inhibitory postsynaptic potential (IPSP)). An EPSP (resp. IPSP) is caused by the flow of positive (resp. negative) ions into the membrane which is called an excitatory postsynaptic current (EPSC) (resp. inhibitory postsynaptic current (IPSC)). We end this paragraph mentioning that the sequence of EPSPs precedes the spike of a neuron. Indeed, when the membrane potential reaches a certain random threshold, a spike is generated. As we have already evoked, even though we know the mechanisms explaining such a threshold, it depends on parameters making it random (the thermal motions and fluctuations for example, see Section 9 of [73]).

Even if it goes beyond the scope of this thesis, we would like to mention the current growing interest in the cells surrounding the neurons: the glia cells. Indeed, their role seems to be much more important than what biologists expected and may be part of the future brain models, see [36].

1.1.2 Synaptic Plasticity

Synaptic Signalling

Synapses are “specialised regions that permit chemical or electrical signalling between neurons” [73]. There exist many types of synapses but the majority are axodendritic ones: synapses composed of an axon terminal ending on a dendritic spine, see Figure 1.1a. Most synapses are chemical ones. We focus here on these synapses rather than on the electrical ones. They convey information via chemical messengers called neurotransmitters. Following an action potential, some vesicles containing neurotransmitters bind the axon terminal membrane and release neurotransmitters within the synaptic cleft. These neurotransmitters bind the receptors of the postsynaptic side, see Figure 1.2. We do not want to enter into more details as we do not need it in the following. We refer to [80] for more details.

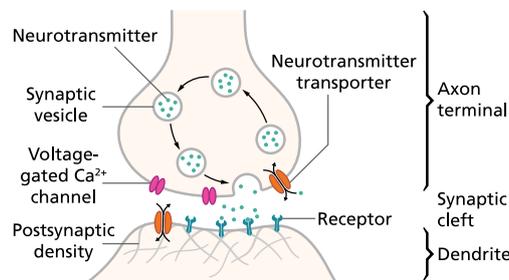


Figure 1.2: Scheme of synaptic signalling taken from wikipedia [132].

Synaptic Weight

We gave the basics of how signals are transmitted between neurons. This transmission is variable. The action potential of a neuron can trigger the release of many neurotransmitters in some of its synapses as it can release none of them in other synapses. In this thesis, we loosely refer to the synaptic weight as the (signed) amplitude of the postsynaptic potential following a presynaptic spike. It depends on the number of postsynaptic receptors (e.g. AMPA), their conductance, etc. This definition is directional in the sense that the synaptic weight linking the presynaptic neuron A to the postsynaptic neuron B (denoted by $W^{B \leftarrow A}$ in what follows) is different from the synaptic weight linking the neuron B to the neuron A ($W^{A \leftarrow B}$). The simplest classification of neurons is certainly to spare neurons in two groups: the inhibitory versus the excitatory. The main difference between them is on their effect postsynaptic neurons. This difference is due to the different types of neurotransmitters they release. As seen previously, excitatory neurons tend to depolarise their postsynaptic neurons whereas inhibitory neurons tend to hyperpolarise theirs. On a modelling point of view, the neuron B is inhibitory (resp. excitatory) when all its outgoing synaptic weights are negative: for any neuron A , $W^{A \leftarrow B}$ is negative (resp. positive). Finally, we mention that, in our model, the synaptic weight $W^{A \leftarrow B}$ takes into account all the axon terminals of the neuron B connected to a dendritic spine of neuron A (there can be until around a dozen of them).

Synaptic Plasticity

The synaptic weights depend on the spine shapes and chemical composition, see [112]. Indeed, the signal transmission depends on the number and density of postsynaptic spine receptors (NMDA and AMPA) as well as the number of neurotransmitters released by the presynaptic neuron, see [80]. Formally, the higher these numbers are, the stronger the connection is. Thus, the larger the synapse is, the bigger these numbers can be and then the stronger the connection is. All of these parameters change over time making the synaptic weights evolve: this phenomenon is called *synaptic plasticity*.

In his book [67], Donald Hebb postulated that memory formation follows three steps: first, synaptic plasticity, second, cell assembly formation (strongly connected neurons) and finally, formation of a phase sequence (series of connected cell assemblies), see [83] for more details. Concerning synaptic plasticity, Hebb mainly detailed his thoughts on the synaptic strengthening. His postulate is usually called the Hebb's rule:

“When an axon of cell A is near enough to excite cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells so that A's efficiency, as one of the cells firing B, is increased.” [67].

However, Hebb did not enter into details on the synaptic weakening which was found experimentally many years later, see [91]. The increase (resp. decrease) of synaptic weights is called potentiation (resp. depression).

Synaptic plasticity can be separated in different phenomena operating at different timescales: short-term plasticity, long-term plasticity and synaptic consolidation [102]. Although, long and short-term plasticity are induced on a similar timescale, which ranges from several seconds to one minute, long term plasticity persists much longer, about hours to days, whereas short term plasticity persists from seconds to minutes, see Section 65 of [73].

Concerning synaptic consolidation, it remains from days to entire life and can be induced really quickly such as with a shock or on a longer timescale like the learning of a poem. Here, we are interested in the long term plasticity.

Long term plasticity is composed of two phenomena: long term potentiation (LTP) and long term depression (LTD). Long lasting changes of the synaptic weight were already evoked more than one century ago by Ramón y Cajal, see [37]. The first experimental evidence of such a plasticity was shown by Bliss and Lømo on LTP, see [20]. Since then, experimental results showed a dependence of synaptic plasticity on different parameters such as the firing rate of the presynaptic neuron, the membrane potential of the postsynaptic neuron, the calcium concentrations, and the precise spiking times of the pre- and postsynaptic neurons. This list is not exhaustive and more details are given in [80]. Here, we are interested in the plasticity resulting from the precise spiking times of the pre- and postsynaptic neurons. It is commonly called Spike-Timing Dependent Plasticity (STDP).

Spike-Timing Dependent Plasticity (STDP)

STDP is a mechanism inducing synaptic plasticity (LTP or LTD) based on the relative timing of pairs of pre- and postsynaptic spikes. The repetition of similar (in causality) pairings leads to synaptic strength changes. STDP is adapted for modelling and such models are often called STDP rules rather than STDP models. An originality of the STDP phenomenon is that it has not been

discovered experimentally but rather by a modelling study in [55]. Then, it was experimentally confirmed some years after in [92, 17] with the classical paradigm: repetition of causal spike pairings (presynaptic spike followed by postsynaptic one) evokes LTP whereas repetition of anticausal spike pairings (postsynaptic spike followed by presynaptic one) leads to LTD. This is not always the case: in the electric fish, the rule is inverted, see for instance [13].

We now detail the typical protocol of the experiments showing this kind of plasticity rules, see [17]. We are interested in the synaptic plasticity from the presynaptic neuron A to the postsynaptic neuron B : $W^{B \leftarrow A}$. Let A and B be forced to emit an action potential in a precise manner: for one minute, every second the neurons A and B spike with a delay Δt between the spikes (which can be negative). The excitatory postsynaptic current (EPSC) after these 60 pairings, $EPSC_{after}$, is compared to the initial one (EPSC before the pairings), $EPSC_{init}$. The quantity

$$\frac{EPSC_{after} - EPSC_{init}}{EPSC_{init}} \approx \frac{W_{after}^{B \leftarrow A} - W_{init}^{B \leftarrow A}}{W_{init}^{B \leftarrow A}}$$

is plotted in function of Δt , see Figure 1.3a. We observe that the change of weight has the same sign as Δt , thus respecting the causality.

Many questions arise from these experiments. Let us evoke three of them. First, how the effect of one pairing on the synaptic weights can be modelled? There exist at least three methods to model these changes: the additive, multiplicative or mixed method. For more details, we refer to [46] (additive and multiplicative methods are presented in 1.2.2). Then, we can ask if it is enough to consider pair of spikes from neurons A and B to model plasticity or is it better to consider longer sequences? For instance, some experiments have been done using triplet of spikes. The results show an advantage given to potentiation. In the experiment leading to the following results, see [130], a triplet of spikes is considered: Δt_1 gives the first delay and Δt_2 the second one. When Δt_1 is positive (resp. negative), two (resp. one) spikes of the presynaptic neuron are considered, see Figure 1.3b. Finally, given the high variability of the results, stochastic models seem reasonable to use.

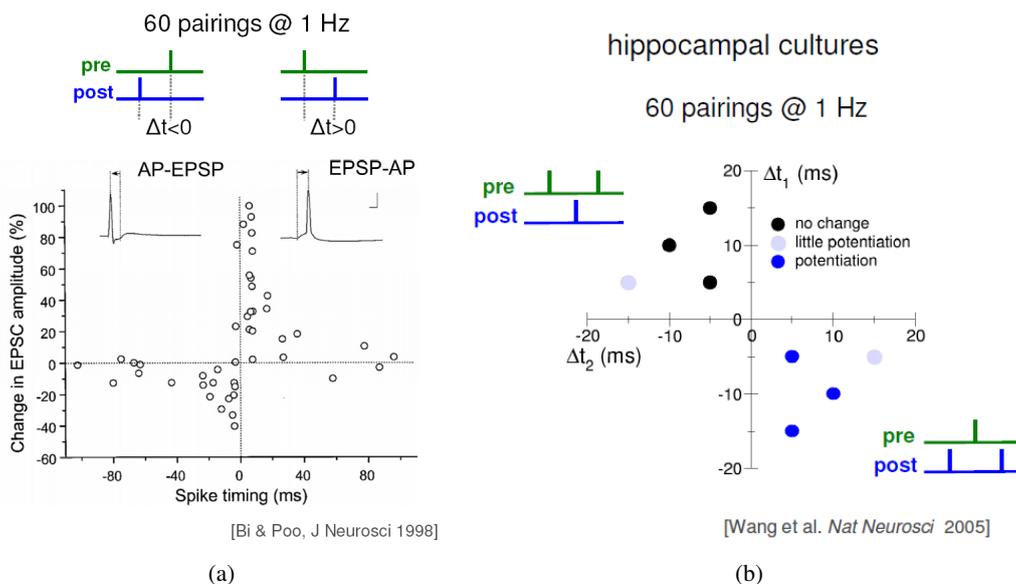


Figure 1.3: (1.3a): Classical experiment to show STDP. (1.3b): Triplet of spikes experiment.

1.2 Modelling Neural Networks with Synaptic Plasticity

We saw that synaptic plasticity is the result of the neural activity. However, before we get this knowledge about plasticity, the idea of flexibility in neural networks as well as some properties of long-term changes were already proposed in 1949 by Hebb [67]. Inspired by his postulates, see page 13, some mathematical models emerged. Because of the lack of biological understanding at this time, they are phenomenological models of plasticity in contrast to biophysical ones. In this section, after giving notations, we present some models of STDP. Then, we illustrate, with the example of the leaky integrate and fire as neuron model, a way to include a plasticity rule in an interacting neural network. Finally, we present some mathematical methods used in this thesis to study such models: timescale separation and mean field approximation.

1.2.1 Notations

In the following, we are interested in modelling neural networks composed of N interacting neurons. We label these neurons from 1 to N . We denote by $V_t^{i,N} \in \mathbb{R}$ the membrane potential at time t of the neuron i and $V_t^N = (V_t^{i,N})_{1 \leq i \leq N}$. The time spent since the last spike of the neuron i is denoted by $S_t^{i,N} \in \mathbb{R}^+$ and the complete vector is $S_t^N = (S_t^{i,N})_{1 \leq i \leq N}$. We denote by $W_t^N = (W_t^{ij,N})_{1 \leq i,j \leq N} \in \mathbb{R}^{N^2}$ the synaptic weight matrix at time t . The strength of the effect of the neuron j on neuron i at time t is given by $W_t^{ij,N}$ ($W_t^{i \leftarrow j,N}$ in the previous notation which is replaced by $W_t^{ij,N}$ in the following). We denote by $I_t^{i,N}$ the synaptic current defined as

$$I_t^{i,N} = \frac{1}{N} \sum_j W_t^{ij,N} V_t^{j,N}.$$

1.2.2 Modelling STDP

On the modelling point of view, some important properties of synaptic plasticity are locality, cooperativity, competition and the boundedness of the weights, see [93, 57]. Locality means that the main variables responsible for plasticity are local: closely linked to the synapses. Cooperativity means that the neurons must be simultaneously active in order to give rise to a synaptic modification. Competition between synapses comes from the fact that resources are finite so that there should be a balance when weights are modified: when there is a synapse potentiated, some other synapses should then be depressed. The boundedness property can be justified with a similar argument. We see how STDP models match with these different properties.

Following the first experiments on STDP, many plasticity models based on the precise spiking times have been proposed. At the beginning, physicists linked STDP models to rate based ones, see for instance [79], and highlighted the difference between them. Precise spiking times have then been showed to play a crucial role in structuring the synaptic weight matrix, see [31, 106, 122]. We give here the general ideas of existing STDP models without entering into the details, see [102] for a very good review.

We consider a fully connected neural network. Therefore, the synapses for which the neuron i is the postsynaptic (resp. presynaptic) neuron are given by the weights $(W_t^{ij,N})_{1 \leq j \leq N}$ (resp. $(W_t^{ji,N})_{1 \leq j \leq N}$). We denote by $(T_k^i)_{k \in \mathbb{N}^*}$ the sequence of spiking times of the neuron i . We consider only pair-based models. This means that as soon as one of the neuron spikes, we only

consider the pairs that it forms with the last spikes (not the previous ones) of the other neurons. We detail this kind of rule formally. Each time a neuron spikes, say neuron i at time T_k^i , we apply the following plasticity rule to every weight associated to the neuron i :

$$W_{T_k^i}^{ij,N} = W_{T_k^{i-}}^{ij,N} + F^+ \left(S_{T_k^{i-}}^{j,N}, W_{T_k^{i-}}^{ij,N} \right) \quad \text{and} \quad W_{T_k^i}^{ji,N} = W_{T_k^{i-}}^{ji,N} - F^- \left(S_{T_k^{i-}}^{j,N}, W_{T_k^{i-}}^{ji,N} \right),$$

where F^+ and F^- are two functions from $\mathbb{R}^+ \times \mathbb{R}$ to \mathbb{R}^+ and T_k^{i-} represents the time just before T_k^i . Most models have used functions with separable variables, that is to say, for all $(s, w) \in \mathbb{R}^+ \times \mathbb{R}$:

$$F^+(s, w) = F_1^+(s)F_2^+(w) \quad \text{and} \quad F^-(s, w) = F_1^-(s)F_2^-(w).$$

The classical functions F_1^+ and F_1^- are derived from an approximation of the Figure 1.3a. Thus, for the two parameters τ^+ , $\tau^- \in \mathbb{R}^+$ these functions are given by

$$F_1^+(s) = F_1^+(0) \exp\left(\frac{-s}{\tau^+}\right) \quad \text{and} \quad F_1^-(s) = F_1^-(0) \exp\left(\frac{-s}{\tau^-}\right).$$

Note that the variable s represents the absolute value of Δt of Figure 1.3a. Other functions have been proposed according to experimental results, see for instance [2, 63]. Finally, depending on the functions F_2^+ and F_2^- different plasticity rules have been proposed:

- the additive STDP rule, see [79]:

$$F_2^+(w) \equiv \lambda \quad \text{and} \quad F_2^-(w) \equiv \lambda\alpha,$$

where λ is the learning rate and α is a parameter to create asymmetry between potentiation and depression,

- the multiplicative STDP rule, see [118]:

$$F_2^+(w) = \lambda(1 - w) \quad \text{and} \quad F_2^-(w) = \lambda\alpha w,$$

where $w \in [0, 1)$,

- the power law STDP rule, see [65]:

$$\mu > 0, \quad F_2^+(w) = \lambda w^\mu \quad \text{and} \quad F_2^-(w) = \lambda\alpha w.$$

STDP rules have the first three properties : locality, cooperativity and competition [120]. The weight boundedness is then ensured either by hard or soft bounds. In the so-called hard bound models, the weights are forced to be bounded. Models for which weights are bounded autonomously are called soft bound models, an example of which is the multiplicative STDP model.

These models are usually studied using a slow fast analysis as the weights dynamics is much slower than the neural dynamics. Hence, the assumption $\lambda \ll 1$ is often used meaning that small weight changes are assumed to occur at each spike pair. The neural dynamics are discussed in the next section but it is usually given by modelling spiking times arrival as random variables, see [79]. However, there are only few models introducing randomness into changes of synaptic weights in the STDP rule, see [9, 10, 60]. Yet, there are many sources of randomness in synaptic plasticity, from thermal noise to protein states. We will therefore subsequently look for a simple plasticity rule which is easy to implement in a neural network model and which assumes that the evolution of synaptic weights are random while respecting the results of neurobiologists.

1.2.3 Implementing Plasticity within Spiking Neural Network Models

There exist many single neuron models of the membrane potential that we list in chronological order: Integrate and Fire [84], Leaky Integrate and Fire [77], Hodgkin–Huxley [69], FitzHugh–Nagumo [43, 103], Morris–Lecar [101], Adaptive exponential integrate-and-fire [22]. This list is of course not exhaustive. In these models, the dynamics of the membrane potential V_t of a neuron is given by a differential equation. Sometimes, V_t is coupled to another variable which usually represents the environment of the neuron or some parts of it such as the states of its ion channels. The equation on the membrane potential always depends on the input current. It is the principal element containing neural interaction and also where the plasticity can potentially be implemented. We give a concrete example of a model of spiking neurons in interaction and show how plasticity could be implemented in it. In this example, the model described was proposed in [86, 109] and then a mathematical analysis of it was done in several studies, first with PDE methods [27] and then with probabilistic ones in [35, 38, 39]. Their common approach bears the name of mean field approximation. The basic idea of such an approach is detailed in the next section.

Example 1.1. *For simplicity and to make clearer the example, we do not include the noise term.*

From an external current $I_t : \mathbb{R}^+ \rightarrow \mathbb{R}$, a constant $R > 0$, a resting potential $V_{rest} \in \mathbb{R}$ and a threshold potential $V_{jump} > V_{rest}$, the Leaky Integrate and Fire (LIF) model is defined as:

$$V_0 = V_{rest} \quad \text{and} \quad \forall t \geq 0, \quad T \frac{dV_t}{dt} = -(V_t - V_{rest}) + RI_t$$

with V_t immediately reset to V_{rest} when it reaches V_{jump} .

From this simple model, one can derive the following neural network:

$$V_0^{i,N} = V_{rest} \quad \text{and} \quad \forall t \geq 0, \forall i \in \llbracket 1, N \rrbracket, \quad T \frac{dV_t^{i,N}}{dt} = -(V_t^{i,N} - V_{rest}) + R\tilde{I}_t^{i,N},$$

where $V_t^{i,N}$ is immediately reset to V_{rest} when it reaches V_{jump} and the synaptic current $\tilde{I}_t^{i,N}$ can be derived as

$$\tilde{I}_t^{i,N} = \frac{W}{N} \sum_j \sum_k \delta_0(t - T_k^j),$$

where T_k^j is the time of the k th spike of the neuron j and $W > 0$ is the interaction term (excitatory neural network interacting through a common weight). The membrane potential “kicks” generated by the spike of another neuron represent the EPSCs previously introduced in Section 1.1.2.

Implementing plasticity in this model can be done by making the weights dynamical and different one from each other: hence, W becomes $W_t \in \mathbb{R}_+^{N^2}$ where positivity means that we model a network of only excitatory neurons. Then, we have to specify the weight dynamics, see Section 1.2.2 for examples. Finally, the synaptic input is now given by:

$$\tilde{I}_t^{i,N} = \frac{1}{N} \sum_j \sum_k W_t^{ij} \delta_0(t - T_k^j).$$

We do not enter into more details because we do not use such models in the following. However, it is interesting to note that such models (the one without plasticity) can be approximated at the macroscopic level using probabilistic tools: at this level, the typical neuron dynamics lies on its spiking rate.

The mathematical analysis of neural networks models (in which the neurons are modelled by their membrane potential) is already complex. That is why simpler single neuron models are usually considered in neural network models with plasticity. For instance, we previously referred to rate based models. However, these models cannot take into account STDP as they are a valid description on the timescale of 100 milliseconds and more, see [79]. Hence, most of the models implementing STDP make the use of what is called the Poisson neuron model. This model gives the spiking times of a neuron as the jump times of a Poisson process, see Definition A.4 in the appendix. This model is convenient as it provides the spiking times necessary to use STDP rules. Initially, physicists linked this model of spiking neurons to STDP to rate based models of plasticity, see [79]. In this work, Kempter et al., using a timescale separation (plasticity much slower than the neural dynamics) on a STDP model, obtain a similar rate based model as the one proposed by Linsker in [88]. Nevertheless, they found that an additional term, due to the correlations of the precise spiking times, was present in the equation. Kempter et al. then studied these correlations and their effect on the spiking rate in the case of two input scenarios, see [78]. The main difficulty for extending this analysis to more general spiking neuron models is the computation of the correlations. Moreover, they modelled the dynamics of afferent synapses of one neuron (the output neuron) connected to N other ones (the input neurons). Hence, their study had to be extended to recurrent neural networks. It was done in several studies, see [25, 58, 126], using (as in Kempter et al.) Poisson neurons. These works were then extended in many ways: replacing the plasticity rule by a probabilistic one in [60], using LIF neurons in [107, 106] and adding intrinsic noise to the synapses in [122]. This list is not exhaustive but they all use the simplification obtained by Kempter et al.. None of them get an exact limit dynamics from the timescale separation as they compute the correlations matrix of the spiking times with the approximation method first proposed in [87], see [126] for further details on this method. Hence, there does not exist rigorous mathematical analysis of neural networks with STDP. In Chapters 3 and 4, we propose such an analysis implementing STDP within a neural network model closed to the Poisson neuron model and introduced in [15].

1.2.4 Mathematical methods: slow fast analysis and mean field approximation

We previously referred to some mathematical methods to analyse models either using a separation of timescale (slow-fast analysis) or the increasing number of neurons (mean field approximation). In this section, we detail their main idea and give examples to illustrate them.

Slow-fast analysis

We use slow-fast methods when we can make the assumption of a large difference between the timescales of two or more variables. When such an assumption is justified (by experiments for instance), we propose a new model in which this difference appears clearly in the new dynamical equations. These equations are usually parametrised by a small variable ε . In this new framework, when ε tends to zero, on the fast timescale, the fast variable “sees” the slow variable as being constant. On the slow timescale, the slow variable “sees” the fast variable under its stationary state (if it exists). Therefore, when it exists, the limit of this new dynamical model when ε tends to zero is then determined: it gives the dynamics of the slow variable in function of the stationary behaviour of the fast one. Usually, the limit system is simpler than the initial one and thus easier to handle with. Its analysis can bring new insights in the initial model. We give a really simple example of such a timescale separation.

Example 1.2. We take an example presented in [51] (Example 2.2). Let consider two coupled variables $(v_t^\varepsilon, w_t^\varepsilon)_{t \geq 0} \in \mathbb{R}^2$, a Brownian motion $(B_t)_{t \geq 0}$ and parameters $\varepsilon > 0$, $\sigma \in \mathbb{R}^+$ such that

$$dv_t^\varepsilon = -\frac{1}{\varepsilon}v_t^\varepsilon dt + \sigma \frac{1}{\sqrt{\varepsilon}}dB_t \quad \text{and} \quad dw_t^\varepsilon = (v_t^\varepsilon)^2 dt - w_t^\varepsilon dt. \quad (1.1)$$

Thus, when ε tends to zero, w_t^ε can be approximated by w_t , the solution of

$$\frac{dw_t}{dt} = \int_{\mathbb{R}} (v^2 - w_t) \rho(dv),$$

where ρ is the invariant measure of the process $(v_t^\varepsilon)_{t \geq 0}$. Note that this measure does not depend on ε . Indeed, changing the time scale in equation (1.1) does not modify its invariant measure and thus, on the timescale $u = \frac{t}{\varepsilon}$, we observe that ρ is also the invariant measure of the process $(v_u)_{u \geq 0}$ such that

$$dv_u = -v_u du + \sigma dB_u.$$

The term ‘‘approximated’’ we used here holds in the following sense:

$$\forall \delta > 0, \forall T > 0, \quad \lim_{\varepsilon \rightarrow 0} \mathbb{P} \left(\sup_{t \in [0, T]} \|w_t^\varepsilon - w_t\| > \delta \right) = 0.$$

In this example, the limit system is deterministic but it could be stochastic, see the results of Chapter 3, Theorem 3.16 for instance.

Mean field analysis

The mean field method consists in interpreting a dynamical system from another point of view, using some ‘‘average’’ behaviour. Like the separation of timescale, it addresses the issue of scale change. In particular, its aim is to bridge the gap between the microscopic scale and the macroscopic one. Here, the scale is determined by the size of the microscopic system: to fix the ideas, we consider a system with N neurons. As the number of neurons tends to infinity, if the global effect of the interaction remains finite, one can apply a mean field method and obtain a simplified model (compared to the finite size neural network when N is large) describing the dynamics of the neurons’ statistics. We provide an example which gives a first introduction to the models used in Chapters 3 and 4.

Example 1.3. We model a neural network of N neurons, described by the vector of their states $V_t^N \in \{0, 1\}^N$, and interacting by the mean of one single weight $W \in \mathbb{R}$. The dynamics is summarized by the following reaction:

$$\forall t \geq 0, \forall i \in \llbracket 1, N \rrbracket, \quad 0 \xrightarrow[\beta]{\alpha\left(\frac{W}{N} \sum_i V_t^{i, N}\right)} 1,$$

where α is a smooth function from \mathbb{R}^+ to $[\alpha_m, \alpha_M]$ with $0 < \alpha_m \leq \alpha_M < \infty$. In order to study this model, we denote by $\bar{V}_t^N \stackrel{\text{def.}}{=} \frac{\sum_i V_t^{i, N}}{N}$ the empirical mean of the sequence $(V_t^{i, N})_{1 \leq i \leq N}$. In particular, note that $\alpha\left(\frac{W}{N} \sum_i V_t^{i, N}\right) = \alpha(W\bar{V}_t^N)$. Then, using the results of the paper of Kang and Kurtz [75], it can be shown that

$$\forall \delta > 0, \forall T > 0, \quad \lim_{N \rightarrow +\infty} \mathbb{P} \left(\sup_{t \leq T} |\bar{V}_t^N - \bar{V}_t| > \delta \right) = 0,$$

where $\bar{V}_0 = \lim_{N \rightarrow +\infty} \bar{V}_0^N$ and satisfies the ODE

$$\frac{d\bar{V}_t}{dt} = (1 - \bar{V}_t)\alpha(W\bar{V}_t) - \beta\bar{V}_t.$$

This model is very rich when splitting the neurons in two subgroups, an excitatory group and an inhibitory one, giving rise to the use of four typical weights, see [15] for more details.

1.3 A new model of STDP within a neural network (Chapters 3 and 4)

A large amount of studies have been focused on neural networks dynamics with the aim of reproducing biological phenomena. Thereby, there exist many different individual neuron models from the binary neurons to the adaptive exponential integrate-and-fire, see [56, 72]. As developed in Section 1.2.3, plasticity can be implemented within these models. Over the last few decades, such models have been proposed with goal to explain the experimental results as well as going beyond them. We saw in Section 1.2.3 that there was no rigorous mathematical analysis of models in which the neural dynamics interplays with spike timing dependent plasticity. Moreover, although there are a lot of deterministic studies in neural network modelling, with or without plasticity and mainly using dynamical systems theory, we find much less probabilistic analyses while the brain is far more stochastic than deterministic.

This is changing thanks to the development of numerous probabilistic models. For example, Robert and Vignoud recently worked on a rigorous mathematical analysis of numerous plasticity models belonging to a general class of models that they introduced and on which they found limit dynamics using a separation of timescale method, see [114, 115]. While they analysed the precise synaptic weight between an input and an output neuron, we propose in this thesis a new model of STDP involving a neural network interacting through dynamical synaptic weights. First, we precise our motivations, going beyond the need of a mathematical understanding of plastic neural networks, while presenting our new model. Then, we give the main results of Chapters 3 and 4.

1.3.1 Motivations for a new model and its description

Motivations

Motivations for proposing such a new model are four folds. First, although the mechanisms involved in plasticity are mainly stochastic (for instance the opening of ion channels or activation of proteins, see [74]), the majority of studies on STDP are implemented using a deterministic description or an extrinsic noise source [102, 31, 63]. One exception is the stochastic STDP model proposed by Appleby and Elliott in [9, 10]. The stochasticity of their model lies in the learning window size. They analysed the dynamics of the weights of one target cell innervated by a few Poisson neurons. They performed a fixed point analysis on the dynamics of the weights. This analysis enabled them to show that multispike interaction (more than two-spike interaction) are required to get stable competitive weight dynamics. Instead of the window size, we propose to introduce stochasticity within the STDP curve. Indeed, in the first experiment providing a relation between spikes' pair and the strength of plasticity, see [17], the results does not seem deterministic

but rather probabilistic. Hence, instead of the classical exponential STDP curves combined with a small increment of the weights each time a spike occurs, we use a probabilistic increment: we change the weights linked to the spiking neuron according to Bernoulli variables with parameters given by two functions p^+ and p^- that depend on the absolute value of the time between the spike pair considered and the actual weight, thus $p^+, p^- : \mathbb{R}^+ \times \mathbb{R} \rightarrow [0, 1]$. Thus, it enables us to be close to biological experiments [17], see Figure 1.5.

Secondly, most studies are based on numerical analyses. Thus, there is still a need to find a good mathematical framework, see [52, 89, 107]. Despite previous attempts, rigorous results in plastic neural network are still lacking. In particular, rigorous timescale separation under the experimentally supported assumption that synaptic plasticity is much slower than neural dynamics and a mean field analysis of such models does not exist yet. In order to tackle these problems, we opted for a simple neural dynamics with binary neurons, the probabilistic Wilson-Cowan model [15], in which we implemented a stochastic STDP rule, see below for the presentation of this rule.

Thirdly, long term plasticity timescale ranges from a few minutes to more than an hour. On the other hand, a spike lasts a few milliseconds [102]. Thus, there is a need to understand how to bridge this timescale gap between the synapses level and the network one [45, 136, 128]. Although this timescale difference has already been exploited in other studies, the long term behaviour of the limit model remains unclear. The assumption of discrete weights and the probabilistic framework leads us to a simple Markov chain as limit model.

Finally, the interplay between the dynamics of the weights and the neural dynamics is not yet fully understood. We think that the mathematical study of biologically plausible model is necessary to bring the basis of this understanding. The Wilson-Cowan model has been widely studied [21, 15, 89] and reproduces many biological features of a network such as oscillations and bi-stability. Thus, we obtain a model rich enough to reproduce biological phenomena, simple enough to be mathematically tractable and easy to be simulated with thousands of neurons.

The new model

We now detail the model more precisely. We study a neural network of N binary neurons, $V_t^N \in \{0, 1\}^N$, connected by weights, W_t^N , living on a grid of size Δw . We obtain a Markov process by adding the time from the last spikes of neurons, $S_t^N \in \mathbb{R}_+^N$. Therefore, we consider the Markov process $(V_t^N, S_t^N, W_t^N)_{t \geq 0}$. The model of individual neuron is a simple binary chain: the state $V_t^{i,N}$ of the neuron i jumps from 0 to 1 at rate $\alpha_i(V_t^N, W_t^N)$ and jumps from 1 to 0 at constant rate β . A spike occurs when one of the components of V_t^N jumps from 0 to 1. When a neuron spikes, say neuron j_0 , its post- and presynaptic weights can be changed according to a probability depending on S_t^N and W_t^N : $W_t^{j_0 k, N}$ is potentiated (incremented of Δw) with probability $p^+(S_t^{k, N}, W_t^{j_0 k, N})$ and $W_t^{k j_0, N}$ is depressed (decremented of Δw) with probability $p^-(S_t^{k, N}, W_t^{k j_0, N})$. This dynamics is represented in Figure 1.4. In Figure 1.4a, the dynamics of one neuron is plotted. In Figure 1.4b, the dynamics of the time from the last spike of this neuron is added to the previous plot. Finally, in Figure 1.4c, the dynamics of the weight linking this neuron to another one is added to the previous plot.

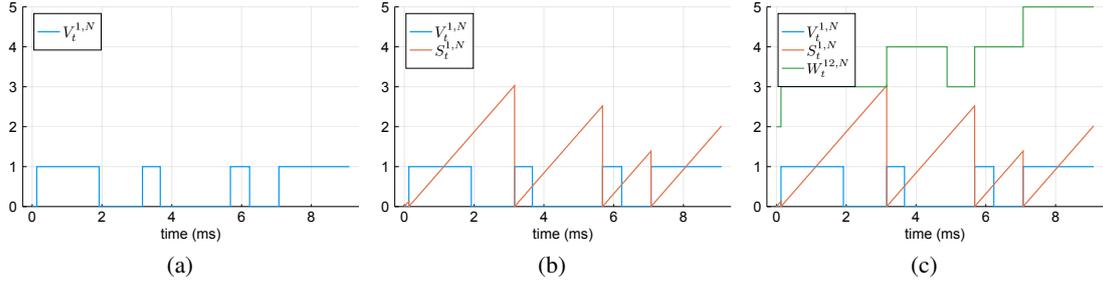


Figure 1.4: (1.4a): Typical trajectory of $(V_t^{1,N})_{t \geq 0}$. (1.4b): Typical trajectory of $(V_t^{1,N}, S_t^{1,N})_{t \geq 0}$. (1.4c): Typical trajectory of $(V_t^{1,N}, S_t^{1,N}, W_t^{12,N})_{t \geq 0}$.

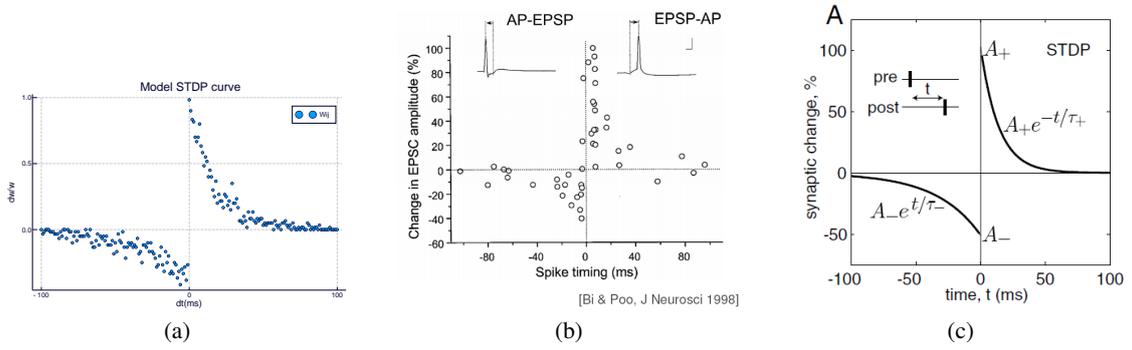


Figure 1.5: (1.5a) Mean change of weights in the new model after 60 pairs of spikes with the functions $p^+(s, w) = A_+ e^{\frac{-s}{\tau_+}}$ and $p^-(s, w) = A_- e^{\frac{-s}{\tau_-}}$ where $A_+ = 1, A_- = 0.4, \tau_- = 34\text{ms}, \tau_+ = 17\text{ms}$ as in [61]. (1.5b) Bi-Poo experiment results obtained in [17]. (1.5c) Classical STDP curve in deterministic STDP modelling.

1.3.2 Slow-Fast Analysis

In Chapter 3, we study the slow-fast phenomenon. The number N of neurons is fixed. To alleviate the notations, we do not keep it in the notation. Thus, we study the Markov process $(V_t, S_t, W_t)_{t \geq 0}$ in the context of long term plasticity: synaptic weight dynamics is much slower than the neural network one. In previous works, see [102], the timescale separation assumption is formally implemented as follows: at each spike pair, every weight linking the spiking neurons is increased or decreased of a value proportional to the small scaling parameter ε that tends to zero. Hence, several spike pairs are necessary to observe a macroscopic change of the weights. We propose here a different approach: at each spike pair, every weight linking the spiking neurons is increased or decreased of a macroscopic value with a certain probability tending to 0 with ε . This approach can be discussed but some studies have shown that it would be more realistic as synapses have a finite number of states and thus should be discrete and not continuous, see [108]. Our scaling parameter thus influences the jumping probability of weights at each spike pair. The slow-fast analysis then consists in firstly studying the fast process and its long time behaviour and secondly describing the slow dynamics on its own timescale. Hence, we first consider the process $(V_t, S_t | W_t \equiv w)_{t \geq 0}$ (fast process) and show that it converges to a unique invariant measure that we denote π_w . Then, we define the scaled process $(V_t^\varepsilon, S_t^\varepsilon, W_t^\varepsilon)_{t \geq 0}$ indexed by a small parameter $\varepsilon \in (0, 1)$ such that the jumping probability of the weights tends to 0 with ε . Note that the invariant measure of the fast process $(V_t^\varepsilon, S_t^\varepsilon | W_t^\varepsilon \equiv w)_{t \geq 0}$ is still π_w for all ε . We obtain the dynamics of the weights

depending on this invariant measure in the limit ε tends to zero. Finally, we study the long time behaviour of this asymptotic dynamics.

Main results (informal)

We assume that $\Delta w = 1$. For every $\varepsilon > 0$, the process $(W_t^\varepsilon)_{t \geq 0}$ takes values in the space

$$E_2 \stackrel{\text{def.}}{=} \{w \in \mathbb{N}_*^{N^2} : \forall i, w^{ii} = 0\}.$$

When the neuron i spikes, say at time t , we denote by $\phi_i^\varepsilon : \mathbb{R}_+^N \times E_2 \times E_2 \rightarrow [0, 1]$ the transition probability of the weights conditionally to the times from the last spikes S_{t-}^ε , see Assumptions 3.9 and 3.10. In particular, we show in Proposition 3.11 that assuming one of these two assumptions leads to the same form of transition probabilities: for all i , there exist bounded functions φ_i such that

$$\phi_i^\varepsilon = \varepsilon \varphi_i + \mathcal{O}(\varepsilon).$$

Then, Theorem 3.16 gives the limit dynamics of $(W_t^\varepsilon)_{t \geq 0}$ when ε tends to zero. This process converges to the Markov chain $(\bar{W}_t)_{t \geq 0} \in E_2$ with transition rate function r such that for all $w, \tilde{w} \in E_2, w \neq \tilde{w}$,

$$r(\tilde{w}, w) = \sum_i \int_{\{0,1\}^N \times \mathbb{R}_+^N} \alpha_i(v, w) \delta_0(v^i) \varphi_i(s, \tilde{w}, w) \pi_w(dv, ds)$$

gives the jump rate from w to \tilde{w} . Hence, the generator \mathcal{C}_{av} of $(\bar{W}_t)_{t \geq 0}$ is such that for all f in its domain and $w \in E_2$,

$$\mathcal{C}_{av}f(w) = \sum_{\tilde{w} \in E_2 \setminus \{w\}} (f(\tilde{w}) - f(w)) r(\tilde{w}, w).$$

Therefore, when ε tends to 0, the process $(W_t^\varepsilon)_{t \geq 0}$ converges to the process $(\bar{W}_t)_{t \geq 0}$ which is a Markov chain on E_2 whose transition probabilities depends on the invariant measure of the fast process.

Once this result is shown, we continue with the study of the long time behaviour of $(\bar{W}_t)_{t \geq 0}$. To do so, we use some properties of the functions φ_i . In particular, we have the following relationship:

$$\forall j \neq i, \quad \sum_{\tilde{w}^{ij} = w^{ij} + 1} \varphi_i(s, \tilde{w}, w) = p^+(s^j, w^{ij}) \quad \text{and} \quad \sum_{\tilde{w}^{ji} = w^{ji} - 1} \varphi_i(s, \tilde{w}, w) = p^-(s^j, w^{ji}).$$

Hence, using the following notations,

$$\begin{aligned} \forall w \in E_2, \quad r_{ij}^+(w) &\stackrel{\text{def.}}{=} \sum_{\tilde{w}^{ij} = w^{ij} + 1} r(\tilde{w}, w) = \int_{\{0,1\}^N \times \mathbb{R}_+^N} \alpha_i(v, w) \delta_0(v^i) p^+(s^j, w^{ij}) \pi_w(dv, ds), \\ r_{ij}^-(w) &\stackrel{\text{def.}}{=} \sum_{\tilde{w}^{ij} = w^{ij} - 1} r(\tilde{w}, w) = \int_{\{0,1\}^N \times \mathbb{R}_+^N} \alpha_j(v, w) \delta_0(v^j) p^-(s^i, w^{ij}) \pi_w(dv, ds), \end{aligned}$$

we obtain a criterion on r_{ij}^+ and r_{ij}^- for transience and positive recurrence of the chain $(\bar{W}_t)_{t \geq 0}$, see Proposition 3.24. Let us enlighten that if p^+ and p^- are separable functions (product of a function depending only on w and a function depending only on s) with exponential dependence on s , the Laplace transform of π_w is natural to use in the definition of r_{ij}^+ and r_{ij}^- . More generally, assume that p^+ and p^- are separable functions with functions on s that can be written as Laplace transforms of some functions $P^{+/-}$:

$$p^{+/-}(s, w) = p_1^{+/-}(s) p_2^{+/-}(w) \quad \text{where} \quad p_1^{+/-}(s) = \mathcal{L}\{P^{+/-}\}(s).$$

Then, we can transfer this Laplace transform to the one of π_w in the transition rates, see Remark 3.35. Therefore, we compute the Laplace transform of π_w to study the process $(\overline{W}_t)_{t \geq 0}$. Then we obtain a criterion based on exact computations that we did at the end of the chapter on an example. Surprisingly, we observe transience even when $p^+ < p^-$ and more importantly, simulations showed that the increasing size of the network seems to prevent divergence of the weights under weaker assumption than the one needed in the case of only two neurons.

1.3.3 Mean Field Analysis

In Chapter 4, we study the large N asymptotic behaviour of the Markov process $(V_t^N, S_t^N, W_t^N)_{t \geq 0}$ where for all i , $\alpha_i(V_t^N, W_t^N) = \alpha(\frac{1}{N} \sum_j W_t^{ij,N} V_t^{j,N})$ with α a strictly positive and bounded function. In particular, we are interested in the dynamics of *new* variables $X_t^{1,N}, \dots, X_t^{N,N}$ describing the neural network such that for all i ,

$$\xi_t^{i,N} = \frac{1}{N} \sum_j \delta_{(V_t^{j,N}, S_t^{j,N}, W_t^{ij,N})} \in \left\{ \text{probability measures on } \underbrace{\{0, 1\} \times \mathbb{R}^+ \times \mathbb{Z}}_{\stackrel{\text{def.}}{=} E_m} \right\} \stackrel{\text{def.}}{=} \mathcal{P}(E_m)$$

is an empirical measure and $X_t^{i,N} = (V_t^{i,N}, S_t^{i,N}, \xi_t^{i,N}) \in E = \{0, 1\} \times \mathbb{R}^+ \times \mathcal{P}(E_m)$. The empirical measure obtained from these variables is denoted by μ_t^N . Finally, we denote by I the function from the space of the measures on E_m to \mathbb{R}^+ such that for such a measure ξ ,

$$I(\xi) = \int w v \xi(dv, ds, dw).$$

In particular, $I(\xi_t^{i,N})$ is the average synaptic current onto the neuron i at time t .

Main results (informal)

Assume that for all $T > 0$, in the large N asymptotic, the empirical measures $(\mu^N)_{N \in \mathbb{N}^*}$ on $\mathcal{P}(D_E[0, T])$ (càdlàg functions from $[0, T]$ to E , see Definition 3.15) converge in law (see Definition A.1 in the appendix) to a deterministic measure $\mu^* \in P(D_E[0, T])$. Moreover, assume that for all $t \geq 0$, μ_t^* and ξ_t^* admit densities of class C^1 in s and in particular at time $t = 0$, for all $w \in \mathbb{Z}$, the following densities satisfy the boundary conditions for μ_0^* almost all ξ_0^* ,

$$\begin{aligned} \xi_0^{*0}(0, \{w\}) &= 0 \\ \xi_0^{*1}(0, \{w\}) &= \int_{\mathbb{R}^+ \times \mathcal{P}(E_m)} \alpha(I(\xi')) \frac{\xi_0^{*0}(s', \{w\})}{\xi_0^{*0}(s', \mathbb{Z})} \mu_0^*(\{0\}, s', d\xi') ds', \end{aligned}$$

where we use the abuse of notation

$$\mu_t^*(\{v\}, ds, d\xi) = \mu_t^*(\{v\}, s, d\xi) ds \quad \text{and} \quad \xi_t^{*v}(ds, \{w\}) = \xi_t^*(\{v\}, s, \{w\}) ds.$$

Then, we conjecture (in **Conjecture 4.18**) that a typical neuron $X_t^* = (V_t^*, S_t^*, \xi_t^*)$, with law μ_t^* , satisfies the following McKean-Vlasov stochastic differential equation (SDE):

1. $dS_t^* = dt,$

2. ξ_t^* admits a density in s such that

$$\begin{cases} \frac{d\xi_t^{*0}}{dt}(s, \{w\}) = -\partial_s \xi_t^{*0}(s, \{w\}) + \beta \xi_t^{*1}(s, \{w\}) - \int_{\mathcal{P}(E_m)} \alpha(I(\xi')) \frac{\xi_t^{*0}(s, \{w\})}{\xi_t^{*0}(s, \mathbb{Z})} \mu_t^*(\{0\}, s, d\xi') \\ \xi_t^{*0}(0, \{w\}) = 0, \end{cases}$$

$$\begin{cases} \frac{d\xi_t^{*1}}{dt}(s, \{w\}) = -\partial_s \xi_t^{*1}(s, \{w\}) - \beta \xi_t^{*1}(s, \{w\}) \\ \xi_t^{*1}(0, \{w\}) = \int_{\mathbb{R}^+ \times \mathcal{P}(E_m)} \alpha(I(\xi')) \frac{p^-(S_t^*, w) \xi_t^{*0}(s', \{w+1\}) + (1-p^-(S_t^*, w)) \xi_t^{*0}(s', \{w\})}{\xi_t^{*0}(s', \mathbb{Z})} \mu_t^*(\{0\}, s', d\xi') ds', \end{cases}$$

3. At rate $\beta \mathbb{1}_{\{V_{t^-}^* = 1\}}$, $(V_{t^-}^*, S_{t^-}^*, \xi_{t^-}^*)$ jumps to $(0, S_{t^-}^*, \xi_{t^-}^*)$,

4. At rate $\alpha(I(\xi_{t^-}^*)) \mathbb{1}_{\{V_{t^-}^* = 0\}}$, $(V_{t^-}^*, S_{t^-}^*, \xi_{t^-}^*)$ jumps to $(1, 0, \nu^+(\xi_{t^-}^*))$ where

$$\begin{aligned} \nu^+(\xi_{t^-}^*)(\{v\}, A, \{w\}) &= \int_A p^+(s, w-1) \xi_{t^-}^*(\{v\}, s, \{w-1\}) ds \\ &\quad + \int_A (1-p^+(s, w)) \xi_{t^-}^*(\{v\}, s, \{w\}) ds \end{aligned}$$

for all $(\{v\}, A, \{w\}) \in \mathcal{B}(E_m)$.

We give some details on this limit system. The first and second points are the drift terms and both last points are the jump terms. The drift on ξ_t^* is composed of three parts. First, the linear increase of s which is the term in $-\partial_s$. Second, the mass transport from ξ_t^{*1} to ξ_t^{*0} at rate β . Finally, the mass transport from ξ_t^{*0} to ξ_t^{*1} at a rate depending on μ_t^* with the reset of s in 0 giving the boundary condition in $s = 0$ for ξ_t^{*1} .

1.4 Plasticity and memory (Chapter 2)

The role played by synaptic plasticity in the processes of memory formation and maintenance has been shown in many experiments from several years now, see [99] and [111]. However, synaptic plasticity does not seem to be the only mechanism responsible for memory. Some other mechanisms are put forward such as intrinsic neuron mechanisms, see [125], or stabilization of the memory at the network level, see [100]. The idea of the network level memory is implemented in the Hopfield model in which memories are attracting states of the neural dynamics: these states are recurrently visited and by this way, their trace is enhanced. After a presentation of various models of memory taking roots in the Hopfield model [70], we give more details on the model we use in Chapter 2 and our results.

1.4.1 Network Models of Memory

One of the most important question in neuroscience is about the neural code. How does the brain represent, store and then restore signals? Since the discovery of synaptic plasticity and its implication in memory, many memory models have been proposed. A lot of work has been done from the proposal of what is called the Hopfield model [70] in 1982. Hopfield gave rise to this

model by applying the networks of threshold automata (NTA) to memory problems. There exist many variants of this model and more importantly, it is at the basis of many researches on the topic of memory retrieving and memory capacity. This research diverged in two different branches. In the first one, models are based on a fixed synaptic weight matrix in which a given number of stimuli are stored, see [133, 70, 54, 127]. The second branch, see [49, 85, 48, 129, 42, 117, 16], followed the work of Amit and Fusi [5] in which they proposed a model to study the memory capacity of a plastic neural network. We give the main ideas of these two branches.

ANN models

One of the most studied aspects of memory is the associative memory: capacity to recall an information when it is presented again entirely or partially. We can refer to the Pavlovian reflex as a famous example. Associative memory is usually modelled by attractor neural networks (ANN) in which the signals (or stimuli) to learn generate neural activity patterns. These signals are effectively learnt as soon as their corresponding neural activity patterns are stable fixed points of the neural network dynamics. This dynamics is governed by a synaptic weight matrix. Thus, the information is stored in this matrix. The function of the patterns which determines this matrix is the main difference between the various ANN models. Among the most famous, we can quote the models of Willshaw [133], Hopfield [70] and Amari [3]. The question that naturally arises is: how many patterns can be stored by such models? Physicists as well as mathematicians studied extensively this problem, especially on the Hopfield model. A capacity of $0,14N$ (where N is the size of the network) was put forward by physicists in [6] whereas rigorous results were proved on the lower bound of this capacity: it has been successively proved to be superior to $0,056N$ in [104], $0,071N$ in [90] and $0,084N$ in [124]. Much better results have been found in the sparse neural coding case: a small proportion of neurons are activated by external signals. In this case, the memory capacity has been shown to scale as $\frac{N^2}{(\log N)^2}$, see [64] and references therein. Similar results have been found in models with plasticity.

SNR analysis of Plastic Models of Memory

How to simply model memory formation and how to test it? In the second branch of memory models, the general framework consists in considering a sequence of independent random signals presented to the network thereby changing its structure such that at least the last signals are remembered. Plasticity rules are responsible for modifying the structure and signal theory methods are used to test the quality of the learning. The retrieval criteria proposed in [5, 49, 48, 117, 16] are similar and based on a signal to noise ratio (SNR) analysis. There is no neural dynamics in these models and their differences lie in the weight dynamics that we detail in the following paragraph. In order to compute the SNR, a memory signal of a stored memory is defined. We assume that a sequence of memories are stored at different times that we denote by $(t^l)_{l \in \mathbb{N}}$. Let us consider the memory stored at time t^k . It is the overlap between the actual weights, noted $(w_t^j)_{1 \leq j \leq N}$, and the possible weight changes induced by the k_{th} signal, noted $(\Delta w_{t^k}^j)_{1 \leq j \leq N}$. For all $t \geq t^k$, we denote it by $Z_{t,k}$:

$$Z_{t,k} = \frac{1}{N} \sum_{j=1}^N w_t^j \Delta w_{t^k}^j. \quad (1.2)$$

This signal is noisy as the presented stimuli are random as well as the weight dynamics. The SNR is then the expectation of the signal divided by its standard deviation. In this framework, the SNR

is a quantity that theoretically measures the trace of a signal. Here, by theoretically we mean that there is no direct link between this SNR and any specific retrieval process. The only test that could be done on this quantity is to compare it to its asymptotic value in time. We could then for instance use a threshold estimator of the signal trace: when the SNR is above the threshold, the signal is still in memory, and below this threshold, the signal is not in memory anymore.

Now, we deal with the differences between these models. First, except from [5] where the synaptic weights depend on the neural interpretation of the signal, in other models, there is no neural network considered and only N weights are modelled. The binary synapse models give poor capacity results except when the coding level (probability that a neuron is selective to a signal) is of the order of $\frac{\log(N)}{N}$ (sparse coding), see [5]. Moreover, this capacity can be improved, without sparse coding, either by adding internal states to the weights (which can be called metaplasticity) see [49], or by partitioning the synapses into subgroups with different levels of plasticity, see [117]. However, considering multi-states synapses does not seem to give better results than the binary case, except under really specific conditions on the parameters, see [48]. Finally, in [16], defining the weight dynamics using a decay kernel, they found much more efficient models. There is a general review on these models proposed in [47].

1.4.2 Model and Results

Motivations

First, the SNR is defined in different ways depending on the application. The definition used here refers to the one used in signal detection theory. The idea behind this quantity is to measure the ability to separate the signal from the surrounding noise. In general, it is linked to the probability of error in distinguishing the signal from the noise. In the current framework, it is related to the probability of answering correctly to the question: have you ever seen this signal before? The principal advantage of the SNR is to be easily tractable and then adapted to a rigorous mathematical analysis, see [82] for such an analysis in the case of only synaptic weight dynamics.

However, it does not provide a way of retrieving the signal (or the pattern activity generated by the signal) as the ANN models do. The SNR analysis is not the only possibility to get information on the memory of the signal left. Several other methods have been proposed based on global variables in order to compute the memory capacity, see [85] and [42] for example. Similarly to the SNR, they do not give precise information on the memory retrieval.

The first study providing such an analysis is [8] by extending the model of [5]. Using a similar idea as ANN, they computed what they called a retrieval probability. Then, a precise analysis of the same model has been done in [41] in the large N asymptotic. In order to compute the probability of error, both [8] and [41] used strong assumptions. In [8] they used Gaussian assumption and independence between the synaptic currents. In [41] they assumed that the synapses are independent. Therefore, there does not exist rigorous analysis of the memory lifetime in the Amit-Fusi model. Moreover, in the SNR analysis previously proposed, the asymptotic value in time of the standard deviation of the memory signal is used instead of the time dependent standard deviation. However, it is not obvious that for any time this standard deviation is of the same order (in N) as its asymptotic (in time). Finally, the SNR is based on the synaptic current received by only one neuron and then get rid of the difficulty generated by correlations.

In Chapter 2, we extend the previous mathematical results on the Amit-Fusi model on many points:

we compute the exact variances of the synaptic currents, we find the spectrum of the transition matrix of the synaptic current and of the weight matrix and finally, we give a lower bound on the time that a simple threshold estimator of the neural activity spend below a given probability of error.

Model and Results

We present here the model and, in an informal way, the results obtained in Chapter 2. The model is designed according to a learning protocol. The two components of this model are the neural network of N neurons, described by the neuron potentials $(V_t^N)_{t \in \mathbb{Z}} \in \{0, 1\}^N$, and the weight matrix without the diagonal terms, $(W_t^N)_{t \in \mathbb{Z}} \in \{0, 1\}^{N(N-1)}$. The time is discrete as the variables evolve upon the presentation of a sequence of external signals. The signals or stimuli (we use both terms in the following) modify the neural network. This change leads to probabilistic modifications of the synaptic weights.

The external signals are assumed to be random, independent and identically distributed (*i.i.d.*) as well as their neural responses $(V_t^N)_{t \in \mathbb{Z}}$. We denote by $f_N \in (0, 1)$ the coding level such that

$$\forall i \in \llbracket 1, N \rrbracket, \quad \mathbb{P}\left(V_t^{i,N} = 1\right) = 1 - \mathbb{P}\left(V_t^{i,N} = 0\right) = f_N.$$

Knowing V_t^N and W_t^N we have for all $i \neq j$:

- if $(V_t^{i,N}, V_t^{j,N}, W_t^{ij,N}) = (1, 1, 0)$, $\mathbb{P}\left(W_{t+1}^{ij,N} = 1\right) = q^+$,
- if $(V_t^{i,N}, V_t^{j,N}, W_t^{ij,N}) = (1, 0, 1)$ (resp. $(0, 1, 1)$), $\mathbb{P}\left(W_{t+1}^{ij,N} = 0\right) = q_{10}^{-,N}$ (resp. $q_{01}^{-,N}$).

Adding the assumption that at time $t = 0$, the network has reached its stable state (in law), ends the description of the Amit-Fusi model introduced in [5].

In this framework, we study how to measure the ability of this model to keep in memory a given signal. We propose to change the protocol in order to test the effect of multiple presentation of a given signal to learn. Thus, we denote by $r \in \mathbb{N}^*$ the number of times this signal is presented and $V_0^N \in \{0, 1\}^N$ its neural response. We assume that (V_{-r}^N, W_{-r}^N) is in its stable state (in law) and that $V_t^N = V_0^N$ for $t \in \{-r + 1, \dots, 0\}$.

Our results hold on the synaptic currents onto the neurons. Indeed, in order to measure the ability of the network to remember the signal to learn, we compute the estimation of its neural response at time $t > 0$, that is to say after the presentation of some other signals that are noise for the one to learn. This estimation is based on the synaptic currents that would be obtained if the signal to learn is presented again at time $t > 0$. We denote by $(h_t^N)_{t \in \mathbb{Z}}$ such a sequence of synaptic currents onto the neuron 1:

$$h_t^N = \sum_{j=2}^N W_t^{1j,N} V_0^{j,N}.$$

Denoting by $(h_{t,K}^N)_{t \in \mathbb{N}^*}$ the process such that $h_{t,K}^N = h_t^N$ knowing that $\sum_{j=2}^N V_0^{j,N} = K$, we show in Chapter 2 that this process is Markov and satisfies

Proposition 1.4. (Proposition 2.12)

The spectrum of the transition matrix P_K^N of $(h_{t,K}^N)_{t \geq 1}$ is

$$\Sigma(P_K^N) = \{(1 - f_N)(\Lambda_0^N)^i + f_N(\Lambda_1^N)^i, 0 \leq i \leq K\},$$

where $\Lambda_0^N = 1 - f_N q_{01}^{-,N}$ and $\Lambda_1^N = 1 - (1 - f_N)q_{10}^{-,N} - f_N q^+$.

Moreover, we show that

Proposition 1.5. (Proposition A.11)

The spectrum of the transition matrix $M_{V,W}^N$ of $(V_t^N, (W_t^{1j,N})_{2 \leq j \leq N})_{t \geq 1}$ is

$$\Sigma(M_{V,W}^N) = \{\lambda_i^N, 0 \leq i \leq N\} \cup \{0\},$$

where λ_i^N has multiplicity $\binom{N-1}{i}$ and 0 has multiplicity $2^{N-1}(2^N - 1)$.

For $y \in \{0, 1\}$, denoting by $(h_t^{y,N})_{t \in \mathbb{N}^*}$ the process such that $h_t^{y,N} = h_t^N$ knowing that $V_0^{1,N} = y$, we obtain

Proposition 1.6. (Proposition 2.15)

The expectations and variances of $h_t^{0,N}$ and $h_t^{1,N}$ satisfy:

$$\mathbb{E}[h_t^{0,N}] = N f_N (\mu^{*,N} - \mu^{0,N} (\lambda_1^N)^{t-1}), \quad \text{with } \mu^{0,N} = \left(1 - (1 - q_{01}^{-,N})^r\right) \frac{f_N^2 q^+}{1 - \lambda_1^N},$$

$$\mathbb{E}[h_t^{1,N}] = N f_N (\mu^{*,N} + \mu^{1,N} (\lambda_1^N)^{t-1}), \quad \text{with } \mu^{1,N} = (1 - (1 - q^+)^r) \left(1 - \frac{f_N^2 q^+}{1 - \lambda_1^N}\right),$$

$$\text{Var}[h_t^{y,N}]^2 = \mathbb{E}[h_t^{y,N}] (1 - \mathbb{E}[h_t^{y,N}]) + f_N^2 N(N-1) \left(C_0^{*,N} + C_1^{y,N} (\lambda_1^N)^{t-1} + C_2^{y,N} (\lambda_2^N)^{t-1}\right),$$

where $C_0^{*,N}$, $C_1^{y,N}$ and $C_2^{y,N}$ are constants depending on the parameters and $\mu^{*,N} = \frac{f_N^2 q^+}{1 - \lambda_1^N}$.

Finally, the main results and contributions of this Chapter is the following.

Main Results - Contributions (informal)

For any fixed error $\epsilon \in (0, 1)$, there is an unbounded set of couples $(N, r) \in \mathbb{N}^{*2}$ for which we control the time that the signal estimation error spends above ϵ , see Theorem 2.27. An explicit formula of this time is given in Remark 2.28 for fixed potentiation and depression probabilities. Another formula of this time is given in Theorem 2.30 for depression probabilities depending on N . In particular, assuming that the depression probabilities are proportional to the coding level f_N , we obtain that this time is of order $\frac{1}{f_N^2}$.

Chapter 2

A Mathematical Analysis of Memory Lifetime in a simple Network Model of Memory

We study the learning of an external signal by a neural network and the time to forget it when this network is submitted to noise. The presentation of an external stimulus to the recurrent network of binary neurons may change the state of the synapses. Multiple presentations of a unique signal leads to its learning. Then, during the forgetting time, the presentation of other signals (noise) may also modify the synaptic weights. We construct an estimator of the initial signal using the synaptic currents and define by this way a probability of error. In our model, these synaptic currents evolve as Markov chains. We study the dynamics of these Markov chains and obtain a lower bound on the number of external stimuli that the network can receive before the initial signal is considered as forgotten (probability of error above a given threshold). Our results hold for finite size networks as well as in the large size asymptotic and they are based on a finite time analysis rather than large time asymptotic. We finally present numerical illustrations of our results.

The main results of this chapter have been published in *Neural Computation* [68].

2.1 Introduction

Amit and Fusi proposed in [5] a model to study the memory capacity of neural networks. The main novelty of their work was the online learning and forgetting of a sequence of random signals. Indeed, in previous models (e.g. [133] or [70]), signals are stored in a fixed weight matrix. This matrix is determined as a function of signals to learn. These models are called associative or attractor neural network (ANN) models: a stimulus is said to be stored if its neural representation is an attractor of the neural dynamics. The maximum storage capacity of ANN models have been widely studied. Gardner and Derrida computed this capacity for the optimal synaptic weight matrix in [54]. They showed that maximal storage is obtained for sparse coding. Moreover, there have been studies of the robustness to noise in the synaptic weight matrix and in the initial input. Sommer and Dayan proposed Bayesian retrieval processes for a stochastic version of the Willshaw model in [119]. However, beyond the maximum number of stimuli learnt, blackout catastrophe (forgetting of all memories) appears in ANN models. One way of avoiding this blackout is to allow the plasticity of the synapses.

The following experimental protocol is introduced in [5]: a neural network, with both binary synapses and binary neurons, receives and learns new random stimuli while forgetting the previous ones. Every signal may affect the synaptic weights. After a certain amount of time, the first

stimulus is presented again (priming) and the ability of the network to recognize it is questioned: how many stimuli can be presented before it forgets the initial signal? To provide an answer, Amit and Fusi ([5]) performed a signal-to-noise ratio (SNR) analysis. The signal under consideration is the sum of the synaptic currents onto one neuron when the network receives the priming. As Gardner and Derrida found in the case of the ANN models in [54], Amit and Fusi concluded that the coding of the stimuli needs to be sparse in order to optimise the storage capacity. They proposed a scaling of the coding level f_N (probability that a neuron is selective to a signal) as a function of the size N of the network. According to their retrieval criterion, the optimal coding level is on the order of $f_N \sim \frac{\log(N)}{N}$. In the large N asymptotic, what they called the storage capacity is then on the order of $\frac{1}{f_N^2}$ for depression probabilities proportional to f_N .

Extensions and approaches different from SNR have then been studied. First, Brunel, Carusi and Fusi studied a different protocol in [23]: they fixed the number of random stimuli and presented them randomly multiple times. Their analysis relied on the comparison of two quantities: the mean potentiation (MP) and the intra-class potentiation (ICP). MP is the mean of synaptic weights. ICP is the mean of synaptic weights among synapses involved in the learning of a stimulus. Intuitively, when ICP is much larger than MP, the trace of a stimulus in the synaptic weights is still non negligible. They found two possible loading regimes, a low-loading (resp. high-loading) regime with a memory capacity on the order $\frac{1}{f_N}$ (resp. $\frac{1}{f_N^2}$). Dubreuil, Amit and Brunel did in [41] a deeper analysis of the multiple presentations model introduced in [23] and the one shot learning model of [5], in the asymptotic N large and f_N small. Then, Elliott considered in [42] the mean number of signals presented before the synaptic current crosses a fixed threshold: the mean first passage time (MFPT). More complex and biologically plausible models have been proposed and analysed numerically in the following studies, see [7, 98, 135]. Finally, to the best of our knowledge, the first article to present a precise way to retrieve stimuli is [8]. In this article, they insisted on the role played by the synaptic correlations and proposed a way to compute numerically an approximation of the distributions of the synaptic currents. It enables them to introduce a new retrieval criterion based on what they called the retrieval probability.

Inspired by this last article, we study here a statistical test based on the synaptic currents. In particular, we study the probability of error associated to this test. Such an error has been studied before under some additive assumptions on the distribution of synaptic currents. Amit and Huang did a numerical analysis with a Gaussian approximation in [8]. In [41], Dubreuil et al. gave analytical results on the probability of no error in the large N asymptotic, assuming independence of the synapses (which leads the synaptic currents to follow Binomial distributions). Here, we perform an analytical study of this error without such approximations and we manage to control it by extending previous analytical studies of [5, 8] on some points. First, we give properties of the synaptic current process such as the spectrum of its transition matrix (Propositions 2.10 and 2.12) and the exact expectations and variances of the synaptic currents (Proposition 2.15). Moreover, we study the case of multiple presentations of the signal to be learnt. Finally, we give in Remark 2.28 and Theorem 2.30 explicit bounds on the time during which a given signal is kept in memory (probability of error below a given threshold). These results deal with a broader range of depression probabilities than in the previous studies. We summarize our asymptotic results in Remark 2.31.

This chapter is organised as follows. We expose the model and the statistical test in Section 2.2. After learning one specific signal, the network is submitted to random signals responsible for its forgetting. The statistical test consists in estimating (using a threshold estimator) the initial signal from the presynaptic inputs caused by priming. We measure whether the signal is still in memory by computing the error associated to this test. After the formal definition of this

error, the main results are presented. Then, Section 2.3 is devoted to the presentation of the results. The main result consists in deriving a lower bound on the maximum number of stimuli one can present while reasonably remembering the initial signal. This derivation relies on the fact that, asymptotically, as time goes to infinity, synaptic currents converge in law to a Binomial mixture (Corollary 2.23). We assume that, before learning, the synaptic currents follow their stationary distributions. Afterwards, the learning phase splits the network in two groups: the neurons activated by the signal and the others. Then, during the forgetting phase, the laws of the synaptic currents of these two groups are shown to remain Binomial mixtures with an explicit dynamics on their mixing distributions (Proposition 2.19). We evaluate the probability of error of the statistical test and the maximum number of stimuli one can present before the test fails (Remark 2.28 and Theorem 2.30). The computations are based on estimates on the support and on the tail of the mixing distributions. The short proofs follow the statement of the results whereas Section 2.4 is devoted to the long proofs. Finally, we perform numerical simulations in Section 2.5. We give in Appendix A.2 a beginning of the study of the weights matrix dynamics which is based on the Kronecker products of square matrices of dimension 2.

2.2 The model and the estimator

First, we present the neural network model and the protocol followed for learning and forgetting. Then, we define the estimator, derive the equations describing the dynamics of the synaptic currents and detail the main assumptions. Finally, we present typical numerical simulations at the end of this section.

2.2.1 The neural network and the protocol

In order to ease the introduction of the different variables, we suggest the reader to see the model as describing an experiment on a person's ability to learn a stimulus. In particular, we ask for how long a learnt signal can persist in memory when the person is exposed to some other signals which we term as noise.

Let us assume that we present a sequence of external stimuli to a network of $N + 1$ neurons. Thus, we sum over N external synaptic currents to get the total synaptic current onto one neuron. We do not study the dynamics of the membrane potential nor the firing rate of neurons, but rather we consider their neural activities that we note $V_t^N \in \{0, 1\}^{N+1}$ at time t . Hence, the neurons do not have their own dynamics but instead they follow the dynamics of the signals. We say that the neuron i is selective (resp. not selective) to a signal presented at time t if its neural response is $V_t^{i,N} = 1$ (resp. $V_t^{i,N} = 0$). We assume that a given signal uniquely determines the neural response, see [23] for a study of learning similar neural activity patterns. Therefore, we refer in an equivalent way to stimulus/signal or neural response in the following. Signals are assumed to be random and we denote by $(\dots, V_{-1}^N, V_0^N, V_1^N \dots)$ the corresponding sequence. We call t the time at which the t^{th} signal after V_0^N is shown. We assume that for different values of t , the vectors V_t^N are independent and identically distributed (*i.i.d.*) with values in $\{0, 1\}^{N+1}$. Moreover, for each t , the components $V_t^{1,N}, \dots, V_t^{N+1,N}$ of V_t^N are themselves *i.i.d.* with Bernoulli distribution with parameter f_N :

$$\forall t, \quad \forall i \in \llbracket 1, N + 1 \rrbracket, \quad f_N = \mathbb{P} \left(V_t^{i,N} = 1 \right) = 1 - \mathbb{P} \left(V_t^{i,N} = 0 \right).$$

The synaptic weight from neuron j to neuron i at time t is denoted by $W_t^{ij,N}$. It can only take two values $W_- < W_+$ and we denote by $W_t^N = \{W_t^{ij,N}, i \neq j\} \in \{W_-, W_+\}^{N(N+1)}$ the matrix of synaptic weights. We consider a plasticity rule which can be viewed as a classic Hebbian rule. The law of W_{t+1}^N only depends on W_t^N and V_t^N . The corresponding transition probabilities are

- $\mathbb{P}\left(W_{t+1}^{ij,N} = W_+ \mid W_t^{ij,N} = W_-, (V_t^{i,N}, V_t^{j,N}) = (1, 1)\right) = q^+$,
- $\mathbb{P}\left(W_{t+1}^{ij,N} = W_- \mid W_t^{ij,N} = W_+, (V_t^{i,N}, V_t^{j,N}) = (0, 1)\right) = q_{01}^{-,N}$,
- $\mathbb{P}\left(W_{t+1}^{ij,N} = W_- \mid W_t^{ij,N} = W_+, (V_t^{i,N}, V_t^{j,N}) = (1, 0)\right) = q_{10}^{-,N}$.

The transition probabilities not mentioned here and involving the change of state of a synaptic weight are equal to zero. For example, $\mathbb{P}\left(W_{t+1}^{ij,N} = W_- \mid W_t^{ij,N} = W_+, (V_t^{i,N}, V_t^{j,N}) = (0, 0)\right) = 0$. In order to simplify the notation and without loss of generality, we set:

$$W_- = 0 \text{ (weak synapse) and } W_+ = 1 \text{ (strong synapse).}$$

Moreover, in order to avoid critical cases, we also assume that

$$f_N \in (0, 1) \quad \text{and} \quad q_{01}^{-,N}, q_{10}^{-,N}, q^+ \in (0, 1]. \quad (2.1)$$

The parameters $q_{01}^{-,N}$ and $q_{10}^{-,N}$ represent respectively the homosynaptic and heterosynaptic depressions, see [23].

We now give the protocol to learn and then forget a signal. We denote by V_0^N the signal to be learnt. Before presenting it, we assume that the network has received a lot of random signals thereby driving the law of the synaptic weight matrix in its ‘‘stable’’ state at time $t = -r + 1$ (we prove in Proposition 2.8 that there is a unique invariant measure for the weight matrix). The learning phase consists in performing r presentations of V_0^N . In order to be consistent with the previous description, the sequence of presented stimuli is then $(\dots, V_{-r}^N, \underbrace{V_0^N, \dots, V_0^N}_{r \text{ times}}, V_1^N, V_2^N \dots)$ that

is $V_t^N = V_0^N$ for $t \in \llbracket -r+1, 0 \rrbracket$. The presentation of the subsequent signals leads to the forgetting of V_0^N .

Notation 2.1. We note $(\Omega, \mathcal{F}, \mathbb{P})$ the probability space on which the Markov process $(V_t^N, W_t^N)_{t \in \mathbb{Z}}$ is defined.

Remark 2.2. The model presented slightly differs with the model introduced by Amit and Fusi in [5]: it allows multiple presentation of the initial signal.

2.2.2 Presentation of the estimator

We study the consistency through time of the response of one neuron to the initial signal. To do so, we consider the previous protocol. After the repetitive presentation of V_0^N , the signal has left a certain footprint in the matrix W_1^N . This trace is subsequently erased by the presentation of the following signals. How much information from a stimulus learnt is left at time t ? As an answer, we define an estimator of this stimulus and compute the error associated to it. At time t , this estimation is done from the observation of the synaptic currents onto neurons when presenting

again V_0^N : this current onto the neuron i is $\sum_{j \neq i} W_t^{ij,N} V_0^{j,N}$. This is a fictive presentation. By fictive, we mean that the synaptic weights do not change following this fictitious presentations. In this framework, neurons are similar. Hence, in order to simplify the notation and without loss of generality, our study focuses on neuron $i = 1$. We denote by h_t^N ,

$$h_t^N \stackrel{def.}{=} \sum_{j=2}^{N+1} W_t^{1j,N} V_0^{j,N}, \quad (2.2)$$

the synaptic current onto neuron 1 when presenting again V_0^N at time t . Note that the process $(h_t^N)_{t \geq 0}$ strongly depends on the initial number K of active neurons

$$K \stackrel{def.}{=} \sum_{j=2}^{N+1} V_0^{j,N}. \quad (2.3)$$

We denote by $h_{t,K}^N$ the process equal in law to h_t^N knowing K , $h_{t,K}^N \stackrel{\mathcal{L}}{=} (h_t^N | K)$. The process $h_{t,K}^N$ is Markovian, see Proposition 2.5.

Remark 2.3. *We thus study the probability of error associated to the estimator of only one neuron response. We are currently studying the extension to the estimation of the entire stimulus, see Appendix A.2 for the first results.*

The estimator

We define a threshold estimator $\hat{V}_t^N : \llbracket 0, N \rrbracket \rightarrow \{0, 1\}$ such that $\hat{V}_t^N(\theta) = \mathbb{1}_{h_t^N > \theta}$ with the associated probability of errors:

$$\begin{aligned} p_e^{0,N}(t, \theta) &\stackrel{def.}{=} \mathbb{P} \left(\hat{V}_t^N(\theta) = 1 \mid V_0^{1,N} = 0 \right) = \mathbb{P} \left(h_t^N > \theta \mid V_0^{1,N} = 0 \right), \\ p_e^{1,N}(t, \theta) &\stackrel{def.}{=} \mathbb{P} \left(\hat{V}_t^N(\theta) = 0 \mid V_0^{1,N} = 1 \right) = \mathbb{P} \left(h_t^N \leq \theta \mid V_0^{1,N} = 1 \right). \end{aligned}$$

Notation 2.4. *We denote by $h_t^{y,N} \stackrel{\mathcal{L}}{=} (h_t^N \mid V_0^{1,N} = y)$ and $h_{t,K}^{y,N} \stackrel{\mathcal{L}}{=} (h_t^N \mid V_0^{1,N} = y, K)$.*

The probability of error $p_e^{0,N}(t, \theta) = \mathbb{P}(h_t^{0,N} > \theta)$ (resp. $p_e^{1,N}(t, \theta) = \mathbb{P}(h_t^{1,N} \leq \theta)$) corresponds to the probability that the estimator responds positively (resp. negatively) to the priming presented at time $t > 0$ whereas the neuron was not activated (resp. activated) initially. We aim at evaluating these errors: for fixed $\epsilon \in (0, 1)$, we estimate the largest time t_* such that both $p_e^{0,N}$ and $p_e^{1,N}$ are smaller than ϵ up to time t_* ,

$$t_*(\epsilon, r, N) \stackrel{def.}{=} \max_{\theta \in \llbracket 0, N \rrbracket} \left(\inf \{ t \geq 1, p_e^{0,N}(t, \theta) \vee p_e^{1,N}(t, \theta) \geq \epsilon \} \right), \quad (2.4)$$

where $x \vee y = \max(x, y)$ and $x \wedge y = \min(x, y)$.

Main Results - Contributions (informal)

*For any fixed error $\epsilon \in (0, 1)$, there is an unbounded set of couples $(N, r) \in \mathbb{N}^{*2}$ for which we show the existence of a threshold $\theta_{\epsilon,r,N} \in \{0, 1, \dots, N\}$ ensuring*

$$t_*(\epsilon, r, N) \geq \inf \{ t \geq 1, p_e^{0,N}(t, \theta_{\epsilon,r,N}) \vee p_e^{1,N}(t, \theta_{\epsilon,r,N}) \geq \epsilon \} \geq \hat{t}(\epsilon, r, N),$$

where an explicit formula of \hat{t} is given in Remark 2.28 for fixed potentiation and depression probabilities. Another formula of \hat{t} is given in Theorem 2.30 for depression probabilities depending on N . In particular, assuming that the depression probabilities are proportional to the coding level f_N , we obtain that $\hat{t}(\epsilon, r, N)$ is on the order of $\frac{1}{f_N^2}$.

The proofs of the results presented in Section 2.3 rely on the study of the Markov chains $(h_{t,K}^N)_{t \geq 1}$ and $(h_{t,K}^{y,N})_{t \geq 1}$.

Proposition 2.5. *The chains $(h_{t,K}^N)_{t \geq 1}$ and $(h_{t,K}^{y,N})_{t \geq 1}$ are Markovian. At the end of the learning phase, we have*

$$h_{1,K}^N \stackrel{\mathcal{L}}{=} h_{-r+1,K}^N + V_0^{1,N} \text{Bin}(K - h_{-r+1,K}^N, 1 - (1 - q^+)^r) - (1 - V_0^{1,N}) \text{Bin}(h_{-r+1,K}^N, 1 - (1 - q_{01}^{-,N})^r) \quad (2.5)$$

where, conditionally on $h_{-r+1,K}^N$, the two Binomial random variables are independent. And during the forgetting phase, for all $t \geq 1$:

$$h_{t+1,K}^N \stackrel{\mathcal{L}}{=} h_{t,K}^N + V_t^{1,N} \left[\text{Bin}(K - h_{t,K}^N, f_N q^+) - \text{Bin}(h_{t,K}^N, (1 - f_N) q_{10}^{-,N}) \right] - (1 - V_t^{1,N}) \text{Bin}(h_{t,K}^N, f_N q_{01}^{-,N}) \quad (2.6)$$

where, conditionally on $h_{t,K}^N$, the three Binomial random variables are independent.

The Markov chains $(h_{t,K}^{y,N})_{t \geq 1}$ for $y \in \{0, 1\}$, satisfy the equation (2.5) with $V_0^{1,N} = y$ and the equation (2.6).

Proof. In order to study the jump from $h_{t,K}^N$ to $h_{t+1,K}^N$, we count the synapses that potentiate and the ones that depress upon presenting a signal V_t^N . From the definitions (2.2) and (2.3) of h_t^N and $h_{t,K}^N$, we only need to consider the K synapses $W_t^{1j,N}$ with $j \geq 2$ such that $V_0^{j,N} = 1$. Among these synapses, at time t , there are $h_{t,K}^N$ strong synapses and $K - h_{t,K}^N$ weak synapses. Given $V_t^{1,N}$ and $h_{t,K}^N$, every synapse evolves independently following a Bernoulli law.

Equation (2.5) follows from the fact that from time $-r + 1$ to 1:

- if $V_0^{1,N} = 0$, every strong synapse is r times candidate to depression so it has probability $1 - (1 - q_{01}^{-,N})^r$ to depress,
- if $V_0^{1,N} = 1$, every weak synapse is r times candidate to potentiation so it has probability $1 - (1 - q^+)^r$ to potentiate.

Equation (2.6) follows from the fact that between time $t \geq 1$ and $t + 1$:

- if $V_t^{1,N} = 0$, the probability that a strong synapse depresses is $f_N q_{01}^{-,N}$,
- if $V_t^{1,N} = 1$, the probability that a weak synapse potentiates is $f_N q^+$ and the probability that a strong synapse depresses is $(1 - f_N) q_{10}^{-,N}$.

By definition of $h_{t,K}^{y,N}$, the chain satisfies equation (2.5) with $V_0^{1,N} = y$ and equation (2.6). \square

Corollary 2.6. *Assume that (2.1) holds. Then, for all $K \in \llbracket 0, N \rrbracket$, the Markov chain $(h_{t,K}^N)_{t \geq 1}$ admits a unique invariant measure π_K^N with support in $\llbracket 0, K \rrbracket$. Moreover, for any initial condition $h_{0,K}^N$, the Markov chain $(h_{t,K}^N)_{t \geq 1}$ converges in law to π_K^N . In addition, the chain $(h_t^N)_{t \geq 1}$ converges in law to $\pi_\infty^N = \sum_{K=0}^N \mathbb{P}(\hat{K} = K) \pi_K^N$ where $\hat{K} = \sum_{j=2}^{N+1} V_0^{j,N}$.*

Proof. By (2.1), the Markov chain $(h_{t,K}^N)_{t \geq 1}$ is irreducible and aperiodic on a finite state space. Thus, it admits a unique invariant measure towards which it converges.

Let $\hat{K} = \sum_{j=2}^{N+1} V_0^{j,N}$. From the Bayes' formula we get that for all $l \in \llbracket 0, N \rrbracket$,

$$\lim_{t \rightarrow \infty} \mathbb{P}(h_t^N = l) = \lim_{t \rightarrow \infty} \left(\sum_{K=0}^N \mathbb{P}(\hat{K} = K) \mathbb{P}(h_{t,K}^N = l) \right) = \sum_{K=0}^N \mathbb{P}(\hat{K} = K) \pi_K^N(l).$$

\square

Remark 2.7. *The Markov chains $(h_{t,K}^{0,N})_{t \geq 1}$ and $(h_{t,K}^{1,N})_{t \geq 1}$ have the same transition matrix as $(h_{t,K}^N)_{t \geq 1}$. They differ by their distribution at time $t = 1$. Hence, they both converge to π_K^N . Moreover, both $(h_t^{0,N})_{t \geq 1}$ and $(h_t^{1,N})_{t \geq 1}$ converge in law to π_∞^N .*

Proposition 2.8. *Under the assumption (2.1), the process $(V_t^N, W_t^N)_{t \geq 1}$ converges to its unique invariant measure. We denote it by ρ_∞^N .*

Proof. Same argument as for Corollary 2.6. \square

We now give the main assumptions.

Assumption 2.9.

2.9.1 $(V_0^N, W_{-r+1}^N) \stackrel{\mathcal{L}}{=} \rho_\infty^N$ and in particular $h_{-r+1,K}^N, h_{-r+1,K}^{0,N}, h_{-r+1,K}^{1,N} \stackrel{\mathcal{L}}{=} \pi_K^N$.

2.9.2 Assume that $\lim_{N \rightarrow \infty} f_N = 0$ and $\lim_{N \rightarrow \infty} N f_N = +\infty$.

2.9.3 Let $q_{01}^{-,N} = a_N f_N$ and $q_{10}^{-,N} = b_N f_N$ with $a_N, b_N : N \in \mathbb{N}^* \rightarrow \mathbb{R}$ such that a_N, b_N both converge in $[0, +\infty)$. However, we assume that at least one of the two limits is not 0 and

$$\lim_{N \rightarrow \infty} q_{01}^{-,N} = \lim_{N \rightarrow \infty} q_{10}^{-,N} = \lim_{N \rightarrow \infty} \frac{b_N^2}{N f_N a_N} = \lim_{N \rightarrow \infty} \frac{b_N}{N f_N} = 0, \quad \lim_{N \rightarrow \infty} N f_N a_N = +\infty.$$

We consider a general paradigm in which before receiving the stimulus V_0^N , many stimuli have already been sent $(\dots, V_{-r-2}^N, V_{-r-1}^N, \dots)$. We assume that the process $(V_t^N, W_t^N)_{t \leq -r+1}$ has reached its invariant measure at time $t = -r + 1$ by Assumption 2.9.1. Then, one key parameter is the coding level f_N . We assume that it depends on N in the analysis of the large N asymptotic: Assumption 2.9.2. This assumption refers to sparse coding as f_N tends to 0. An additional constraint put forward is that the mean number of selective neurons, $N f_N$, needs to be large enough: Assumption 2.9.2. In this context, we are interested to see how the dependence on N of the depression probabilities can affect the memory lifetime, see Assumption 2.9.3. This assumption gives conditions on the large N asymptotic behaviours of the depression probabilities.

2.2.3 First illustrations

In this subsection, we illustrate the dynamics of $\left(h_{t,K}^{y,N}\right)_{t \geq 0}$ and $\left(h_{t,K}^N\right)_{t \geq 0}$. In particular, we are interested in the effects of the coding level f_N on these synaptic currents. Let us assume that the signal V_0^N is of size $K = \lfloor Nf_N \rfloor$, where the floor function $\lfloor x \rfloor$ is equal to $k \in \mathbb{Z}$ if $k \leq x < k+1$. Let us have a look at the expected size of jumps of $h_{t,K}^N$ from the formulas (2.5), (2.6).

$$\begin{aligned} \text{For } t = 1, \quad \mathbb{E} \left[h_{1,K}^N - h_{-r+1,K}^N | h_{-r+1,K}^N, V_0^{1,N} = 0 \right] &= -h_{-r+1,K}^N (1 - (1 - q_{01}^{-,N})^r), \\ \mathbb{E} \left[h_{1,K}^N - h_{-r+1,K}^N | h_{-r+1,K}^N, V_0^{1,N} = 1 \right] &= (K - h_{-r+1,K}^N) (1 - (1 - q^+)^r), \\ \forall t \geq 1, \quad \mathbb{E} \left[h_{t+1,K}^N - h_{t,K}^N | h_{t,K}^N \right] &= (K - h_{t,K}^N) f_N^2 q^+ - h_{t,K}^N f_N (1 - f_N) (q_{10}^{-,N} + q_{01}^{-,N}). \end{aligned}$$

From these equations, we note that the average jump size strongly depends on f_N . When f_N is close to 1, the reception of V_0^N has a large impact on the weight matrix, easy to detect. However, the following average jump size are close to the initial one. Thus, as soon as some other stimuli are presented, the initial signal is forgotten: the distributions of $h_{t,K}^{0,N}$ and $h_{t,K}^{1,N}$ quickly overlap. Conversely, when f_N is close to 0, the average jump size is significantly different between the learning (relatively big jumps) and the forgetting (relatively small jumps) phases. As a consequence, the convergence to the stationary distribution, and thus forgetting, is slower. However, the learning still occurs: the initial jump is still big. In order to illustrate these phenomena, we plot simulation results obtained with a high coding level, $f_N = 0.8$ in Figure 2.1, and a low coding level, $f_N = 0.1$ in Figure 2.2.

Figure 2.1a shows that the size of jumps is effectively big for $f_N = 0.8$, just after learning as well as during forgetting time. Figure 2.1b illustrates the separation between the initial distributions of $h_{t,K}^{0,N}$ and $h_{t,K}^{1,N}$. Indeed, at time $t = -r + 1 = 0$, both $h_{0,K}^{0,N}$ and $h_{0,K}^{1,N}$ follow the invariant measure plotted in black. Then, after the reception of V_0^N , the distribution of $h_{1,K}^{0,N}$ is shifted to the left and the distribution of $h_{1,K}^{1,N}$ to the right. Initially, the signal is learnt because the distributions are well separated, see Figure 2.1b. Figures 2.1c and 2.1d exhibit the fast overlapping of these two distributions. Indeed, following the learning phase, the reception of new stimuli makes the two distributions converge back quickly to the invariant distribution. At time $t = 5$, the signal is already forgotten.

Figure 2.2 illustrates the advantages of a low coding level. Indeed, even at time $t = 20$, the two distributions do not overlap a lot and they remained uni-modal. This makes the choice of a threshold estimator reasonable. Moreover, such an estimator allows a tractable analysis.

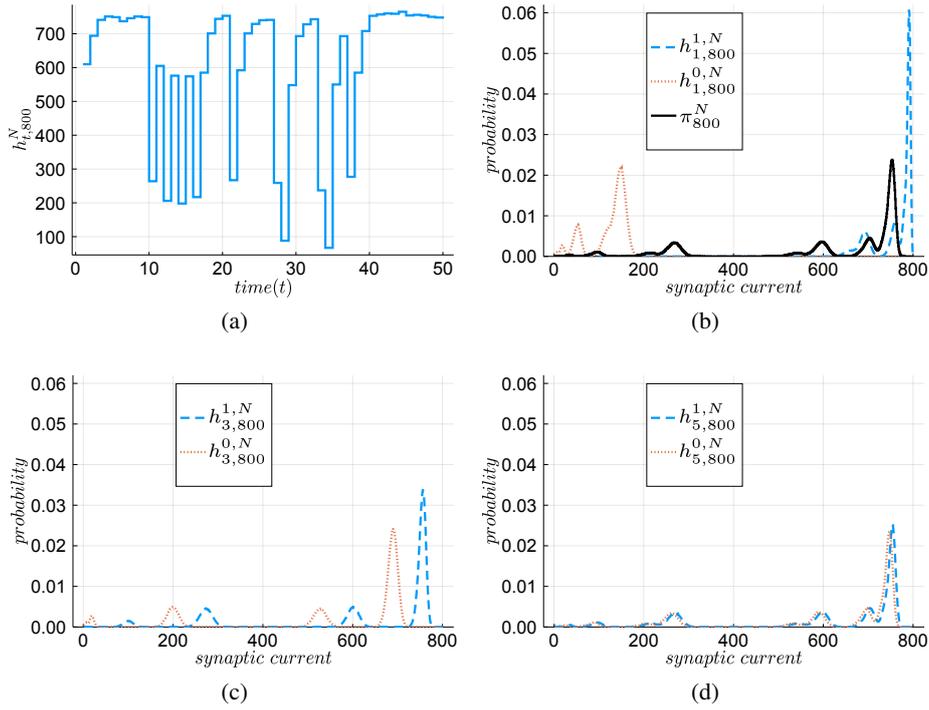


Figure 2.1: (2.1a): A typical trajectory of $h_{t,800}^N$. (2.1b): The distributions of $h_{1,800}^{0,N}$ and $h_{1,800}^{1,N}$ and the invariant measure π_{800}^N . (2.1c),(2.1d): The distributions of $h_{t,800}^{0,N}$ and $h_{t,800}^{1,N}$ at time $t = 3$ and $t = 5$. Parameters: $r = 1$, $N = 1000$, $K = 800$, $f_N = 0.8$, $q^+ = 0.8$, $q_{01}^{-,N} = 0.8$ and $q_{10}^{-,N} = 0.2$.

2.3 Results

In this section, we present our results. First, we give the spectrum of the transition matrix of the chains $(h_{t,K}^{0,N})_{t \geq 1}$ and $(h_{t,K}^{1,N})_{t \geq 1}$, see Notation 2.4, and their expectations and variances at any time $t \geq 1$. Then, we expose some properties satisfied by the distributions of $h_{t,K}^{0,N}$ and $h_{t,K}^{1,N}$, and their invariant measure π_K^N . They are the key to the proof of our main results, Theorems 2.27 and 2.30, which are presented in the last part of this Section.

2.3.1 Spectrum and SNR analysis

Spectrum

Let $P_K^{y,N}$ be the transition matrix of the synaptic current $(h_{t,K}^{y,N})_{-r+1 < t \leq 1}$. We denote by

$$\nu_{t,K}^{y,N} = \left[\nu_{t,K}^{y,N}(0), \nu_{t,K}^{y,N}(1), \dots, \nu_{t,K}^{y,N}(K) \right]$$

the distribution of $h_{t,K}^{y,N}$. We can then write $\nu_{1,K}^{y,N} = \nu_{0,K}^{y,N} P_K^{y,N} = \nu_{-r+1,K}^{y,N} \left(P_K^{y,N} \right)^r$.

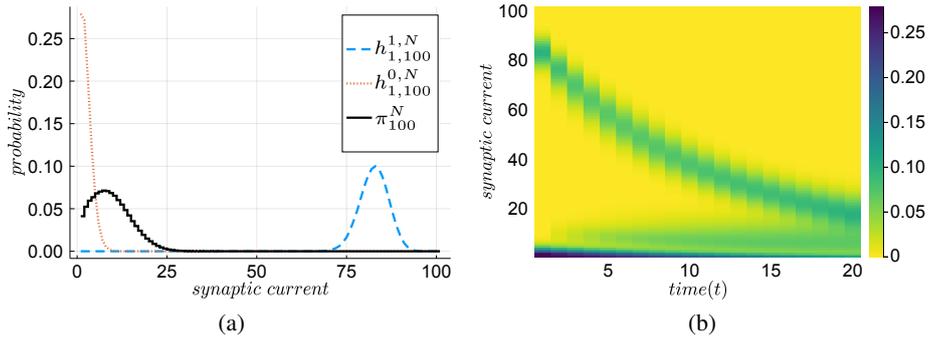


Figure 2.2: (2.2a): The distributions of $h_{t,K}^{0,N}$ and $h_{t,K}^{1,N}$ at time $t = 1$ and the invariant measure π_{100}^N . (2.2b): The sum of the two distributions $h_{t,K}^{0,N}$ and $h_{t,K}^{1,N}$ for $t \in [1, 20]$. The colour bar gives the probability values. Parameters: $r = 1$, $N = 1000$, $K = 100$, $f_N = 0.1$, $q^+ = 0.8$, $q_{01}^{-,N} = 0.8$, and $q_{10}^{-,N} = 0.2$.

Proposition 2.10. *The spectra of $P_K^{0,N}$ and $P_K^{1,N}$ are*

$$\Sigma \left(P_K^{0,N} \right) = \left\{ \left(1 - q_{01}^{-,N} \right)^i, 0 \leq i \leq K \right\} \quad \text{and} \quad \Sigma \left(P_K^{1,N} \right) = \left\{ \left(1 - q^+ \right)^i, 0 \leq i \leq K \right\}.$$

Proof. The dynamics give for all $j > i$, $P_{K,ij}^{0,N} = P_{K,ji}^{1,N} = 0$. So the matrices are triangular. Their spectra are given by the diagonal elements:

$$\forall i, \quad P_{K,ii}^{0,N} = \left(1 - q_{01}^{-,N} \right)^i \quad \text{and} \quad P_{K,ii}^{1,N} = \left(1 - q^+ \right)^i.$$

□

In the following, we use the

Notation 2.11. *We denote by $\Lambda_0^N = 1 - f_N q_{01}^{-,N}$, $\Lambda_1^N = 1 - (1 - f_N) q_{10}^{-,N} - f_N q^+$ and*

$$\forall i \in \llbracket 0, N \rrbracket, \quad \lambda_i^N = \left(1 - f_N \right) \left(\Lambda_0^N \right)^i + f_N \left(\Lambda_1^N \right)^i.$$

Proposition 2.12. *The spectrum of the transition matrix P_K^N of $\left(h_{t,K}^N \right)_{t \geq 1}$ is*

$$\Sigma \left(P_K^N \right) = \left\{ \lambda_i^N, 0 \leq i \leq K \right\}.$$

This proposition is proved in Section 2.4.2.

We deduce from Proposition 2.12 the rate of convergence of the law of $h_{t,K}^N$.

Corollary 2.13. *For all $0 \leq K \leq N$, the sequence of the distributions of the synaptic currents, $\left(\nu_{t,K}^N \right)_{t \geq 1}$, converges exponentially fast to the unique invariant measure π_K^N . In particular, there exists $c_N \in \mathbb{R}^+$ such that the distance in total variation between $\nu_{t,K}^N$ and π_K^N satisfies:*

$$\forall t \geq 1, \quad \left\| \nu_{t,K}^N - \pi_K^N \right\|_{TV} \stackrel{\text{def.}}{=} \frac{1}{2} \sum_{l=0}^K \left| \nu_{t,K}^N(l) - \pi_K^N(l) \right| \leq c_N \left(\lambda_1^N \right)^t.$$

We discuss, in the second paragraph of Section 2.6, the role played by this eigenvalue λ_1^N in our main results.

SNR Analysis

We propose to extend the previous SNR analysis done in [5, 116, 8]. The memory signal used here is neuron specific: we associate to each neuron a memory signal. It refers to the memory of the first stimulus. For the neuron 1, it is obtained from equation (1.2) by replacing $w_t^j \Delta w_0^j$ by $\left(W_t^{1j,N} V_0^{j,N} | V_0^{1,N} = 1\right)$ minus $\left(W_t^{1j,N} V_0^{j,N} | V_0^{1,N} = 0\right)$:

$$Z_{t,K}^N = h_{t,K}^{1,N} - h_{t,K}^{0,N}. \quad (2.7)$$

Notation 2.14. From the notations

$$\mu_t^{y,N} \stackrel{\text{def.}}{=} \mathbb{E} \left[h_t^{y,N} \right] \quad \text{and} \quad (\sigma_t^{y,N})^2 \stackrel{\text{def.}}{=} \text{Var} \left[h_t^{y,N} \right],$$

we define the SNR at time t as:

$$SNR_t^N \stackrel{\text{def.}}{=} \frac{\mu_t^{1,N} - \mu_t^{0,N}}{\sqrt{\sigma_{t,K}^{1,N} + \sigma_t^{0,N}}}. \quad (2.8)$$

In the previous article, they used the large time asymptotic value of the variance and do not give the explicit one. Here, we give an explicit formula of $\mu_t^{y,N}$ and $(\sigma_t^{y,N})^2$.

Proposition 2.15. Under Assumption 2.9, the expectations and variances of $h_t^{0,N}$ and $h_t^{1,N}$ satisfy:

$$\mu_t^{0,N} = N f_N (\mu^{*,N} - \mu^{0,N} (\lambda_1^N)^{t-1}), \quad \text{with} \quad \mu^{0,N} = \left(1 - (1 - q_{01}^{-,N})^r\right) \frac{f_N^2 q^+}{1 - \lambda_1^N}, \quad (2.9)$$

$$\mu_t^{1,N} = N f_N (\mu^{*,N} + \mu^{1,N} (\lambda_1^N)^{t-1}), \quad \text{with} \quad \mu^{1,N} = (1 - (1 - q^+)^r) \left(1 - \frac{f_N^2 q^+}{1 - \lambda_1^N}\right), \quad (2.10)$$

$$(\sigma_t^{y,N})^2 = \mu_t^{y,N} (1 - \mu_t^{y,N}) + f_N^2 N (N - 1) \left(C_0^{*,N} + C_1^{y,N} (\lambda_1^N)^{t-1} + C_2^{y,N} (\lambda_2^N)^{t-1}\right), \quad (2.11)$$

where $C_0^{*,N}$, $C_1^{y,N}$ and $C_2^{y,N}$ are constants depending on the parameters and $\mu^{*,N} = \frac{f_N^2 q^+}{1 - \lambda_1^N}$.

Moreover, in the case $a_N = a$ and $b_N = b$, the standard deviations stay of the same order of their large time asymptotic value that we denote by $\sigma^{*,N}$ and which satisfies:

$$(\sigma^{*,N})^2 = f_N N (\sigma_1^{*,N})^2 + f_N^3 N^2 (\sigma_2^{*,N})^2,$$

where $(\sigma_1^{*,N})^2 \approx_{N\infty} \frac{q^+}{q^+ + a + b}$ and $(\sigma_2^{*,N})^2 \approx_{N\infty} \frac{a^2 (q^+)^2}{2(q^+ + a + b)^3}$.

By this proposition, we justify the use of $\sigma^{*,N}$ instead of $\sigma_t^{y,N}$ for the computation of the SNR in the large N asymptotic. It was not obvious. Indeed, if one looks individually at the time-dependent terms of $\sigma_t^{y,N}$, they show the possibility of a term in $N^2 f_N^2$ instead of $N^2 f_N^3$. What saves us is that $\max_{t \geq 1} \{(\lambda_2^N)^t - (\lambda_1^N)^{2t}\}$ scales in f_N , see the proof in Section 2.4.2.

2.3.2 Binomial Mixture

The proofs of the results on the spectrum, the SNR and the main results are based on the various properties of the synaptic currents' dynamics that are presented in this section.

We denote by $F_{[0,1]}$ the set of cumulative distribution functions associated to the set $\mathcal{P}([0, 1])$ of probability measures on $[0, 1]$.

Definition 2.16. *The distribution of X is said to be a Binomial mixture with mixing distribution $g \in \mathcal{P}([0, 1])$ and size parameter K , denoted by $\text{BinMix}(K, g)$, if*

$$\forall j \in \llbracket 0, K \rrbracket, \quad \mathbb{P}(X = j) = \binom{K}{j} \int_0^1 u^j (1-u)^{K-j} g(du).$$

Remark 2.17.

- $X \stackrel{\mathcal{L}}{=} \text{BinMix}(K, g)$ is equivalent to $X|Y \stackrel{\mathcal{L}}{=} \text{Bin}(K, Y)$ where Y is a random variable independent of the Binomial and with law g . Indeed

$$\mathbb{P}(X = j) = \int_0^1 \mathbb{P}(X = j|Y = u)g(du) = \binom{K}{j} \int_0^1 u^j (1-u)^{K-j} g(du).$$

We use both notations $X \stackrel{\mathcal{L}}{=} \text{BinMix}(K, g)$ and $X \stackrel{\mathcal{L}}{=} \text{BinMix}(K, Y)$.

- The law of X is fully characterized by the moments $\mathbb{E}(Y), \mathbb{E}(Y^2), \dots, \mathbb{E}(Y^K)$. Hence, if $\tilde{g} \in \mathcal{P}([0, 1])$ is such that $\forall k \in \llbracket 0, K \rrbracket, \int_0^1 u^k \tilde{g}(du) = \int_0^1 u^k g(du)$, then $\text{BinMix}(K, \tilde{g}) \stackrel{\mathcal{L}}{=} \text{BinMix}(K, g)$.

First, we show that the set of Binomial mixtures is stable by the Markov chain $h_{t,K}^N$: assume that $h_{t,K}^N \stackrel{\mathcal{L}}{=} \text{BinMix}(K, g_t^N)$ for some $g_t^N \in \mathcal{P}([0, 1])$, then there exists a probability g_{t+1}^N , function of g_t^N , such that $h_{t+1,K}^N \stackrel{\mathcal{L}}{=} \text{BinMix}(K, g_{t+1}^N)$. Moreover, denoting by G_t^N the cumulative distribution function associated to g_t^N , we show that for all $t \geq 1$, $G_{t+1}^N(x) = \mathcal{R}(G_t^N)(x)$ where

Notation 2.18. For all $\Gamma \in F_{[0,1]}$ and $u \in \mathbb{R}$, \mathcal{R} is defined by

$$\mathcal{R}(\Gamma)(u) \stackrel{\text{def.}}{=} f_N \Gamma \left(\frac{u - f_N q^+}{\Lambda_1^N} \right) + (1 - f_N) \Gamma \left(\frac{u}{\Lambda_0^N} \right).$$

Proposition 2.19. *Let us assume that $h_{-r+1,K}^N \stackrel{\mathcal{L}}{=} \text{BinMix}(K, g_{-r+1}^N)$, for $g_{-r+1}^N \in \mathcal{P}([0, 1])$. Then for all $t \geq 1$, $\exists g_t^N, g_t^{0,N}, g_t^{1,N} \in \mathcal{P}([0, 1])$ such that $h_{t,K}^N \stackrel{\mathcal{L}}{=} \text{BinMix}(K, g_t^N)$ and $h_{t,K}^{y,N} \stackrel{\mathcal{L}}{=} \text{BinMix}(K, g_t^{y,N})$ for $y = 0, 1$. Moreover, at time $t = 1$,*

$$G_1^N(u) = f_N G_{-r+1}^N \left(\frac{u-1}{(1-q^+)^r} + 1 \right) + (1-f_N) G_{-r+1}^N \left(\frac{u}{(1-q_{01}^{-,N})^r} \right), \quad (2.12)$$

$$G_1^{1,N}(u) = G_{-r+1}^N \left(\frac{u-1}{(1-q^+)^r} + 1 \right) \quad \text{and} \quad G_1^{0,N}(u) = G_{-r+1}^N \left(\frac{u}{(1-q_{01}^{-,N})^r} \right), \quad (2.13)$$

$$\text{and } \forall t \geq 1, \quad G_{t+1}^N(u) = \mathcal{R}(G_t^N)(u) \quad \text{and} \quad G_{t+1}^{y,N}(u) = \mathcal{R}(G_t^{y,N})(u). \quad (2.14)$$

Finally, we deduce that

$$h_t^N \stackrel{\mathcal{L}}{=} \text{BinMix}(N, f_N g_t^N) \quad \text{and} \quad h_t^{y,N} \stackrel{\mathcal{L}}{=} \text{BinMix}(N, f_N g_t^{y,N}). \quad (2.15)$$

The equations of (2.15) derive from the application of the

Lemma 2.20. Assume $K \stackrel{\mathcal{L}}{=} \text{Bin}(N, f_N)$ and conditionally on K , X is independent of K with law $\text{Bin}(K, p)$, then $X \stackrel{\mathcal{L}}{=} \text{Bin}(N, f_N p)$.

Remark 2.21. First, we note that g_t^N does not depend on K . This is crucial for the proof of Theorems 2.27 and 2.30. Then, let the assumptions of the previous Proposition hold and denote by Y_t^N a random variable with distribution g_t^N . Knowing Y_t^N , we have $h_{t,K}^N \stackrel{\mathcal{L}}{=} \text{Bin}(K, Y_t^N)$. In particular, the mean synaptic current is given by $\mathbb{E}(h_{t,K}^N) = K \mathbb{E}(Y_t^N) = K \mathbb{P}(W_t^{1j,N} = 1)$. Indeed, $\mathbb{E}(h_{t,K}^N)$ is the mean number of strong synapses $W_t^{1j,N}$ (with j such that $V_0^{j,N} = 1$) at time t .

Finally, we show that \mathcal{R} is contracting and characterises π_K^N .

Proposition 2.22. The application \mathcal{R} acting on $F_{[0,1]}$ is contracting for the norm $L^1(0, 1)$. Moreover, there exists a unique $G^{*,N} \in F_{[0,1]}$ invariant for \mathcal{R} .

Propositions 2.19 and 2.22 are proved in the Section 2.4.1.

Corollary 2.23. Let $G^{*,N}$ be the unique fixed point of \mathcal{R} and $g^{*,N}$ its associated distribution. The invariant measure π_K^N of the Markov chain $h_{t,K}^N$ satisfies

$$\pi_K^N = \text{BinMix}(K, g^{*,N}).$$

We deduce that the invariant measure π_∞^N of h_t^N is given by

$$\pi_\infty^N = \text{BinMix}(N, f_N g^{*,N}).$$

Finally, the smallest interval $[m_\infty^N, M_\infty^N]$ containing the support of $g^{*,N}$ verifies

$$\text{supp}(g^{*,N}) \subset [m_\infty^N, M_\infty^N] \stackrel{\text{def.}}{=} \left[0, \frac{f_N q^+}{1 - \Lambda_1^N}\right]. \quad (2.16)$$

Proof. Let $g^{*,N} \in \mathcal{P}([0, 1])$ be a probability distribution such that its cumulative distribution function $G^{*,N}$ satisfies $\mathcal{R}(G^{*,N}) = G^{*,N}$. Then, by Proposition 2.19, $\text{BinMix}(K, g^{*,N})$ is invariant for $(h_{t,K}^N)_{t \geq 1}$. The result on π_∞^N follows from Corollary 2.6.

Now, let $[m_\infty^N, M_\infty^N]$ be the convex envelop of the support of $g^{*,N}$, then $\text{supp}(g^{*,N}) \subset [m_\infty^N, M_\infty^N] \subset [0, 1]$. Thus by the equation $\mathcal{R}(G^{*,N}) = G^{*,N}$, we get

$$m_\infty^N = m_\infty^N \Lambda_0^N \wedge (m_\infty^N \Lambda_1^N + f_N q^+), \quad \text{and} \quad M_\infty^N = M_\infty^N \Lambda_0^N \vee (M_\infty^N \Lambda_1^N + f_N q^+).$$

As $\Lambda_0^N = 1 - f_N q_{01}^{-,N} < 1$, the first equation implies that $0 \leq m_\infty^N \leq m_\infty^N \Lambda_0^N$ so $m_\infty^N = 0$. The second equation implies that $M_\infty^N = M_\infty^N \Lambda_1^N + f_N q^+$, thus $M_\infty^N = \frac{f_N q^+}{1 - \Lambda_1^N}$. \square

Remark 2.24. Propositions 2.19, 2.22, and the first part of Corollary 2.23 are in [8] with $q_{10}^{-,N} = 0$ and $r = 1$. We prove them here with a different method.

2.3.3 Main results

We end this section devoted to the presentation of results by the main results. The learning and the forgetting phases are both described by the Markov chains $(h_{t,K}^{y,N})_{-r+1 \leq t \leq 1}$ and $(h_{t,K}^{y,N})_{t \geq 1}$. The previous part of this Section gave us enough information on these chains to propose a precise control on their overlapping. This overlap is directly link to the probability of error associated to the estimator defined in Section 2.2.

Under Assumption 2.9.1, $h_{-r+1,K}^N$ follows its invariant distribution π_K^N , a Binomial mixture by Corollary 2.23. Thus, by Proposition 2.19, the processes $(h_{t,K}^{y,N})_{t \geq 1}$ follow also Binomial mixtures. Combining the inequality provided by Lemma 2.25, inequalities on Binomial tails (Lemma 2.26) and a control on the tail of the mixing distribution $g^{*,N}$ and on the support of $g_t^{1,N}$, we prove Theorem 2.27.

Lemma 2.25. *Under Assumption 2.9.1, for all $\theta \in \llbracket 0, N \rrbracket$, $\mathbb{P} \left(h_t^{0,N} > \theta \right) \leq \mathbb{P} \left(\pi_\infty^N > \theta \right)$.*

Proof. The proof is recursive and relies on the functional equation for the cumulative distribution of the synaptic currents (2.13) under Assumption 2.9.1. From (2.14), we have for all $x \in [0, 1]$, $G_1^{0,N}(x) = G^{*,N} \left(\frac{x}{\Lambda_0^N} \right) \geq G^{*,N}(x)$. Then,

$$\begin{aligned} G_2^{0,N}(x) &= f_N G_1^{0,N} \left(\frac{x - f_N q^+}{\Lambda_1^N} \right) + (1 - f_N) G_1^{0,N} \left(\frac{x}{\Lambda_0^N} \right) \\ &\geq f_N G^{*,N} \left(\frac{x - f_N q^+}{\Lambda_1^N} \right) + (1 - f_N) G^{*,N} \left(\frac{x}{\Lambda_0^N} \right) = G^{*,N}(x), \end{aligned}$$

and so forth so that for all $t \geq 1$ and x , $G_t^{0,N}(x) \geq G^{*,N}(x)$. It implies that for all K , $\theta \in \mathbb{N}$, $\mathbb{P} \left(\text{BinMix}(K, g_t^{0,N}) > \theta \right) \leq \mathbb{P} \left(\text{BinMix}(K, g^{*,N}) > \theta \right)$, which ends the proof. \square

Lemma 2.26. *Let $S_N \stackrel{\mathcal{L}}{=} \text{Bin}(N, p)$. Then, for all $\varepsilon \in (0, 1)$*

$$\mathbb{P} (S_N \geq Np(1 + \varepsilon)) \leq \exp \left(\frac{-Np \varepsilon^2}{2 + \varepsilon} \right), \quad (2.17)$$

$$\mathbb{P} (S_N \leq Np(1 - \varepsilon)) \leq \exp \left(\frac{-Np \varepsilon^2}{2} \right). \quad (2.18)$$

Proof. We use the method of [29]. Let S_N be the sum of X_1, X_2, \dots, X_N which are independent Bernoulli random variables of parameter p .

For all $\varepsilon \in (0, 1)$, $u \in \mathbb{R}^+$,

$$\begin{aligned} \mathbb{P} (S_N \geq Np(1 + \varepsilon)) &= \mathbb{P} \left(e^{u S_N} \geq e^{Np(1+\varepsilon)u} \right) \leq \frac{\mathbb{E} (e^{u S_N})}{e^{Np(1+\varepsilon)u}} = \frac{\prod_{i=1}^N \mathbb{E} (e^{u X_i})}{e^{Np(1+\varepsilon)u}} \\ &\leq \frac{(1 + p(e^u - 1))^N}{e^{Np(1+\varepsilon)u}} \leq \frac{e^{Np(e^u - 1)}}{e^{Np(1+\varepsilon)u}} = e^{Np(e^u - 1 - (1+\varepsilon)u)}. \end{aligned}$$

The minimum of the last term is reached for $u = \log(1 + \varepsilon)$ so

$$\mathbb{P} (X > Np(1 + \varepsilon)) \leq \left(\frac{e^\varepsilon}{(1 + \varepsilon)^{1+\varepsilon}} \right)^{Np} = \exp (Np(\varepsilon - (1 + \varepsilon) \log(1 + \varepsilon))).$$

From the inequality, $\forall z > 0$, $\log(1+z) \geq \frac{2z}{2+z}$, we obtain (2.17). In order to show (2.18), we proceed with the same method and use the inequality $\log(1+z) \geq \frac{z}{2} \frac{2+z}{1+z}$ whenever $-1 < z \leq 0$. \square

Theorem 2.27. For $y \in \{0, 1\}$, let $(h_t^{y,N})_{t \geq 1}$ be the solutions of (2.5) with $V_0^{1,N} = y$ and (2.6). Let us assume that Assumptions 2.9.1 and 2.9.2 hold and that $q_{01}^{-,N}$ and q^+ are fixed in $(0, 1]$ and $q_{10}^{-,N}$ in $[0, 1]$.

Then, for all $0 < \epsilon < 1$ and $r \in \mathbb{N}^*$, there exists $N(\epsilon, r) \in \mathbb{N}$ such that for all $N \geq N(\epsilon, r)$, there exist $\theta_{\epsilon,N} \in \llbracket 0, N \rrbracket$ and $\hat{t}(\epsilon, r, N)$ such that for all $1 \leq t \leq \hat{t}(\epsilon, r, N)$,

$$\mathbb{P}\left(h_t^{0,N} > \theta_{\epsilon,N}\right) \vee \mathbb{P}\left(h_t^{1,N} \leq \theta_{\epsilon,N}\right) \leq \epsilon.$$

This theorem is proved in Section 2.4.3.

In particular, we have $t_*(\epsilon, r, N) \geq \hat{t}(\epsilon, r, N)$. This result relies on the study of the mixing distributions $g^{*,N}$ and $g_t^{1,N}$. Thanks to Lemma 2.25, we know that as long as $g_t^{1,N}$ is far enough from $g^{*,N}$, the probability of error,

$$\mathbb{P}\left(h_t^{0,N} > \theta\right) \vee \mathbb{P}\left(h_t^{1,N} \leq \theta\right) \leq \mathbb{P}\left(\text{BinMix}(K, g^{*,N}) > \theta\right) \vee \mathbb{P}\left(\text{BinMix}(K, g_t^{1,N}) \leq \theta\right),$$

is small enough. This condition appears as an inequality depending both on the time and the accepted error ϵ . As long as this inequality holds, there exists a threshold θ such that the probability of error is below ϵ for all previous times. We give in the following Remark an explicit formula for the lower bound \hat{t} on t_* for any couple (ϵ, r) .

Remark 2.28. Recall that $M_\infty^N = \frac{f_N q^+}{1 - \Lambda_1^N}$, $m_1^{1,N} = 1 - (1 - q^+)r$, $M_\epsilon^N = \frac{f_N q^+ \left(1 + q_{01}^{-,N} \sqrt{\frac{2f_N}{\epsilon(1-\lambda_2^N)}}\right)}{f_N q^+ + (1 - f_N)(q_{01}^{-,N} + q_{10}^{-,N})}$, $\Lambda_0^N = 1 - f_N q_{01}^{-,N}$, $\Lambda_1^N = 1 - f_N q^+ - (1 - f_N)q_{10}^{-,N}$ and $\lambda_2^N = f_N(\Lambda_1^N)^2 + (1 - f_N)(\Lambda_0^N)^2$.

We proved that under Assumptions 2.9.1 and 2.9.2, for all $\epsilon, r, N \geq N(\epsilon, r)$ (N for which the two conditions given by (2.30) are satisfied), there exists $\theta_{\epsilon,N} \in \llbracket 0, N \rrbracket$ and \hat{t} such that for all $1 \leq t \leq \hat{t}(\epsilon, r, N)$, $\mathbb{P}\left(h_t^{0,N} > \theta_{\epsilon,N}\right) \vee \mathbb{P}\left(h_t^{1,N} \leq \theta_{\epsilon,N}\right) \leq \epsilon$.

In particular, if $q_{01}^{-,N}, q_{10}^{-,N}, q^+ \in (0, 1]$

$$\hat{t}(\epsilon, r, N) - 1 = \left\lfloor \frac{\log\left(\frac{2\sqrt{-2\log(\frac{\epsilon}{2})} N f_N M_\infty^N - 4\log(\frac{\epsilon}{2})}{N f_N (m_1^{1,N} - M_\infty^N)}\right)}{\log(\Lambda_1^N)} \right\rfloor \wedge \left\lfloor \frac{\log\left(\frac{f_N^2 q^+ q_{01}^{-,N}}{(1 - \Lambda_1^N)(\Lambda_0^N - \Lambda_1^N)(m_1^{1,N} - M_\infty^N)}\right)}{\log(\Lambda_1^N)} \right\rfloor,$$

$$\text{and if } q_{10}^{-,N} = 0, \hat{t}(\epsilon, r, N) - 1 = \left\lfloor \frac{\log\left(\frac{\left(\sqrt{M_\epsilon^N N f_N} + \sqrt{-2\log(\frac{\epsilon}{2})}\right)^2 - \frac{3}{2}\log(\epsilon)}{N f_N}\right)}{\log(\Lambda_0^N)} \right\rfloor.$$

Example 2.29. We give here a detailed result for a particular choice of parameters. Let $q^+ = q_{01}^{-,N} = 1$, $q_{10}^{-,N}$ small enough, and $f_N = \frac{q_{10}^{-,N}}{3 + q_{10}^{-,N}}$. Explicit computations give

$$\hat{t}(\epsilon, r, N) = \left\lfloor \frac{\log\left(\frac{1}{9}\right) \vee \log\left(\frac{\sqrt{-2\log(\frac{\epsilon}{2})} N f_N - 16\log(\frac{\epsilon}{2})}{3N f_N}\right)}{\log(1 - 4f_N)} \right\rfloor.$$

For instance, for $q_{10}^{-,N} = 0.005$ we get $f_N = 0.00167$ and

$$\hat{t}(\epsilon = 0.001, r = 1, N = 2.10^5) = 246 \quad \text{and} \quad \theta_{\epsilon, N} = 122.$$

We also give a formula when the depression probabilities depend on N in Theorem 2.30.

Theorem 2.30. Assume Assumptions 2.9.1, 2.9.2 and 2.9.3 are satisfied. Then, for all $\epsilon \in (0, 1)$, r large enough, there exists $N(\epsilon, r) \in \mathbb{N}$ such that for all $N \geq N(\epsilon, r)$,

$$\hat{t}(\epsilon, r, N) = t_c^N + \left\lfloor \frac{\log(C(\epsilon, r, N))}{\log(\Lambda_0^N)} \right\rfloor,$$

with t_c^N defined in (2.28) and $C(\epsilon, r, N) \in (0, 1)$ satisfies $\frac{\log(C(\epsilon, r, N))}{f_N} \rightarrow +\infty$. Moreover, if $\lim a_N$, and $\lim b_N$ exist and are finite, $\frac{\log(C(\epsilon, r, N))}{\log(\Lambda_0^N)}$ is on the order of $\frac{1}{f_N^2}$.

This theorem is proved in Section 2.4.3.

We note that $\log(\Lambda_0^N) = \log(1 - a_N f_N^2) \approx_{N\infty} -a_N f_N^2$. Concerning $C(\epsilon, r, N)$ (and then $\hat{t}(\epsilon, r, N)$), it mainly depends on the different large N asymptotic of a_N and b_N . We detail in Remark 2.31 the different large N asymptotic of $\hat{t}(\epsilon, r, N)$.

Remark 2.31. In the large N asymptotic (under Assumptions 2.9.2 and 2.9.3), we can compute the terms equivalent to \hat{t} in the different a_N, b_N cases ($a \in \mathbb{R}_*^+$ and $b \in \mathbb{R}^+$):

conditions on a_N, b_N and r	$\hat{t}(\epsilon, r, N)$ for large N
$b_N \rightarrow +\infty, b_N = \mathcal{O}(a_N), \forall r$	$\frac{\log\left(\sqrt{\frac{b_N}{a_N}} + 2\sqrt{\frac{-\log(\frac{\epsilon}{2})b_N}{q^+ N f_N}}\right)}{f_N^2 a_N}$
$a_N, b_N \rightarrow +\infty$ of same order, $\forall r$	$-\frac{\log\left(1 - \frac{a_N}{a_N + b_N}\right)}{2f_N^2 a_N}$
$a_N = \mathcal{O}(b_N), b_N \rightarrow +\infty, \forall r$	$\frac{1}{2f_N^2 b_N}$
$a_N \rightarrow +\infty, b_N \rightarrow b \in \mathbb{R}^+, \forall r$	$\frac{-\log\left(\sqrt{\frac{q^+}{(1-(1-q^+)^r) \wedge \frac{q^+}{q^+ + b}}} a_N + 2\sqrt{\frac{-\log(\frac{\epsilon}{2})(q^+ + b)}{q^+ N f_N}}\right)}{f_N^2 a_N}$
$a_N \rightarrow 0, b_N \rightarrow b > 0, \forall r > r_0$	$\frac{1}{2f_N^2 (q^+ + b)}$
$a_N = a, b_N \rightarrow 0$ ou $b_N = 0, \forall r > r_0$	$-\frac{\log\left(\frac{q^+}{(1-(1-q^+)^r)(q^+ + a)}\right)}{2f_N^2 a}$
$a_N = a, b_N = b, \forall r > r_0$	$-\frac{\log\left(1 - \frac{a}{q^+ + a + b}\right)}{2f_N^2 a}$

Table 2.1: The large N equivalent of $\hat{t}(\epsilon, r, N)$ in function of a_N and b_N .

Remark 2.32. Note that we have also proved the following result:

For every $\epsilon > 0$ and N large enough, there exists r_0 such that, if the initial signal is presented at least r_0 times, then it is well memorized after at least $\hat{t}(\epsilon, r, N)$ presentations of noisy signals. Moreover, in the large r asymptotic, $h_{1,K}^{0,N} \stackrel{\mathcal{L}}{=} \delta_0$ (dirac in 0) and $h_{1,K}^{1,N} \stackrel{\mathcal{L}}{=} \delta_K$ (dirac in K). Thus, the initial error is null. However, the \hat{t} increases with r until reaching a threshold value which is given by the expression of Remarks 2.28 and 2.31 replacing the quantities $m_1^{1,N}$ by 1.

2.4 Proofs

2.4.1 Proofs linked with the Binomial Mixture behaviour

Proof of Proposition 2.19

Notation 2.33. Let g be a distribution on \mathbb{R} with cumulative distribution function G . For all $a, b \in \mathbb{R}$, we denote by $g_{(a,b)}$ the distribution associated to the cumulative distribution function $G_{(a,b)}$ such that

$$\forall u \in \mathbb{R}, \quad G_{(a,b)}(u) = G\left(\frac{u-b}{a-b}\right).$$

Proposition 2.19 relies on the following

Lemma 2.34. Let Z be a mixture of Binomial $Z = \text{BinMix}(K, Y_Z)$. Let $0 \leq b < a < 1$. Conditionally on Z , consider two independent Binomial distributions $\text{Bin}(Z, a)$ and $\text{Bin}(K - Z, b)$ and define $X = \text{Bin}(Z, a) + \text{Bin}(K - Z, b)$. Then

$$X \stackrel{\mathcal{L}}{=} \text{BinMix}(K, Y_X) \quad \text{with} \quad Y_X = (a - b)Y_Z + b. \quad (2.19)$$

In particular, $G_X(u) = G_{Z,(a,b)}(u)$, where G_X (resp. G_Z) is the cumulative distribution function of X (resp. Z).

Proof. Let \tilde{U} , $(U_i)_{1 \leq i \leq K}$, $(V_i)_{1 \leq i \leq K}$, $(\eta_i)_{1 \leq i \leq K}$ and $(W_i)_{1 \leq i \leq K}$ be *i.i.d.* random variables following the uniform law on $[0, 1]$. By the first point of Remark 2.17, Z is the sum of $(Z_i)_{1 \leq i \leq K}$ *i.i.d.* Bernoulli of parameter $Y_Z = G_Z^{-1}(\tilde{U})$. Thus, we obtain that conditionally on Z ,

$$X = \underbrace{\sum_{i=1}^K Z_i \mathbb{1}_{\{V_i \leq a\}}}_{\stackrel{\mathcal{L}}{=} \text{Bin}(Z, a)} + \underbrace{\sum_{i=1}^K (1 - Z_i) \mathbb{1}_{\{\eta_i \leq b\}}}_{\stackrel{\mathcal{L}}{=} \text{Bin}(K - Z, b)}$$

where the Binomials are independent. Then, let consider $\forall i, Z_i = \mathbb{1}_{\{U_i \leq G_Z^{-1}(\tilde{U})\}}$. Thus,

$$\begin{aligned} X &= \sum_{i=1}^K \mathbb{1}_{\{U_i \leq G_Z^{-1}(\tilde{U})\}} \mathbb{1}_{\{V_i \leq a\}} + \sum_{i=1}^K \mathbb{1}_{\{U_i > G_Z^{-1}(\tilde{U})\}} \mathbb{1}_{\{\eta_i \leq b\}}, \\ &= \sum_{i=1}^K \mathbb{1}_{\{U_i \leq G_Z^{-1}(\tilde{U}), V_i \leq a\} \cup \{U_i > G_Z^{-1}(\tilde{U}), \eta_i \leq b\}}. \end{aligned} \quad (2.20)$$

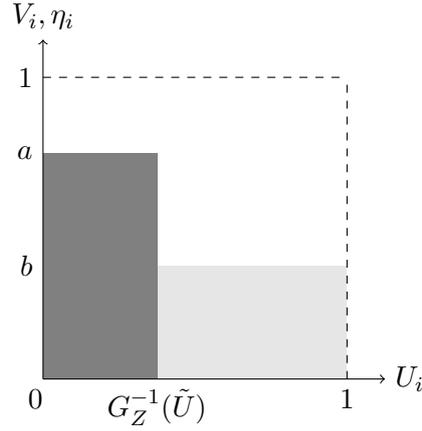


Figure 2.3: In gray, the domain to which the couple (U_i, V_i, η_i) needs to belong to from equation (2.20).

For all Borel set $\mathcal{D} \subset [0, 1]^3$, $\mathbb{P}((U_i, V_i, \eta_i) \in \mathcal{D}) = \mathcal{V}(\mathcal{D})$ where $\mathcal{V}(\mathcal{D})$ is the volume of \mathcal{D} . Let $W_i \stackrel{\mathcal{L}}{=} \mathcal{U}([0, 1])$, then $\mathbb{P}((U_i, V_i, \eta_i) \in \mathcal{D}) = \mathbb{P}(W_i \leq \mathcal{V}(\mathcal{D}))$. In Figure 2.3, we put V_i and η_i on the same axis as they do not depend one on the other so that the volume

$$\mathcal{V}\left(\left\{U_i \leq G_Z^{-1}(\tilde{U}), V_i \leq a\right\} \cup \left\{U_i > G_Z^{-1}(\tilde{U}), \eta_i \leq b\right\}\right)$$

is equal to the sum of the two grey areas. We deduce that

$$X = \sum_{i=1}^K \mathbb{1}_{\{W_i \leq b + (a-b)G_Z^{-1}(\tilde{U})\}} = \sum_{i=1}^K \mathbb{1}_{\{G_Z\left(\frac{W_i - b}{a-b}\right) \leq \tilde{U}\}} = \sum_{i=1}^K \mathbb{1}_{\{G_X(W_i) \leq \tilde{U}\}},$$

with $G_X(w) = G_{Z,(a,b)}(w)$. We conclude that (2.19) is satisfied. \square

We now prove of Proposition 2.19:

Proof of Proposition 2.19. We first show (2.12) and (2.14) for $h_{i,K}^N$, then the rest follows.

At $t = 1$, from equation (2.5) we get

$$\begin{aligned} \mathcal{L}\left(h_{1,K}^N \mid V_0^{1,N} = 1, h_{-r+1,K}^N\right) &= \text{Bin}\left(h_{-r+1,K}^N, 1\right) + \text{Bin}\left(K - h_{-r+1,K}^N, 1 - (1 - q^+)^r\right) \\ \mathcal{L}\left(h_{1,K}^N \mid V_0^{1,N} = 0, h_{-r+1,K}^N\right) &= \text{Bin}\left(h_{-r+1,K}^N, (1 - q_{01}^{-,N})^r\right). \end{aligned}$$

Applying twice Lemma 2.34 with $(a, b) = (1, 1 - (1 - q^+)^r)$ and then $(a, b) = (1 - (1 - q_{01}^{-,N})^r, 0)$, we obtain, using Notation 2.33,

$$\begin{aligned} \mathcal{L}\left(h_{1,K}^N \mid V_0^{1,N} = 1, h_{-r+1,K}^N\right) &\stackrel{\mathcal{L}}{=} \text{BinMix}\left(K, g_{-r+1, (1, 1 - (1 - q^+)^r)}\right) \\ \mathcal{L}\left(h_{1,K}^N \mid V_0^{1,N} = 0, h_{-r+1,K}^N\right) &\stackrel{\mathcal{L}}{=} \text{BinMix}\left(K, g_{-r+1, (1 - (1 - q_{01}^{-,N})^r, 0)}\right). \end{aligned}$$

Thus,

$$\begin{aligned} \mathbb{P}(h_{1,K}^N = j | h_{-r+1,K}^N) &= \mathbb{P}(V_0^{1,N} = 1) \mathbb{P}(h_{1,K}^N = j | V_0^{1,N} = 1, h_{-r+1,K}^N) \\ &\quad + \mathbb{P}(V_0^{1,N} = 0) \mathbb{P}(h_{1,K}^N = j | V_0^{1,N} = 0, h_{-r+1,K}^N) \\ &= \binom{K}{j} \int_0^1 u^j (1-u)^{K-j} (f_N g_{-r+1, (1, 1 - (1-q^+)^r)}(du) + (1-f_N) g_{-r+1, (1 - (1-q_0^{-,N})^r, 0)}(du)), \end{aligned}$$

which enables us to get (2.12).

Now, assume that $h_{t,K}^N \stackrel{\mathcal{L}}{=} \text{BinMix}(K, g_t^N)$, for some fixed $t \geq 1$. Then, by equation (2.6),

$$\begin{aligned} \mathcal{L}(h_{t+1,K}^N | V_t^{1,N} = 1, h_{t,K}^N) &= \text{Bin}(K - h_{t,K}^N, f_N q^+) + \text{Bin}(h_{t,K}^N, 1 - (1-f_N)q_{10}^{-,N}), \\ \mathcal{L}(h_{t+1,K}^N | V_t^{1,N} = 0, h_{t,K}^N) &= \text{Bin}(h_{t,K}^N, 1 - f_N q_{01}^{-,N}), \end{aligned}$$

where the Binomials are independent conditionally on $h_{t,K}^N$. Applying twice Lemma 2.34, first with $(a, b) = (1 - (1-f_N)q_{10}^{-,N}, f_N q^+)$ and then with $(a, b) = (1 - f_N q_{01}^{-,N}, 0)$, we get

$$\begin{aligned} \mathcal{L}(h_{t+1,K}^N | V_t^{1,N} = 1) &\stackrel{\mathcal{L}}{=} \text{BinMix}(K, g_{t, (1 - (1-f_N)q_{10}^{-,N}, f_N q^+)}) \\ \mathcal{L}(h_{t+1,K}^N | V_t^{1,N} = 0) &\stackrel{\mathcal{L}}{=} \text{BinMix}(K, g_{t, (1 - f_N q_{01}^{-,N}, 0)}). \end{aligned}$$

Hence, $h_{t+1,K}^N \stackrel{\mathcal{L}}{=} \text{BinMix}(K, f_N g_{t, (1 - (1-f_N)q_{10}^{-,N}, f_N q^+)} + (1-f_N)g_{t, (1 - f_N q_{01}^{-,N}, 0)})$, and we deduce that $h_{t+1,K}^N \stackrel{\mathcal{L}}{=} \text{BinMix}(K, g_{t+1}^N)$ with $G_{t+1}^N(x) = \mathcal{R}(G_t^N)(x)$.

For the processes $(h_{t,K}^{y,N})_{t \geq 0}$, we proceed exactly with the same method with the fact that $V_0^{1,N} = y$ in Proposition 2.5.

Finally, equations of (2.15) derives from the application of Lemma 2.20. \square

Proof of Proposition 2.22

Proof. **1. The map \mathcal{R} is a contraction**

Let $\Gamma_1, \Gamma_2 \in F_{[0,1]}$. We recall that $\Lambda_1^N = 1 - (1-f_N)q_{10}^{-,N} - f_N q^+$, $\Lambda_0^N = 1 - f_N q_{01}^{-,N}$.

$$\begin{aligned} &\|\mathcal{R}(\Gamma_2) - \mathcal{R}(\Gamma_1)\|_{L^1(0,1)} \\ &\leq \int_0^1 f_N \left| \Gamma_2 \left(\frac{u - f_N q^+}{\Lambda_1^N} \right) - \Gamma_1 \left(\frac{u - f_N q^+}{\Lambda_1^N} \right) \right| + (1-f_N) \left| \Gamma_2 \left(\frac{u}{\Lambda_0^N} \right) - \Gamma_1 \left(\frac{u}{\Lambda_0^N} \right) \right| du \\ &= f_N \int_{f_N q^+}^{1 - (1-f_N)q_{10}^{-,N}} \left| \Gamma_2 \left(\frac{(u - f_N q^+)}{\Lambda_1^N} \right) - \Gamma_1 \left(\frac{(u - f_N q^+)}{\Lambda_1^N} \right) \right| du \\ &\quad + (1-f_N) \int_0^{\Lambda_0^N} \left| \Gamma_2 \left(\frac{u}{\Lambda_0^N} \right) - \Gamma_1 \left(\frac{u}{\Lambda_0^N} \right) \right| du \\ &= f_N \Lambda_1^N \int_0^1 |\Gamma_2(u) - \Gamma_1(u)| du + (1-f_N) \Lambda_0^N \int_0^1 |\Gamma_2(u) - \Gamma_1(u)| du \\ &= \underbrace{(f_N \Lambda_1^N + (1-f_N) \Lambda_0^N)}_{\lambda_1^N} \|\Gamma_2 - \Gamma_1\|_{L^1(0,1)}. \end{aligned}$$

As $\lambda_1^N < 1$, the map \mathcal{R} acting on $F_{[0,1]}$ is strictly contracting in $L^1(0, 1)$.

2. Existence and uniqueness of a fixed point

We now prove the second point of the Lemma. For all $\Gamma_0 \in F_{[0,1]}$, by contraction of \mathcal{R} , $(\mathcal{R}^n(\Gamma_0))_{n \geq 0}$ is a Cauchy sequence for the $L^1(0, 1)$ norm. By completeness of $L^1(0, 1)$, this sequence converges to some $\Gamma \in L^1(0, 1)$. It remains to prove that Γ can be chosen in $F_{[0,1]}$. First, any limit Γ is non decreasing almost everywhere. Define $G^{*,N}(x) = \lim_{y \rightarrow x_+} \Gamma(y)$. The function $G^{*,N}$ is càdlàg and satisfies for every $x \leq 0$, $G^{*,N}(x) = 0$ and for every $x \geq 1$, $G^{*,N}(x) = 1$. Thus $G^{*,N} \in F_{[0,1]}$ and $\mathcal{R}(G^{*,N}) = G^{*,N}$. Finally, the uniqueness of $G^{*,N}$ is deduced from the fact that \mathcal{R} is strictly contracting. \square

2.4.2 Proofs of SNR and Spectrum results

Spectrum: proof of Proposition 2.12

Lemma 2.35. *Let X and Y be two random variables in $[0, 1]$ with cumulative distribution functions G_X and G_Y . We assume that there exist $\eta \in [0, 1]$, $a, \bar{a} \in [0, 1)$ and $b, \bar{b} \in (0, 1]$ with $a + b \leq 1$, $\bar{a} + \bar{b} \leq 1$ such that*

$$G_Y(u) = \eta G_X\left(\frac{u-a}{b}\right) + (1-\eta)G_X\left(\frac{u-\bar{a}}{\bar{b}}\right). \quad (2.21)$$

Then $\forall k \in \mathbb{N}$, $\mathbb{E}[Y^k] = \eta \mathbb{E}[(a+bX)^k] + (1-\eta)\mathbb{E}[(\bar{a}+\bar{b}X)^k]$.

Proof. First, note that $G_X\left(\frac{u-a}{b}\right)$ is the cumulative distribution function of $a+bX$. Second, for all random variables U, V, W , we have

$$G_U(z) = \eta G_V(z) + (1-\eta)G_W(z) \implies \mathbb{E}[U^k] = \eta \mathbb{E}[V^k] + (1-\eta)\mathbb{E}[W^k].$$

This last result is obtained by differentiation, multiplication by z^k and integration. It ends the proof of the lemma. \square

In the proof below, we use the classical convention $\binom{i}{j} = 0$ when $j > i$ or $j < 0$.

Proof of Proposition 2.12. We denote by $\nu_{i,K}^N = [\nu_{i,K}^N(0), \nu_{i,K}^N(1), \dots, \nu_{i,K}^N(K)]$ the distribution of $h_{i,K}^N$. Its transition matrix $P_K^N = (P_K^{ij,N})_{0 \leq i,j \leq K}$ can be derived from Proposition 2.5:

$$\begin{aligned} P_K^{ij,N} &= (1-f_N) \binom{i}{i-j} (f_N q_{01}^{-,N})^{i-j} (1-f_N q_{01}^{-,N})^j \\ &+ f_N \sum_{l=0}^i \binom{i}{l} ((1-f_N) q_{10}^{-,N})^l (\Lambda_1^N + f_N q^+)^{i-l} \binom{K-i}{j-i+l} (f_N q^+)^{j-i+l} (1-f_N q^+)^{K-j-l}. \end{aligned}$$

Let us define the two matrices \tilde{P}_K^N and Q_K^N such that for all $0 \leq i, j \leq K$:

$$\tilde{P}_K^{ij,N} = f_N \binom{j}{i} (\Lambda_1^N)^i (f_N q^+)^{j-i} + (1-f_N) \delta_{ij} (\Lambda_0^N)^i \quad \text{and} \quad Q_K^{ij,N} = \binom{K}{i} \binom{i}{j} (-1)^{i-j}.$$

Then, assuming that $\nu_{t,K}^N \stackrel{\mathcal{L}}{=} \text{BinMix}(K, g_t^N)$ and denoting by $U_t^N = [U_t^{0,N}, U_t^{1,N}, \dots, U_t^{K,N}]$ with $U_t^{k,N} = \int u^k g_t^N(du)$, we get by definition 2.16: $\nu_{t,K}^N = U_t^N Q_K^N$. Moreover, by Lemma 2.35 we have $U_{t+1}^N = U_t^N \tilde{P}_K^N$. Finally, by definition we have $\nu_{t+1,K}^N = \nu_{t,K}^N P_K^N$, so we obtain:

$$\nu_{t+1,K}^N = U_{t+1}^N Q_K^N = U_t^N \tilde{P}_K^N Q_K^N = U_t^N Q_K^N (Q_K^N)^{-1} \tilde{P}_K^N Q_K^N = \nu_{t,K}^N (Q_K^N)^{-1} \tilde{P}_K^N Q_K^N = \nu_{t,K}^N P_K^N.$$

A straightforward computation shows that $Q_K^N P_K^N = \tilde{P}_K^N Q_K^N$. Thus P_K^N and \tilde{P}_K^N have the same spectrum. Finally, \tilde{P}_K^N is a triangular matrix with λ_i^N as diagonal elements. \square

SNR: proof of Proposition 2.15

From equation (2.14), we can link the associated cumulative distribution function $G_{t+1}^{y,N}$ to the previous one $G_t^{y,N}$. From this result and Lemmas 2.35 and 2.36, we can derive a simple expression of the first two moments of $g_t^{y,N}$ and an expression of $(\sigma_{t,K}^{y,N})^2 = \text{Var} [h_{t,K}^{y,N}]$.

We remind a classical result, see Theorem 4.4.7 in [28].

Lemma 2.36. (Conditional variance identity) For any two random variables X and Y ,

$$\text{Var} X = \mathbb{E}(\text{Var}(X | Y)) + \text{Var}(\mathbb{E}(X | Y))$$

provided that the quantities are finite.

Remark 2.37. Recall that $X \stackrel{\mathcal{L}}{=} \text{BinMix}(N, f_N g_x)$ is equivalent to $X|Y_x \stackrel{\mathcal{L}}{=} \text{Bin}(N, f_N Y_x)$ where Y_x is a random variable with distribution g_x . So, Lemma 2.36 reads

$$\text{Var} [X] = \mathbb{E} \left[\underbrace{\text{Var} [X|Y_x]}_{N f_N Y_x (1-f_N Y_x)} \right] + \text{Var} \left[\underbrace{\mathbb{E} [X|Y_x]}_{N f_N Y_x} \right] = N (f_N \mathbb{E} [Y_x] - f_N^2 \mathbb{E} [Y_x^2]) + N^2 f_N^2 \text{Var} [Y_x].$$

We now prove Proposition 2.15.

Proof of Proposition 2.15. We denote by $Y^{*,N}$ a random variable with distribution $g^{*,N}$ and by $Y_t^{y,N}$ the random variables with distributions $g_t^{y,N}$. From Assumption 2.9.1 and Corollary 2.23, for all $t \geq 0$, $\mathcal{L} (h_t^{y,N}) = \text{BinMix} (N, f_N g_t^{y,N})$. Thus, with notation 2.14 and from Remark 2.37:

$$\mu_t^{y,N} = N f_N \mathbb{E} [Y_t^{y,N}], \quad (2.22)$$

$$\left(\sigma_t^{y,N} \right)^2 = N f_N \left(\mathbb{E} [Y_t^{y,N}] - f_N \mathbb{E} [(Y_t^{y,N})^2] \right) + N^2 f_N^2 \left(\mathbb{E} [(Y_t^{y,N})^2] - \mathbb{E} [Y_t^{y,N}]^2 \right). \quad (2.23)$$

We now compute $\mathbb{E} [Y_t^{y,N}]$ and $\mathbb{E} [(Y_t^{y,N})^2]$. To do so, we use Lemma 2.35 and the equations of Proposition 2.19. From (2.13) and (2.14), we obtain

$$\begin{aligned} \mathbb{E} [Y_1^{0,N}] &= (1 - q_0^{-,N})^r \mathbb{E} [Y_{-r+1}^{0,N}] \\ \mathbb{E} [Y_1^{1,N}] &= 1 - (1 - q^+)^r + (1 - q^+)^r \mathbb{E} [Y_{-r+1}^{1,N}] \\ \mathbb{E} [Y_{t+1}^{y,N}] &= f_N \left(f_N q^+ + \Lambda_1^N \mathbb{E} [Y_t^{y,N}] \right) + (1 - f_N) \Lambda_0^N \mathbb{E} [Y_t^{y,N}] \\ &= \lambda_1^N \mathbb{E} [Y_t^{y,N}] + f_N^2 q^+. \end{aligned}$$

We find that

$$\mathbb{E} \left[Y_t^{y,N} \right] = \left(\mathbb{E} \left[Y_1^{y,N} \right] - \mathbb{E} \left[Y^{*,N} \right] \right) (\lambda_1^N)^{t-1} + \mathbb{E} \left[Y^{*,N} \right] \quad \text{and} \quad \mathbb{E} \left[Y^{*,N} \right] = \frac{f_N^2 q^+}{1 - \lambda_1^N}.$$

From Assumption 2.9.1, we have $\mathbb{E} \left[Y_{-r+1}^{y,N} \right] = \mathbb{E} \left[Y^{*,N} \right]$, so

$$\mathbb{E} \left[Y_1^{0,N} \right] - \mathbb{E} \left[Y^{*,N} \right] = - \left(1 - (1 - q_{01}^{-,N})^r \right) \mathbb{E} \left[Y^{*,N} \right] \quad (2.24)$$

$$\mathbb{E} \left[Y_1^{1,N} \right] - \mathbb{E} \left[Y^{*,N} \right] = (1 - (1 - q^+)^r) (1 - \mathbb{E} \left[Y^{*,N} \right]). \quad (2.25)$$

This ends the proof for the expectations. We now evaluate the second order moments in order to compute the variances using (2.23).

$$\begin{aligned} \mathbb{E} \left[(Y_1^{0,N})^2 \right] &= (1 - q_{01}^{-,N})^{2r} \mathbb{E} \left[(Y_{-r+1}^{0,N})^2 \right] \\ \mathbb{E} \left[(Y_1^{1,N})^2 \right] &= (1 - q^+)^{2r} \mathbb{E} \left[(Y_{-r+1}^{1,N})^2 \right] + 2(1 - q^+)^r (1 - (1 - q^+)^r) \mathbb{E} \left[Y_{-r+1}^{1,N} \right] \\ &\quad + (1 - (1 - q^+)^r)^2 \\ \mathbb{E} \left[(Y_{t+1}^{y,N})^2 \right] &= f_N \left[(f_N q^+)^2 + 2f_N q^+ \Lambda_1^N \mathbb{E} \left[Y_t^{y,N} \right] + (\Lambda_1^N)^2 \mathbb{E} \left[(Y_t^{y,N})^2 \right] \right] \\ &\quad + (1 - f_N) (\Lambda_0^N)^2 \mathbb{E} \left[(Y_t^{y,N})^2 \right] \\ &= \lambda_2^N \mathbb{E} \left[(Y_t^{y,N})^2 \right] + 2f_N^2 q^+ \Lambda_1^N \mathbb{E} \left[Y_t^{y,N} \right] + f_N^3 q^{+2}. \end{aligned}$$

Thus

$$\mathbb{E} \left[(Y^{*,N})^2 \right] = \frac{f_N^3 q^{+2} + 2f_N^2 q^+ \Lambda_1^N \frac{f_N^2 q^+}{1 - \lambda_1^N}}{1 - \lambda_2^N}.$$

Then, introducing the notation $\gamma \stackrel{def.}{=} \frac{2f_N^2 q^+ \Lambda_1^N}{\lambda_1^N - \lambda_2^N}$, we find

$$\mathbb{E} \left[(Y_{t+1}^{y,N})^2 \right] - \gamma \mathbb{E} \left[Y_{t+1}^{y,N} \right] = \lambda_2^N \left(\mathbb{E} \left[(Y_t^{y,N})^2 \right] - \gamma \mathbb{E} \left[Y_t^{y,N} \right] \right) + f_N^3 q^{+2} - \gamma f_N^2 q^+.$$

Thus,

$$\begin{aligned} \mathbb{E} \left[(Y_t^{y,N})^2 \right] &= \underbrace{\left(\mathbb{E} \left[(Y_1^{y,N})^2 \right] - \gamma \mathbb{E} \left[Y_1^{y,N} \right] - \left(\mathbb{E} \left[(Y^{*,N})^2 \right] - \gamma \mathbb{E} \left[Y^{*,N} \right] \right) \right)}_{C_2^{y,N}} (\lambda_2^N)^{t-1} \\ &\quad + \gamma \underbrace{\left(\mathbb{E} \left[Y_1^{y,N} \right] - \mathbb{E} \left[Y^{*,N} \right] \right)}_{C_1^{y,N}} (\lambda_1^N)^{t-1} + \underbrace{\mathbb{E} \left[(Y^{*,N})^2 \right]}_{C_0^{*,N}}. \end{aligned}$$

We conclude the proof using equation (2.23). For their exact computation, one can use Assumption 2.9.1 which gives $\mathbb{E} \left[(Y_{-r+1}^{y,N})^2 \right] = \mathbb{E} \left[(Y^{*,N})^2 \right]$.

We now show the second point which deals with comparing $\sigma_t^{y,N}$ to its large time asymptotic value $\sigma^{*,N}$. To do so, we use Taylor expansions in f_N to derive the leading terms of $(\sigma_t^{y,N})^2$. We first

rewrite $\sigma_t^{y,N}$:

$$\begin{aligned}
(\sigma_t^{y,N})^2 &= \mu_t^{y,N}(1 - \mu_t^{y,N}) + f_N^2 N(N-1) \left(C_0^{*,N} + C_1^{y,N} (\lambda_1^N)^{t-1} + C_2^{y,N} (\lambda_2^N)^{t-1} \right) \\
&= (\sigma^{*,N})^2 + N f_N \left(\left\{ \mathbb{E} [Y_1^{y,N}] - \mathbb{E} [Y^{*,N}] - f_N C_1^{y,N} \right\} (\lambda_1^N)^{t-1} - f_N C_2^{y,N} (\lambda_2^N)^{t-1} \right) \\
&\quad + N^2 f_N^2 \left(\left\{ C_1^{y,N} - 2\mathbb{E} [Y^{*,N}] (\mathbb{E} [Y_1^{y,N}] - \mathbb{E} [Y^{*,N}]) \right\} (\lambda_1^N)^{t-1} \right. \\
&\quad \quad \left. + C_2^{y,N} (\lambda_2^N)^{t-1} - (\mathbb{E} [Y_1^{y,N}] - \mathbb{E} [Y^{*,N}])^2 (\lambda_1^N)^{2(t-1)} \right) \\
&= (\sigma^{*,N})^2 + N f_N \left(\left\{ (\mathbb{E} [Y_1^{y,N}] - \mathbb{E} [Y^{*,N}]) (1 - \gamma f_N) \right\} (\lambda_1^N)^{t-1} - f_N C_2^{y,N} (\lambda_2^N)^{t-1} \right) \\
&\quad + N^2 f_N^2 \left(\left\{ (\gamma - 2\mathbb{E} [Y^{*,N}]) (\mathbb{E} [Y_1^{y,N}] - \mathbb{E} [Y^{*,N}]) \right\} (\lambda_1^N)^{t-1} \right. \\
&\quad \quad \left. + C_2^{y,N} (\lambda_2^N)^{t-1} - (\mathbb{E} [Y_1^{y,N}] - \mathbb{E} [Y^{*,N}])^2 (\lambda_1^N)^{2(t-1)} \right) \\
&\stackrel{def.}{=} (\sigma^{*,N})^2 + N f_N \textcircled{1}^y + N^2 f_N^2 \textcircled{2}^y.
\end{aligned}$$

The aim is thus to check that for all time t we have $(\sigma_t^{y,N})^2 - (\sigma^{*,N})^2 = \mathcal{O}_{N\infty}(f_N N + f_N^3 N^2)$. To do so, we show that $\textcircled{1}^y = \mathcal{O}_{N\infty}(1)$ and $\textcircled{2}^y = \mathcal{O}_{N\infty}(f_N)$.

First, for all $i \geq 1$, $0 \leq \lambda_i^N \leq 1$. Then, $1 - \gamma f_N \approx_{N\infty} 1$ and $C_2^{y,N} = \mathcal{O}_{N\infty}(1)$, so $\textcircled{1}^y = \mathcal{O}_{N\infty}(1)$.

Second, $\gamma - 2\mathbb{E} [Y^{*,N}] = \mathcal{O}_{N\infty}(f_N)$ so from equalities (2.24) and (2.25), we have:

$$\left\{ (\gamma - 2\mathbb{E} [Y^{*,N}]) (\mathbb{E} [Y_1^{y,N}] - \mathbb{E} [Y^{*,N}]) \right\} (\lambda_1^N)^{t-1} = \mathcal{O}_{N\infty}(f_N).$$

Then, $C_2^{0,N} = \mathcal{O}_{N\infty}(f_N^2)$ and $(\mathbb{E} [Y_1^{0,N}] - \mathbb{E} [Y^{*,N}])^2 = \mathcal{O}_{N\infty}(f_N^2)$, so we deduce that $\textcircled{2}^0 = \mathcal{O}_{N\infty}(f_N)$.

Finally, $C_2^{1,N} = \mathcal{O}_{N\infty}(1)$ and $(\mathbb{E} [Y_1^{0,N}] - \mathbb{E} [Y^{*,N}])^2 = \mathcal{O}_{N\infty}(1)$ with the same leading term:

$$C_2^{1,N} - (\mathbb{E} [Y_1^{1,N}] - \mathbb{E} [Y^{*,N}])^2 = \mathcal{O}_{N\infty}(f_N).$$

We conclude the proof with the computation of the maximum of $t \in \mathbb{R}^+ \mapsto (\lambda_2^N)^t - (\lambda_1^N)^{2t} \in \mathbb{R}^+$ which is maximal in

$$t_{max} = \frac{\log \left(\frac{\log(\lambda_2^N)}{2 \log(\lambda_1^N)} \right)}{2 \log(\lambda_1^N) - \log(\lambda_2^N)}.$$

We obtain

$$\begin{aligned}
(\lambda_2^N)^{t_{max}} &= \exp \left(-1 + \frac{(b + q^+)^2}{4(a + b + q^+)} f_N + \mathcal{O}(f_N) \right) \\
(\lambda_1^N)^{2t_{max}} &= \exp \left(-1 - \frac{(b + q^+)^2}{4(a + b + q^+)} f_N + \mathcal{O}(f_N) \right).
\end{aligned}$$

So

$$(\lambda_2^N)^{t_{max}} - (\lambda_1^N)^{2t_{max}} \approx_{N\infty} e^{-1} \frac{(b + q^+)^2}{2(a + b + q^+)} f_N.$$

Thus, $\textcircled{2}^1 = \mathcal{O}_{N\infty}(f_N)$ which ends the proof. \square

2.4.3 Proofs of Theorems on Memory Lifetime

Proof of Theorem 2.27

Proof of Theorem 2.27. The proof follows these lines: from Lemma 2.25 we have that $\mathbb{P}\left(h_t^{0,N} > \theta\right)$ is inferior or equal to $\pi_\infty^N([\theta, +\infty[)$. Hence, we propose a threshold θ based on the measure π_∞^N such that $\pi_\infty^N([\theta, +\infty[) \leq \epsilon$ and then we bound the time before which $\mathbb{P}\left(h_t^{1,N} \leq \theta\right) \geq \epsilon$.

We split $\pi_\infty^N([\theta, +\infty[)$ in two terms. We recall that $\pi_\infty^N = \text{BinMix}(K, g^{*,N})$ with $K \stackrel{\mathcal{L}}{=} \text{Bin}(N, f_N)$ and $[0, M_\infty^N]$ is the smallest interval containing the support of $g^{*,N}$. So

$$\begin{aligned} \pi_\infty^N([\theta, +\infty[) &= \int_0^{M_\infty^N} \mathbb{P}(\text{Bin}(K, u) > \theta) g^{*,N}(du) = \int_0^{M_\infty^N} \mathbb{P}(\text{Bin}(N, f_N u) > \theta) g^{*,N}(du) \\ &\leq \mathbb{P}(\text{Bin}(N, f_N M_\epsilon^N) > \theta) + \int_{M_\epsilon^N}^{M_\infty^N} g^{*,N}(du). \end{aligned}$$

The second equality comes from the following property: assume $K \stackrel{\mathcal{L}}{=} \text{Bin}(N, f_N)$ and conditionally on K , X is independent of K with law $\text{Bin}(K, p)$, then $X \stackrel{\mathcal{L}}{=} \text{Bin}(N, f_N p)$. Let $Y^{*,N}$ be a random variable with distribution $g^{*,N}$. We propose a value for M_ϵ^N using the Bienaymé-Tchebychev inequality:

$$M_\epsilon^N = \left(\sqrt{\frac{2\text{Var}(Y^{*,N})}{\epsilon}} + \mathbb{E}(Y^{*,N}) \right) \wedge M_\infty^N \quad \Rightarrow \quad \int_{M_\epsilon^N}^{M_\infty^N} g^{*,N}(du) \leq \frac{\epsilon}{2}.$$

We first fix $\theta_{\epsilon,N}$ such that $\mathbb{P}(\text{Bin}(N, f_N M_\epsilon^N) \geq \theta_{\epsilon,N} + 1) \leq \frac{\epsilon}{2}$. To do so, we apply Lemma 2.26 with $\varepsilon = \frac{\theta_{\epsilon,N} + 1}{N f_N M_\epsilon^N} - 1$ and obtain:

$$\theta_{\epsilon,N} = \left\lfloor N f_N M_\epsilon^N + \sqrt{-2 \log\left(\frac{\epsilon}{2}\right) N f_N M_\epsilon^N - \log\left(\frac{\epsilon}{2}\right)} \right\rfloor.$$

We now bound the probability of error $\mathbb{P}\left(h_t^{1,N} \leq \theta_{\epsilon,N}\right)$. Let $[m_t^{1,N}, M_t^{1,N}]$ be the smallest interval containing the support of $g_t^{1,N}$. Then, we get:

$$\mathbb{P}\left(h_t^{1,N} \leq \theta_{\epsilon,N}\right) = \int_{m_t^{1,N}}^{M_t^{1,N}} \mathbb{P}(\text{Bin}(K, u) \leq \theta_{\epsilon,N}) g_t^{1,N}(du) \leq \mathbb{P}\left(\text{Bin}\left(N, f_N m_t^{1,N}\right) \leq \theta_{\epsilon,N}\right).$$

Using Lemma 2.26 with $\varepsilon = 1 - \frac{\theta_{\epsilon,N}}{N f_N m_t^{1,N}}$, we get

$$\mathbb{P}\left(\text{Bin}\left(N, f_N m_t^{1,N}\right) \leq \theta_{\epsilon,N}\right) \leq \exp\left(-\frac{\left(N f_N m_t^{1,N} - \theta_{\epsilon,N}\right)^2}{2 N f_N m_t^{1,N}}\right).$$

Using the inequality $\sqrt{x} + \sqrt{y} \geq \sqrt{x+y}$ for all $x, y > 0$, we obtain that if

$$N f_N m_t^{1,N} \geq \theta_{\epsilon,N} + \sqrt{-2 \log(\epsilon) \theta_{\epsilon,N} - 2 \log(\epsilon)} \quad (2.26)$$

then $\mathbb{P}\left(h_t^{1,N} \leq \theta_{\epsilon,N}\right) \leq \epsilon$. Let us define $m_{\epsilon,N} \stackrel{\text{def.}}{=} \frac{1}{Nf_N} \left(\theta_{\epsilon,N} + \sqrt{-2\log(\epsilon)\theta_{\epsilon,N}} - 2\log(\epsilon)\right)$.

Using the bound $\theta_{\epsilon,N} \leq \left(\sqrt{Nf_N M_\epsilon^N} + \frac{\sqrt{-2\log(\frac{\epsilon}{2})}}{2}\right)^2$ we get

$$Nf_N m_{\epsilon,N} + \frac{3}{2} \log(\epsilon) = \left(\sqrt{\theta_{\epsilon,N}} + \frac{\sqrt{-2\log(\epsilon)}}{2}\right)^2 \leq \left(\sqrt{M_\epsilon^N Nf_N} + \sqrt{-2\log(\frac{\epsilon}{2})}\right)^2. \quad (2.27)$$

We now find $m_t^{1,N}$. From equation (2.14) and the definition of \mathcal{R} (see Notation 2.18), we have

$$\forall t \geq 1, \quad m_{t+1}^{1,N} = m_t^{1,N} \Lambda_0^N \wedge (m_t^{1,N} \Lambda_1^N + f_N q^+).$$

We note that for N large enough such that $m_t^{1,N} > \frac{f_N q^+}{\Lambda_0^N - \Lambda_1^N} \geq M_\infty^N$, we have

$$\frac{f_N q^+}{1 - \Lambda_1^N} = M_\infty^N < m_t^{1,N} \Lambda_1^N + f_N q^+ < m_t^{1,N} \Lambda_0^N < m_t^{1,N}.$$

Denoting by $t_c^N = \inf\{t \in \mathbb{N}^*, m_t^{1,N} \leq \frac{f_N q^+}{\Lambda_0^N - \Lambda_1^N}\}$, we obtain

$$m_t^{1,N} = \left((m_1^{1,N} - M_\infty^N) (\Lambda_1^N)^{(t \wedge t_c^N) - 1} + M_\infty^N \right) (\Lambda_0^N)^{(t - t_c^N) \mathbb{1}_{t > t_c^N}}. \quad (2.28)$$

Let us now consider q^+ , $q_{01}^{-,N}$ and $q_{10}^{-,N}$ fixed in $(0, 1]$. By definition, $M_\epsilon^N \leq M_\infty^N$, hence

$$Nf_N m_{\epsilon,N} \leq \left(\sqrt{M_\infty^N Nf_N} + \sqrt{-2\log\left(\frac{\epsilon}{2}\right)}\right)^2 - \frac{3}{2} \log(\epsilon).$$

Therefore, the inequality (2.26) holds true as long as

$$t - 1 \leq \left(\left[\frac{\log\left(\frac{2\sqrt{-2\log(\frac{\epsilon}{2})Nf_N M_\infty^N - 4\log(\frac{\epsilon}{2})}}{Nf_N(m_1^{1,N} - M_\infty^N)}\right)}{\log(\Lambda_1^N)} \right] \wedge t_c^N \right)_+ \quad \text{with } (x)_+ = x \mathbb{1}_{x \geq 0}. \quad (2.29)$$

But $m_1^{1,N} = 1 - (1 - q^+)^r \geq q^+$ and both $\frac{f_N q^+}{\Lambda_0^N - \Lambda_1^N}$ and $M_\infty^N = \frac{f_N q^+}{f_N q^+ + (1 - f_N)(q_{01}^{-,N} + q_{10}^{-,N})}$ tends to 0 with increasing N so there exists $N(\epsilon, r)$ such that for all $N \geq N(\epsilon, r)$ we can remove “ $(\cdot)_+$ ” in the inequality (2.29): that is to say for all $N \geq N(\epsilon, r)$ such that,

$$\frac{2\sqrt{-2\log(\frac{\epsilon}{2})Nf_N M_\infty^N - 4\log(\frac{\epsilon}{2})}}{Nf_N(m_1^{1,N} - M_\infty^N)} < 1 \quad \text{and} \quad m_1^{1,N} > \frac{f_N q^+}{\Lambda_0^N - \Lambda_1^N}.$$

Using the fact that for all $\epsilon \in (0, 1)$, $\sqrt{-2\log(\frac{\epsilon}{2})} \leq -2\log(\frac{\epsilon}{2})$ and $m_1^{1,N} \geq q^+$ we get the two following conditions on N :

$$2 \exp\left(-\frac{Nf_N(m_1^{1,N} - M_\infty^N)}{4(\sqrt{Nf_N M_\infty^N} + 1)}\right) < \epsilon \quad \text{and} \quad \frac{f_N q^+}{f_N q^+ + (1 - f_N)q_{10}^{-,N} - f_N q_{01}^{-,N}} < q^+. \quad (2.30)$$

In the particular case $q_{10}^{-,N} = 0$, we have $M_\infty^N = 1$ so the dynamics of $m_t^{1,N}$ is simply $m_t^{1,N} = m_1^{1,N} (\Lambda_0^N)^{t-1}$. We compute $\mathbb{E}(Y^{*,N})$ and $\text{Var}(Y^{*,N})$ using Lemma 2.35 and equation (2.14):

$$\mathbb{E}(Y^{*,N}) = \frac{f_N^2 q^+}{1 - \lambda_1^N} = \frac{f_N q^+}{f_N q^+ + (1 - f_N)(q_{01}^{-,N} + q_{10}^{-,N})}, \quad \text{Var}(Y^{*,N}) = \frac{f_N^5 (1 - f_N) q^+{}^2 q_{01}^{-,N 2}}{(1 - \lambda_1^N)^2 (1 - \lambda_2^N)}.$$

Hence,

$$M_\epsilon^N = \frac{f_N q^+ \left(1 + q_{01}^{-,N} \sqrt{\frac{2f_N}{\epsilon(1-\lambda_2^N)}}\right)}{f_N q^+ + (1-f_N)(q_{01}^{-,N} + q_{10}^{-,N})}. \quad (2.31)$$

We note that $1 - \lambda_2^N \approx_{N \rightarrow \infty} 2f_N(q_{01}^{-,N} + q_{10}^{-,N})$, so M_ϵ^N converges to 0 with increasing N . Thus, by inequality (2.27), there exists a $N(\epsilon, r)$ such that for all $N \geq N(\epsilon, r)$, $m_{\epsilon, N} < 1$. We conclude that for all $N \geq N(\epsilon, r)$, the inequality (2.26) holds true as long as

$$t - 1 \leq \left\lfloor \frac{\log \left(\frac{(\sqrt{M_\epsilon^N N f_N} + \sqrt{-2 \log(\frac{\epsilon}{2})})^2 - \frac{3}{2} \log(\epsilon)}{N f_N} \right)}{\log(\Lambda_0^N)} \right\rfloor > 0.$$

□

Proof of Theorem 2.30

Proof. We use the results proved in the proof of Theorem 2.27. From the dynamics of $m_t^{1, N}$ given by equation (2.28) and the bound $m_{t_c}^1 \geq m_1^{1, N} \wedge M_\infty^N$, we obtain that the inequality (2.26) is satisfied as long as

$$t - 1 \leq t_c^N + \left(\left\lfloor \frac{\log \left(\frac{m_{\epsilon, N}}{m_{t_c}^1} \right)}{\log(\Lambda_0^N)} \right\rfloor \right)_+ \leq t_c^N + \left(\left\lfloor \frac{\log \left(\frac{m_{\epsilon, N}}{(m_1^{1, N} \wedge M_\infty^N)} \right)}{\log(\Lambda_0^N)} \right\rfloor \right)_+. \quad (2.32)$$

We can remove “ $()_+$ ” in the last inequality if there exists N_0 such that

$$\forall N \geq N_0, \quad \frac{m_{\epsilon, N}}{(m_1^{1, N} \wedge M_\infty^N)} < 1.$$

Using the inequality (2.27), we deduce that this is the case if

$$C(\epsilon, r, N) = \sqrt{\frac{M_\epsilon^N}{m_1^{1, N} \wedge M_\infty^N}} + 2 \sqrt{\frac{-\log(\frac{\epsilon}{2})}{(m_1^{1, N} \wedge M_\infty^N) N f_N}} < 1. \quad (2.33)$$

From the previous computation of M_ϵ^N , see equation (2.31), we obtain

$$\frac{M_\epsilon^N}{M_\infty^N} = \left(1 - \frac{(1-f_N)a_N}{q^+ + (1-f_N)(a_N + b_N)}\right) \left(1 + a_N f_N \sqrt{\frac{2f_N}{\epsilon(1-\lambda_2^N)}}\right).$$

Thus, we compare the three terms (recalling that $m_1^{1, N} = 1 - (1 - q^+)^r$)

$$\frac{a_N}{q^+ + a_N + b_N}, \quad a_N f_N \sqrt{\frac{2f_N}{\epsilon(1-\lambda_2^N)}} \quad \text{and} \quad \frac{-\log(\frac{\epsilon}{2})}{((1 - (1 - q^+)^r) \wedge M_\infty^N) N f_N}.$$

First, $(1 - \lambda_2^N) \approx_{N \rightarrow \infty} 2f_N^2(a_N + b_N + q^+)$. Then, we have to separate the different cases:

- If b_N tends to $+\infty$, both M_ϵ^N and M_∞^N converge to 0. Hence, $(1 - (1 - q^+)^r) \wedge M_\infty^N = M_\infty^N$ and the Assumption 2.9.3, in particular $\lim_{N \rightarrow \infty} q_{01}^{-,N} = \lim_{N \rightarrow \infty} q_{10}^{-,N} = \lim_{N \rightarrow \infty} \frac{b_N^2}{N f_N a_N} = \lim_{N \rightarrow \infty} \frac{b_N}{N f_N} = 0$, enables us to conclude that if $b_N = \mathcal{O}(a_N)$, $C(\epsilon, r, N) \approx_{N \rightarrow \infty} \sqrt{\frac{b_N}{a_N}} + 2\sqrt{\frac{-\log(\frac{\epsilon}{2})b_N}{q^+ N f_N}} \rightarrow 0$, else, $C(\epsilon, r, N) \approx_{N \rightarrow \infty} \left(1 - \frac{a_N}{a_N + b_N}\right)$ so the inequality (2.33) holds true for any r and for a N large enough.
- If a_N tends to $+\infty$ and not b_N , then M_ϵ^N converges to 0 and M_∞^N converges to 1 (resp. $\frac{q^+}{q^+ + b}$) if b_N converges to 0 (resp. b). Thus, $C(\epsilon, r, N)$ converges to 0 with large N and for any r , the inequality (2.33) is satisfied.
- If a_N tends to 0 and b_N converges to $b > 0$, then M_ϵ^N converges to M_∞^N . Then, there exists r_0 such that $(1 - (1 - q^+)^{r_0}) \wedge M_\infty^N = M_\infty^N$. Using the assumption $\lim_{N \rightarrow \infty} a_N N f_N = +\infty$ we have for all $r \geq r_0$, $C(\epsilon, r, N) \approx_{N \rightarrow \infty} \left(1 - \frac{a_N}{q^+ + b}\right)$.
- In all other cases, M_ϵ^N and M_∞^N converges to a value in $(0, 1)$. Moreover, $M_\epsilon^N < M_\infty^N$ so there exists r_0 such that for all $r \geq r_0$, $(1 - (1 - q^+)^{r_0}) \wedge M_\infty^N > M_\epsilon^N$, so $C(\epsilon, r, N)$ converges in $(0, 1)$ with large N .

□

2.5 Simulations

Our code follows these lines. We draw V_0^N and $K = \sum_{j=2}^{N+1} V_0^{j,N}$. We simulate a trajectory of $h_{t,K}^N$ long enough to be under the invariant measure. We perform r presentations of the signal to be learnt and then compute the trajectories of $h_{t,K}^{y,N}$, $y \in \{0, 1\}$. We reiterate this procedure $N_{MC} = 10^7$ times to get an approximation of the distributions of $h_{t,K}^{0,N}$ and $h_{t,K}^{1,N}$.

The result of Theorem 2.27 is interesting for large values of $N f_N$ (small errors) combined with a small f_N (non-negligible \hat{t}). In this context, we need to compute many trajectories before the synaptic currents cross a reasonable threshold θ .

In Figure 2.4a, the top (resp. bottom) roughly represents the distribution of $h_{t,K}^{1,N}$ (resp. $h_{t,K}^{0,N}$). Before time $t = 50$, the distribution of $h_t^{0,N}$ is highly concentrated in 0. Indeed, looking carefully to Figure 2.4a, we can observe a residue of this high probability (dark blue) for very weak synaptic currents until time $t = 65$, see also Figure 2.5a. This concentration drastically reduces the contrast of the plot. That is why the time axis starts at $t = 50$ in Figure 2.4a. This figure shows that a threshold θ around one hundred is a good choice: it seems to maximises the time for which the threshold estimation holds true. With this threshold, the numerical errors $p_e^{0,N}(t, \theta)$ and $p_e^{1,N}(t, \theta)$ does not exceed 10^{-4} , see Figure 2.4b, before time 15. It is coherent with \hat{t} plotted in Figure 2.4c. Indeed, the time \hat{t} is equal to 12 for errors on the order of 10^{-4} , see Figure 2.4c. Moreover, in Remark 2.31, the result \hat{t} is a maximum between two times. The second one does not depend on the error ϵ (it is called t_c^N in the proof of Theorem 2.27, see equation (2.28)). This explains the plateau starting at an error just before 10^{-3} in Figure 2.4c. Indeed, for this set of parameters and ϵ large enough, the time \hat{t} is equal to t_c^N . Finally, in Figure 2.4d, we note that $p_e^{0,N}$ is above $p_e^{1,N}$ for small values of t . Then, around time $t = 70$, $p_e^{1,N}$ increases quickly until a value close to one whereas $p_e^{0,N}$ stays below 10^{-2} . This is because the majority of the mass of the distribution of $h_{t,K}^{0,N}$

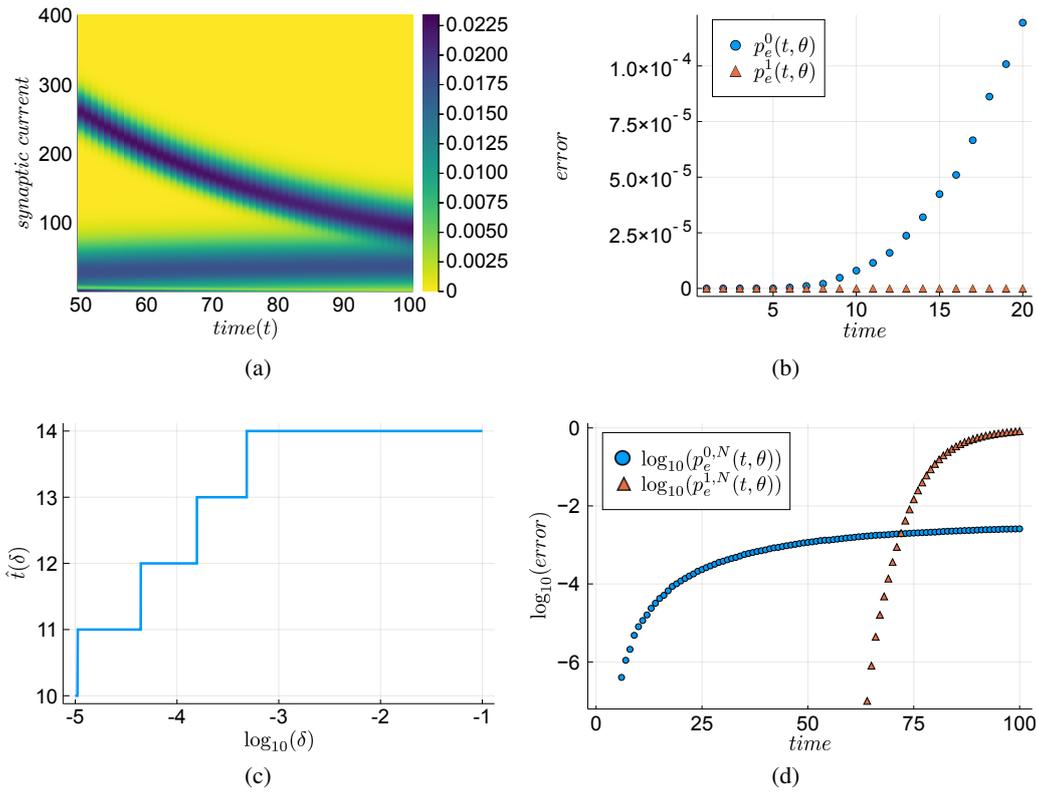


Figure 2.4: (2.4a) The sum of the distributions of $h_{t,K}^{0,N}$ and $h_{t,K}^{1,N}$. The colour bar gives the probability values. (2.4b) (resp. (2.4d)) The numerical errors $p_e^{0,N}(t, \theta)$ and $p_e^{1,N}(t, \theta)$ on a short (resp. large) timescale. (2.4c) \hat{t} as a function of ϵ on the logarithmic to the base ten scale. Parameters: $\theta = 117$, $N = 20\,000$, $f_N = 0.05$, $q^+ = q_{01}^{-,N} = 0.5$, $q_{10}^{-,N} = 0.05$, $r = 3$.

stays less than θ . On the other hand, most of the mass of the distribution of $h_{t,K}^{1,N}$ crosses θ around time $t = 70$. So, the error $p_e^{1,N}$ becomes large. We present the histograms of the distributions of the synaptic currents at certain times.

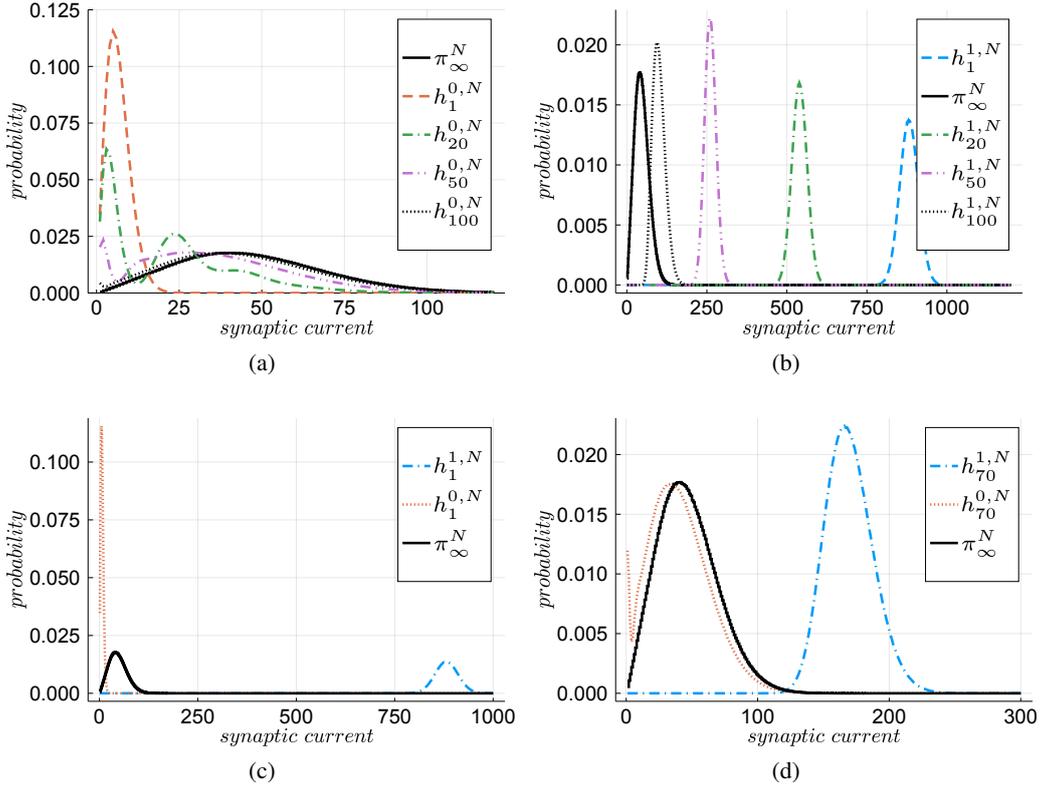


Figure 2.5: (2.5a) Histograms of the distributions of $h_t^{0,N}$ at different times. (2.5b) Histograms of the distributions of $h_t^{1,N}$ at different times. (2.5c) Distributions of $h_1^{1,N}$ just after the learning phase and the invariant measure. (2.5d) Distributions of $h_t^{y,N}$ at $t = 70$ and the invariant measure. Parameters: $N = 20\,000$, $f_N = 0.05$, $q^+ = q_{01}^{-,N} = 0.5$, $q_{10}^{-,N} = 0.05$, $r = 3$.

We note again that the invariant measure is concentrated around small values. This enables the post learning distribution of $h_1^{0,N}$ to have a small variance, see Figures 2.5a and 2.5c. However, the variance of this distribution increases quickly. In particular, the distribution of $h_t^{0,N}$ has a multimodal shape with a high proportion of the mass staying near 0 for more than 50 presentations after learning. On the other hand, the distribution of $h_t^{1,N}$ keeps a unimodal shape with a variance decreasing at the beginning, then increasing before decreasing again, see Figure 2.5b. Distributions stay well separated approximately until time $t = 70$, see Figure 2.5d.

In order to illustrate the role played by the parameter r , we plot the distributions just after the learning phase for different values of r .

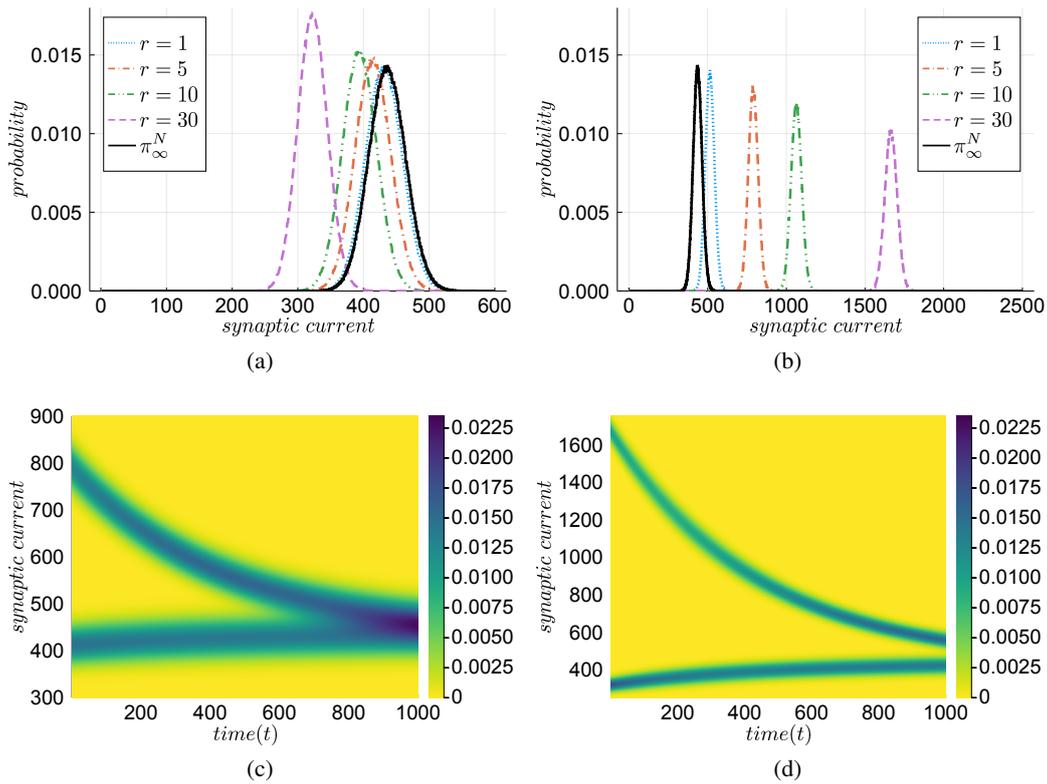


Figure 2.6: (2.6a) Distributions of $h_1^{0,N}$, just after learning, for different values of r and the invariant measure. (2.6b) Distributions of $h_1^{1,N}$, for different values of r and the invariant measure. (2.6c) The sum of the distributions of $h_t^{0,N}$ and $h_t^{1,N}$ for $r = 5$. (2.6d) The sum of the distributions of $h_t^{0,N}$ and $h_t^{1,N}$ for $r = 30$. The colour bar gives the probability values. Parameters: $N_{MC} = 10^6$, $N = 20\,000$, $f_N = 0.1$, $q_{01}^{-,N} = q_{10}^{-,N} = 0.01$ and $q^+ = 0.05$.

Because of the parameters choice, the distributions of $h_1^{0,N}$ are close to the invariant measure π_∞^N whereas the distributions of $h_1^{1,N}$ are further from it, see Figures 2.6a and 2.6b. Moreover, the forgetting is really slow. However, if we want the signal to be learnt correctly with such a small q^+ , then r has to be high enough. This shows the need of a large r in view of a slow forgetting. Figures 2.6c and 2.6d show well the difference brought by a higher value of r : the separation between the two distributions is clearer.

2.6 Discussion

We provide a mathematical framework to study the memory retention of random signals by a recurrent neural network with binary neurons and binary synapses. We thus consider a paradigm linking synaptic plasticity and memory: a stimulus is remembered as long as its trace in the synaptic weights is strong enough. In order to measure the memory of a stimulus, we study the synaptic current onto one neuron during the presentation of this stimulus. First, we compute the spectrum of the transition matrix of the Markov chain associated to the synaptic current. This enables us to conclude that the eigenvalues are strictly different whatever the parameters are. In particular, we can compute the rate of convergence of the chain to its invariant measure, see Corollary 2.13. Then, we carry on the work done by [8] on the dynamics of the distributions of the synaptic current and their invariant distribution. This leads us to control the form of these distributions. Their properties give enough information to find a lower bound on the time a neuron keeps a good estimate on its response to the first stimulus and hence remembers it. We measure the quality of this estimation by performing a statistical test based on the observation of the synaptic current onto one neuron. We define an error associated to this test which depends on two distributions: the distribution of the synaptic current knowing that the neuron was selective to the initial signal and the distribution knowing that the neuron was not selective. Finally, unlike previous studies, we take into account the possibility that heterosynaptic and homosynaptic depressions scale differently in the network size N and we consider the role of presenting several times a signal in the learning phase.

We use the model presented by [5] because of its relative simplicity and its consideration of synapse correlations. Their study focused on the first two moments of the synaptic current. It leads to a result on the memory capacity of the network which depends on a global variable, the so-called signal-to-noise ratio (SNR). In particular, they studied the SNR in the large N asymptotic. They obtained a large SNR when the coding level f_N is low and the depression probabilities are proportional to f_N : $q_{01}^{-,N} \propto q^+ f_N$ and $q_{10}^{-,N} \propto q^+ f_N$. The lowest coding level possible f_N is on the order of $\frac{\log(N)}{N}$ and it gives a memory capacity on the order of $\frac{-1}{\log(\lambda_1^N)} \sim_{N\infty} \frac{1}{f_N^2}$. In [116, 8], they assumed that $q_{10}^{-,N} = 0$ and showed the same result using a Gaussian approximation of the synaptic currents. Under the same assumption as in [5] ($q_{01}^{-,N}, q_{10}^{-,N} \propto q^+ f_N$ and $f_N \rightarrow 0$), our result also predicts a forgetting time on the order of $\frac{-1}{\log(\Lambda_0^N)} \sim_{N\infty} \frac{1}{f_N^2}$, see Theorem 2.30. Moreover, we give a result for depression probabilities not depending on N and our result link the probability of error to the parameters. Note the presence of Λ_0^N in our result rather than λ_1^N in previous studies. This difference comes from our different measure of memory lifetime. The SNR analysis is based on the convergence of the means of the synaptic currents whereas our retrieval criterion requests the knowledge of their entire distributions. Indeed, we search for a memory lifetime obtained with a control on the errors $p_e^{0,N}$ and $p_e^{1,N}$. We conjecture that we could prove similar result as ours with λ_1^N rather than Λ_0^N . Finally, our results do not necessarily need the large N asymptotic. Nevertheless, in this asymptotic, the expression of \hat{t} simplifies, see Remark 2.31.

In this study, we assume that learning is generated by the divergence of the distributions of the synaptic currents $h_t^{0,N}$ and $h_t^{1,N}$ from their invariant distribution, see Figure 2.7. The main role of the number of signal presentations (r) is to separate these two distributions. Indeed, the larger the r , the more separated the support of the mixing distributions $g_1^{0,N}$ and $g_1^{1,N}$ are. In our proofs, we compare $g_t^{1,N}$ to $g^{*,N}$ after showing that as long as $g_t^{1,N}$ is far enough from $g^{*,N}$ it is far enough from $g_t^{0,N}$, see Lemma 2.25. As a consequence, the expression of \hat{t} is an increasing function of $m_1^{1,N} - M_\infty^N$, and so of r .

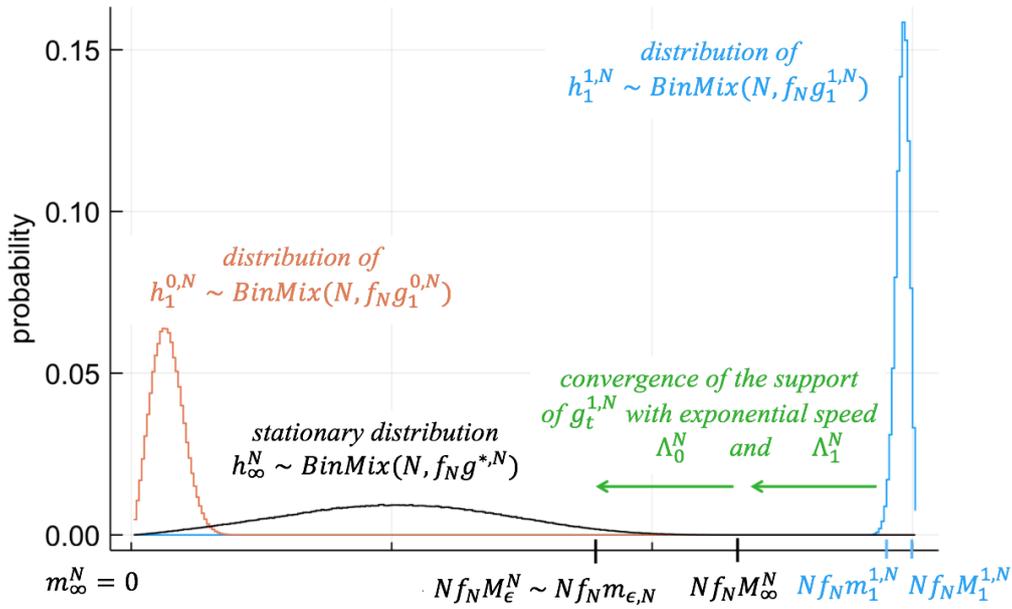


Figure 2.7: Illustration of the notations used in the proofs. Both distributions of $h_t^{0, N}$ (orange) and $h_t^{1, N}$ (blue) converge to the unique stationary distribution π_∞^N (black). The convergence speed of $h_t^{1, N}$ is first Λ_1^N and then Λ_0^N .

Let us now discuss the roles of the coding level, the potentiation and depression probabilities. They affect both learning and forgetting. The coding level directly affects the number of synapses candidate to depression and potentiation. Indeed, looking at an individual synapse, its probability to potentiate is $f_N^2 q^+$ and its probability to depress is $f_N(1 - f_N)(q_{10}^{-, N} + q_{01}^{-, N})$. Thus, when the coding level is close to one, the fluctuations are important and seem to cause a fast forgetting as shown in the illustrations of Section 2.2. Therefore, we used a low coding level, see Assumption 2.9.2. This choice slows down the forgetting. However, f_N cannot be too small because it is detrimental to the learning phase as the distance between the two conditional distributions depends on f_N . More particularly, it depends on $N f_N$ which then need to be large enough, see Assumption 2.9.2. The last parameters we can tune are the potentiation and depression probabilities. As for f_N , there is a compromise between their role in learning and forgetting. Indeed, in order to promote learning, they need to be close enough to one but on the contrary, small probabilities reduce the forgetting rate. So we propose to take a potentiation probability (q^+) on the order of 1, to learn quickly, and small depression probabilities, to forget slowly. Potentiation increases the synaptic currents so it leads to a shift of the distribution of $h_t^{1, N}$ to the right and for the same reasons, depression results in a shift of the distribution of $h_t^{0, N}$ to the left. Therefore, smaller depression than potentiation implies that the distribution of $h_t^{1, N}$ is significantly shifted to the right whereas the distribution of $h_t^{0, N}$ is slightly shifted to the left. In view of learning, the initial separation between distributions can be limited if the invariant distribution π_∞^N is already concentrated on high values of synaptic currents. As there are two depression probabilities, this situation can be avoided by choosing one probability big enough and the other one smaller. For example, when depression probabilities depend on N under Assumption 2.9.3, both $q_{10}^{-, N}$ and $q_{01}^{-, N}$ converge to 0. If they both converge too fast (a_N and b_N converge to 0), the invariant measure is concentrated around one and no learning is possible. However, if either a_N or b_N does not converge to 0, then the invariant measure is not concentrated around 0 and learning is possible. Then, depending on the different large N asymptotic of a_N and b_N , we computed the different memory lifetime sum-

marized in Table 2.1. The best memory lifetimes are on the order of $\frac{1}{f_N^2}$ and are obtained when a_N (resp. b_N) converges to 0 and b_N (resp. a_N) converges to a constant in \mathbb{R}^+ (resp. \mathbb{R}_*^+) or (a_N, b_N) converges to constants in $\mathbb{R}_*^+ \times \mathbb{R}^+$. Thus, if we want to increase the memory lifetime beyond this order, we seem to need a more complex model.

Our study is valid for a classic learning, which needs multiple stimulus presentations, but also for a one shot learning. This last one is possible only with a specific choice of parameters. Indeed, when presenting a stimulus, the synaptic weights between selective neurons need to be potentiated with a high probability (high q^+). When presenting other stimuli, these same weights need to have a very small probability of undergoing depression (low $q_{01}^{-,N}$ and $q_{10}^{-,N}$). As a result, following the presentation of a stimulus, selective neurons develop strong links and then these connections take time to disappear. Thus, the experiment associated with this model would focus on recognition memory. A well-known experiment in this field was carried out by [121]. He showed that humans are able to recognize up to 10,000 images, presented only once, with 90 percent success rate.

Many perspectives can be studied as a follow-up to this study. First, the analysis carried out on the synaptic current onto one neuron could be extended to the entire vector of synaptic currents. The correlations between synaptic weights would then play a major role. In addition, the model could be completed in order to get closer to biology. Indeed, the formation of synaptic memory is far more complex than in our model. In particular, the link between the dynamics of the neurons and the synaptic weight is missing. Improving the model in this direction could be done by considering more structured and complex external signals, adding neural layers and a more realistic membrane potential neural dynamics. In the literature, adding synaptic states does not seem to be successful as the authors stated in [48, 71], whereas meta-plastic transitions brought better SNR results [49, 117, 16]. Adding neural dynamics in such models would be a next challenging step. Nevertheless, the model analysed here illustrates well the trade-off between the plastic and the stable characteristics of memory. Indeed, learning implies changes of synaptic weights (plasticity) as well as mechanisms which maintain them (stability). In mathematical terms, stability is related to the minimal convergence rate and plasticity refers to the sensibility to disturbance. We see that there is a compromise: the more a dynamics is sensitive to disturbances, the less it is stable and vice-versa.

Chapter 3

Slow-Fast Analysis of a New Stochastic STDP model

Abstract

Thought to be responsible for memory, synaptic plasticity has been widely studied in the past few decades. A plasticity mechanism often used to model synaptic plasticity is the Spike-Timing-Dependent Plasticity (STDP). There is a huge literature on STDP models. Their analyses are mainly based on numerical work when only a few has been studied mathematically. Here, we propose a new stochastic STDP rule with discrete synaptic weights. It provides a new framework for using probabilistic tools for an analytical study of plasticity. Moreover, it is biologically plausible and can be easily simulated with thousands of neurons. Based on the assumption that the plasticity is slow compared to the neural network dynamics, we perform a separation of timescale. This separation enables us to derive an equation for the weight dynamics based on the invariant measure of the neural dynamics. The study of long-term behaviour of the weights gives us conditions for its transience and positive recurrence. We illustrate these conditions on simple cases in which we can derive the exact dynamics of the weights thanks to the computation of the Laplace transform of the invariant measure of the fast process. These calculations put forward some counter intuitive results. This model attempts to answer the need for understanding the interplay between the weight and neural dynamics.

3.1 Introduction

At the beginning, plasticity models were based on firing rates [18]. Later on, as suggested by Hebb in his book [67] published in 1949, the crucial role of precise spiking times has been proven experimentally. First in vitro at the end of the twentieth century by Markram et al. in [92] and Bi and Poo in [17], and then in vivo in many cortical areas and different species, see [24] and references therein. This type of plasticity was named Spike-Timing-Dependent Plasticity (STDP) by Song et al. in [120]. Following such a breakthrough, numerous STDP models emerged and their studies helped understanding this plasticity mechanism.

In the earliest studies, the need of bounding the synaptic weights emerged. Different methods have been proposed: from the hard bounds when the weight modification does not depend on its current value (additive STDP) to the soft bounds when this evolution depends on the value of the synaptic weight (multiplicative STDP), see Section 1.2.2 for more details. In the context of one neuron connected to multiple presynaptic neurons, hard bounding leads to the formation of groups of weights at the two extremal bounds with competition between the weights (bimodal distribution), whereas soft bounding leads to a unimodal distribution of the weights, see for instance [79, 120, 118, 78, 65].

In particular, in [79] and [78], a separation of timescales is used. Indeed, long term plasticity is induced on a timescale that ranges from a few minutes to more than one hour. On the other hand, a spike lasts for a few milliseconds [102]. Hence using this timescale difference, the weight dynamics can be simplified by using the averaged behaviour of the neural system. More recently, timescale separation method has been used in numerical study of different recurrent neural networks models. In [58] and [59], the authors have shown the emergence of a network structure and its functional implications. In [107], the authors put forward the need of the precise spiking times for the assembly formation and stability, especially when the spiking rates stay rather constants, see [106]. Despite all these previous works, there is still a need to understand how to bridge this timescale gap between the synapse level and the network one, see [45, 136, 128].

Unfortunately, there is still a lack of rigorous mathematical analyses of such systems. Among deterministic analyses, Perthame et al. analysed the plasticity between populations of LIF neurons interacting through their average activities in [110]. Stochastic studies of STDP models are proposed in [52] and more recently in [114] and [115]. Contrary to the latter, we study the plasticity within a neural network rather than the dynamics of a single synaptic weight and its pre- and postsynaptic neurons.

In this chapter, we present and study our new STDP rule which is implemented in the well-known stochastic Wilson-Cowan model of binary spiking neurons presented in [15]. More precisely, because of the plasticity rule, our model is a Piecewise Deterministic Markov Process (PDMP) [33] and [34] whereas it is a pure point process in [15]. This model is new on three main points. First, instead of a deterministic STDP rule, we introduce a stochastic one: the weights do not change at each spike pair but rather may change or not according to a certain probability depending on the spike pair. Second, the size of the weight jumps is fixed. Our synaptic weights are then discrete as suggested in [108]. Thus, the usual separation of timescale using the assumption of arbitrarily small weight jumps is not possible. Instead, we assume that the probability of such jumps is arbitrarily small. Third, the effect of the spike of a given neuron on the other ones persists for a few milliseconds (as long as the neuron is in state 1) as it was found experimentally. Indeed, it depends on the receptors but the order of magnitude is the millisecond, see [14] for instance.

Thanks to the simplicity of the neural network model, we can perform a mathematical analysis of the Markov process composed of the following three components: the synaptic weight matrix, W_t , the times since the last spike of the neurons, S_t , and the neuron states or potentials, V_t . In particular, assuming that the synaptic dynamics of the weight, W_t , is much slower than the neural network one, (V_t, S_t) , we perform a separation of timescales to obtain the limit dynamics of the weights using results proved by Kurtz in [81]. We show that this dynamics is a Markov chain on the discrete space of weights. Its transitions depend on the invariant measure of the fast process for the current synaptic weight matrix. Once this result is established, we precise the long time behaviour of this chain: conditions of transience and positive recurrence are obtained. In particular, we can use the Laplace transform of the stationary distribution of the fast process to check these conditions. Simulations enable us to illustrate our results and to go beyond them. In particular, we show counter intuitive results such as the divergence of the weights when the STDP curve of potentiation always “wins” against the one of depression. Therefore, this work presents a new biologically plausible model that can be studied mathematically and easily simulated with thousands of neurons.

This chapter is organised as follows. Section 3.2 is devoted to the presentation of the model as well as the introduction of notations. The definition of the generator of the process enables us to explain more formally how the timescale separation is studied. This study is detailed in Section 3.3. First, we show in Proposition 3.8 the exponential convergence of the fast process to its unique invariant measure. Then, we give the main result of this chapter: under the slow-fast assumption, we obtain in Theorem 3.16 the dynamics of the limit process. This result tells us that the limit dynamics of the weights is a non-homogeneous in space Markov chain with kernel transition depending on the invariant measure of the fast process. Once this dynamics is obtained, we study its long time behaviour in Section 3.4. We give simple and intuitive conditions for transience and positive recurrence in Proposition 3.24. After bounding the jump rate of individual weights in Proposition 3.28, we propose a simple condition on the parameters for an example in Proposition 3.31. In the second part of Section 3.4, we show how to compute the Laplace transform of the invariant measure of the fast process, first within a toy minimal model with 2 neurons in Lemma 3.36 and in the general case in Proposition 3.39. Finally, we illustrate our results and go beyond them with simulations in the Section 3.5. In particular, we illustrate the divergence of the weights when the potentiation STDP curve is below the depression one for all value of spike pair delay. Moreover, the stability analysis shows that increasing the size of the network seems to stabilise it. For a fixed set of parameters except the number N of neurons, if the condition of positive recurrence is satisfied for $N = 2$, then, for a larger network, $N > 2$, the weight chain seems to be also positive recurrent. More generally, the set of parameters leading to positive recurrence of the chain for $N = 2$ seems to be included in the one for $N > 2$.

3.2 Presentation of the model and notations

In this chapter, the size N of the network is fixed. To alleviate the notation, we do not mark any parameter with the exponent N .

The model

We use the well-known stochastic Wilson-Cowan model of spiking neurons (see [15] for an analysis with fixed synaptic weights) in which we implement a stochastic STDP rule. We study a net-

work of N binary neurons, $V_t \in \{0, 1\}^N$, all-to-all connected via the matrix of synaptic weights W_t belonging to the grid $(\Delta w \mathbb{Z})^{N^2}$ where $\Delta w \mathbb{Z} \stackrel{\text{def.}}{=} \{m \Delta w : m \in \mathbb{Z}\}$ and $\Delta w > 0$. For all $1 \leq i, j \leq N$, the weight W_t^{ij} represents the effect on the neuron i of a spike emitted by the neuron j at time t . For all $1 \leq i \leq N$, we denote by $W_t^i = (W_t^{1i}, \dots, W_t^{Ni})$ the vector of the postsynaptic weights of the neuron i . A neuron is either excitatory, all its postsynaptic weights are positive, or inhibitory, all its postsynaptic weights are negative. We say that the neuron i spikes at time t if V^i jumps from 0 to 1 at time t , that is $V_t^i = 1$ and $V_{t-}^i = 0$. One notes $S_t^i \in \mathbb{R}^+$ the time spent since the last spike of the neuron i . The dynamics studied in this chapter is a Piecewise Deterministic Markov Process (PDMP). We describe it now.

- The deterministic part between the jumps is very simple, the variables S_t^i increase linearly with time:

$$\forall t \geq 0, \forall i, \quad dS_t^i = dt.$$

- For any i , the neuron i jumps from state 0 to state 1 at rate $\alpha_i(V_t, W_t)$. The function α_i is defined on $\{0, 1\}^N \times \mathbb{R}^{N^2}$ and takes value in $[\alpha_m, \alpha_M]$, with $0 < \alpha_m \leq \alpha_M$. At the spiking time t ,
 - the potential V_t^i jumps from 0 to 1,
 - the clock S_t^i is reset to 0,
 - the presynaptic weights $(W_{t-}^{ij})_{j \neq i}$ are potentiated with probability $p^+(S_{t-}^j, W_{t-}^{ij})$,
 - the postsynaptic weights $(W_{t-}^{ji})_{j \neq i}$ are depressed with probability $p^-(S_{t-}^j, W_{t-}^{ji})$.

Here the functions p^+ and p^- are defined on $\mathbb{R}^+ \times \Delta w \mathbb{Z}$ with values in $[0, 1]$. Potentiation and depression consist of a change of $\pm \Delta w$ in the synaptic weight. We assume moreover that $p^+(s, -\Delta w) = p^-(s, +\Delta w) = 0$, that is the signs of the synaptic weights do not change.

- Finally, the potentials of the neurons jump from 1 to 0 at constant rate $\beta > 0$; at these instants, neither the synaptic weights nor the clocks are modified. Without loss of generality, we assume that

$$0 < \alpha_m < \beta < \alpha_M < \infty. \quad (3.1)$$

We can summarize the previous explanations by giving the dynamics of the potential V_t^i and then the plasticity rule. The potential V_t^i follows the reaction

$$0 \xrightleftharpoons[\beta]{\alpha_i(V_t, W_t)} 1. \quad (3.2)$$

When the neuron i spikes at time t , then S_t^i is reset to 0 and for all $j \neq i$:

$$\mathbb{P}(W_t^{ij} = W_{t-}^{ij} + \Delta w) = p^+(S_{t-}^j, W_{t-}^{ij}) \quad \text{and} \quad \mathbb{P}(W_t^{ji} = W_{t-}^{ji} - \Delta w) = p^-(S_{t-}^j, W_{t-}^{ji}).$$

From the definition of the model, we see that the weights are discrete. In the following, we assume that every neuron is excitatory and without loss of generality, we set $\Delta w = 1$ and W_0 belongs to the space E_2 such that:

$$E_2 \stackrel{\text{def.}}{=} \{w \in \mathbb{N}^{N^2} \text{ s.t. } \forall i, w^{ii} = 0 \text{ and } \forall i \neq j, w^{ij} > 0\}. \quad (3.3)$$

Remark 3.1. *The synaptic weights between two different neurons are positive, so we consider an all-to-all connected network of excitatory neurons. Extending some of our results to inhibitory-excitatory neurons without more mathematical work is possible, see Remark 3.18. Also, we note that W_t^{ii} is null for all i and t .*

Finally, let us remark that the process $(V_t, W_t)_{t \geq 0}$ is not Markovian but the process $(V_t, S_t, W_t)_{t \geq 0}$ is Markovian.

We now study the continuous time Markov process $(V_t, S_t, W_t)_{t \geq 0} \in E \stackrel{\text{def.}}{=} E_1 \times E_2$ where $E_1 \stackrel{\text{def.}}{=} \{0, 1\}^N \times \mathbb{R}_+^N$ and E_2 is defined in (3.3). Here, the difference of timescales between the weight dynamics and the neural network one has piloted the definitions of E_1 and E_2 .

Notation 3.2. *We denote by $(\Omega, \mathcal{B}(E), \mathbb{P})$ the probability space on which the Markov process $(V_t, S_t, W_t)_{t \geq 0}$ is defined.*

Generator of the process

For all $w \in E_2$, assuming that $W_{t-} = w$, we denote by A_w^i the accessible states of W_t after the spike of neuron i at time t ,

$$A_w^i \stackrel{\text{def.}}{=} \left\{ \tilde{w} \in E_2 : \forall j \neq i, \tilde{w}^{ij} \in \{w^{ij}, w^{ij} + 1\}, \tilde{w}^{ji} \in \{w^{ji}, w^{ji} - 1\} \text{ and } k \neq i, \tilde{w}^{jk} = w^{jk} \right\}.$$

We denote by $\phi_i : \mathbb{R}_+^N \times E_2 \times E_2 \rightarrow [0, 1]$ the transition of W conditionally to the vector S_{t-} : for all $(s, \tilde{w}, w) \in \mathbb{R}_+^N \times E_2 \times E_2$,

$$\begin{aligned} \phi_i(s, \tilde{w}, w) &= \mathbb{P}(W_t = \tilde{w} | W_{t-} = w, S_{t-} = s) \\ &= \mathbb{1}_{A_w^i}(\tilde{w}) \prod_{j \neq i} \left([(\tilde{w}^{ij} - w^{ij})p^+(s^j, w^{ij}) + (1 - (\tilde{w}^{ij} - w^{ij}))(1 - p^+(s^j, w^{ij}))] \right. \\ &\quad \left. [(w^{ji} - \tilde{w}^{ji})p^-(s^j, w^{ji}) + (1 - (w^{ji} - \tilde{w}^{ji}))(1 - p^-(s^j, w^{ji}))] \right). \end{aligned} \tag{3.4}$$

Remark 3.3. *We obtain this transition using the fact that the weight changes are independent one to each other.*

In this chapter, when we refer to continuous or more regular functions, the regularity properties hold on the continuous variable S_t living on \mathbb{R}_+^N . Although the domain of the operators we use can be larger, we consider their restriction to the space of continuous and bounded functions on the same state space. This means that if we note $\mathcal{D}(\mathcal{C})$ the largest domain of definition of the operator \mathcal{C} acting on functions on E , we will consider $D(\mathcal{C}) = \mathcal{D}(\mathcal{C}) \cap C_b(E)$.

We denote by $(\mathcal{C}, D(\mathcal{C}))$ the generator of the complete process $(V_t, S_t, W_t)_{t \geq 0}$ where $D(\mathcal{C})$ is the domain of definition of \mathcal{C} . Thus, \mathcal{C} is given by: for all $f \in D(\mathcal{C})$ and $(v, s, w) \in E$,

$$\begin{aligned} \mathcal{C}f(v, s, w) &= \sum_i \partial_{s^i} f(v, s, w) + \sum_i \delta_1(v^i) \beta [f(v - e_i, s, w) - f(v, s, w)] \\ &\quad + \sum_i \delta_0(v^i) \alpha_i(v, w) [f(v + e_i, s - s^i e_i, w) - f(v, s, w)] \phi_i(s, w, w) \\ &\quad + \sum_i \delta_0(v^i) \alpha_i(v, w) \left(\sum_{\tilde{w} \neq w} [f(v + e_i, s - s^i e_i, \tilde{w}) - f(v, s, w)] \phi_i(s, \tilde{w}, w) \right), \end{aligned}$$

where we denote by $e_i = (0, \dots, 0, \underbrace{1}_i, 0, \dots, 0)$ the elements of both $\{0, 1\}^N$ and \mathbb{R}_+^N .

Remark 3.4. We note that $D(\mathcal{C})$ contains the space of bounded and continuously differentiable functions $C_b^1(E)$.

We have split the generator into four terms to facilitate their interpretation. It contains three interacting dynamics:

- the dynamics of the time since the last spike (first term),
- the neural dynamics (second and third term),
- the dynamics of the weights (last term).

We moreover assume that the synaptic weight dynamics is slow compared to the neural network dynamics. As seen in the introduction, this assumption is experimentally supported. Thus, $(V_t, S_t)_{t>0}$ evolves on a faster timescale than $(W_t)_{t>0}$. This assumption can be written more formally as:

$$\forall i \in \llbracket 1, N \rrbracket, \quad \sum_{\tilde{w} \neq w} \phi_i(s, \tilde{w}, w) \ll \phi_i(s, w, w). \quad (3.5)$$

Typically, this means that for small $\varepsilon \in (0, 1)$, we have $\sum_{\tilde{w} \neq w} \phi_i(s, \tilde{w}, w) \approx \varepsilon$ and $\phi_i(s, w, w) \approx 1 - \varepsilon$. This timescale difference is studied in Section 3.3.2. The fast part of the process is studied in Section 3.3.1.

3.3 Slow-fast analysis

First, we show that the fast process has a unique invariant measure. The synaptic weight dynamics is then derived as a function of this invariant measure under the slow-fast assumption.

3.3.1 The fast process invariant measure

Notations

In this subsection, we fix a weight matrix $W_0 = w \in E_2$ and assume that $p^+ \equiv p^- \equiv 0$. We deduce that $W_t \equiv w$ and study the two first components (V_t, S_t) of the process $(V_t, S_t, W_t)_{t \geq 0}$. Hence, for all $w \in E_2$, its generator $(\mathcal{B}_w, D(\mathcal{B}_w))$ is given by: for any $f \in D(\mathcal{B}_w)$ and $(v, s) \in E_1$,

$$\begin{aligned} \mathcal{B}_w f(v, s) = & \sum_{i=1}^N \partial_{s^i} f(v, s) + \sum_{i=1}^N \delta_1(v^i) \beta [f(v - e_i, s) - f(v, s)] \\ & + \sum_{i=1}^N \delta_0(v^i) \alpha_i(v, w) [f(v + e_i, s - s^i e_i) - f(v, s)]. \end{aligned} \quad (3.6)$$

The process $(V_t, S_t)_{t \geq 0}$ is a classic PDMP with two components: a discrete one, $V_t \in \{0, 1\}^N$, and a continuous one, $S_t \in \mathbb{R}_+^N$. A typical trajectory of one of its coordinates is illustrated in Figure 3.1.

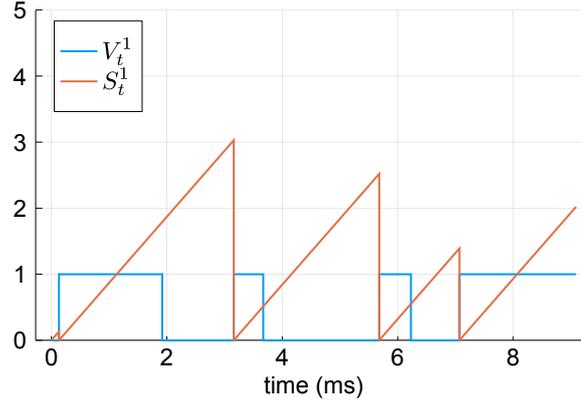


Figure 3.1: A typical trajectory of V_t^1 in blue and S_t^1 in red.

Notation 3.5. We denote by $(P_{t,w})_{t \geq 0}$ the transition probability kernel of the process $(V_t, S_t)_{t \geq 0}$. In particular, for all $x \in E_1$, $A \in \mathcal{B}(E_1)$ and $w \in E_2$, we have

$$P_{t,w}(x, A) = \mathbb{P}((V_t, S_t) \in A | (V_0, S_0, W_0) = (x, w)).$$

For a measure μ on E_1 , a function $f \in C_b(E_1)$, a Borel set $A \in \mathcal{B}(E_1)$ and $x \in E_1$, we use the classical notations:

$$\mu P_{t,w}(A) = \int_{E_1} P_{t,w}(y, A) \mu(dy) \quad \text{and} \quad P_{t,w}f(x) = \int_{E_1} f(y) P_{t,w}(x, dy).$$

Finally, we introduce the number of jumps of the processes.

Notation 3.6. Let $(N_{t,w}^i)_{t \geq 0}$ be the number of jumps of the process $(V_u^i, S_u^i)_{u \geq 0}$ between the times 0 and t . This counter can be formally defined with a family $(\zeta^i(du, dz))_{1 \leq i \leq N}$ of N Poisson random measures on \mathbb{R}_+^2 with intensity $dudz$ (see Definition A.5 in the appendix),

$$N_{t,w}^i = \int_0^t \int_{\mathbb{R}^+} \mathbb{1} \left\{ z \leq \alpha_i(V_{u-}, w) \mathbb{1}_{\{V_{u-}^i=0\}} \right\} + \mathbb{1} \left\{ z \leq \beta^1 \mathbb{1}_{\{V_{u-}^i=1\}} \right\} \zeta^i(du, dz).$$

Convergence of the fast process to its invariant measure

We use the classical Doeblin condition, given by (3.7), to prove the exponential convergence to the unique invariant measure. Under the Doeblin condition, a simple and short proof shows that the probability transition kernel of a Markov process in a general metric state space is contracting. This condition is very useful when the state space is compact or when a compact set of the state space is uniformly accessible from any of its element, see [97, Thm 16.0.2] and [11]. We give the associated Doeblin Theorem. Its proof can be found in [50, Thm 11] and in [26, Thm 2.3], we recall it here for completeness. Then, we show that the process $(V_t, S_t)_{t \geq 0}$ satisfies the Doeblin condition.

In what follows, we use the Jordan decomposition of the signed measures and their total variation, see Definition A.2 in the appendix.

Theorem 3.7. (Doebelin Theorem) Let $(Y_t)_{t \geq 0}$ be a Markov process on a measurable set (X, \mathcal{X}) . Its Markov transition kernel is denoted by $(P_t)_{t \geq 0}$. Assume that there exist $T > 0$, $\kappa \in (0, 1)$ and a probability measure ν on X such that for all $x \in X$,

$$\inf_{x \in X} P_T(x, \cdot) \geq \kappa \nu(\cdot). \quad (3.7)$$

Then, the process $(Y_t)_{t \geq 0}$ admits a unique invariant measure μ_∞ .

Moreover, denoting by $\gamma = -\log(1 - \kappa) > 0$, we have for all probability measure μ on X ,

$$\|\mu P_t - \mu_\infty\|_{TV} \leq 2e^{-\gamma \lfloor \frac{t}{T} \rfloor}.$$

Proof. Let μ_1 and μ_2 be two probability measures on X such that $\|\mu_1 - \mu_2\|_{TV} > 0$. One has,

$$\|(\mu_1 - \mu_2)_+\|_{TV} = \|(\mu_1 - \mu_2)_-\|_{TV} = \frac{1}{2} \|\mu_1 - \mu_2\|_{TV}.$$

We denote $\hat{\mu} = \frac{\mu_1 - \mu_2}{\frac{1}{2} \|\mu_1 - \mu_2\|_{TV}}$. Then, $\hat{\mu}_+$, $\hat{\mu}_-$ are probability measures.

From the triangle inequality and (3.7), we get that for any Borel set A of X ,

$$\begin{aligned} |\hat{\mu} P_T|(A) &= |\hat{\mu}_+ P_T - \hat{\mu}_- P_T|(A) \leq |\hat{\mu}_+ P_T - \kappa \nu|(A) + |\kappa \nu - \hat{\mu}_- P_T|(A) \\ &\leq \hat{\mu}_+ P_T(A) - \kappa \nu(A) + \hat{\mu}_- P_T(A) - \kappa \nu(A). \end{aligned}$$

In particular, for $A = X$ we obtain that

$$\|\hat{\mu} P_T\|_{TV} \leq \|\hat{\mu}_+ P_T\|_{TV} + \|\hat{\mu}_- P_T\|_{TV} - 2\kappa \|\nu\|_{TV} = 2(1 - \kappa).$$

We deduce that

$$\|\mu_1 P_T - \mu_2 P_T\|_{TV} \leq \|\mu_1 - \mu_2\|_{TV} (1 - \kappa).$$

The space of signed measures embedded with the norm $\|\cdot\|_{TV}$ is a Banach space, see [4, Rk 1.7]. Thus, for any initial probability measure μ , the sequence $(\mu P_{nT})_{n \geq 0}$ converges to a unique invariant probability measure (fixed point of P_T) that we denote by μ_∞ . Moreover, from the semi group property of $(P_t)_{t \geq 0}$ we have that for all $t \geq 0$,

$$\mu_\infty P_t P_T = \mu_\infty P_{t+T} = \mu_\infty P_T P_t = \mu_\infty P_t.$$

Thus, $\mu_\infty P_t$ is also invariant for P_T and by uniqueness of μ_∞ , we have $\mu_\infty P_t = \mu_\infty$. Hence, μ_∞ is also invariant for P_t for all $t \geq 0$. We conclude that for all t , denoting by $n = \lfloor \frac{t}{T} \rfloor$, we have

$$\|\mu P_t - \mu_\infty\|_{TV} = \|\mu P_t - \mu_\infty P_t\|_{TV} \leq \|\mu P_{t-nT} - \mu_\infty P_{t-nT}\|_{TV} (1 - \kappa)^n \leq 2(1 - \kappa)^n.$$

□

Proposition 3.8. For all $w \in E_2$, the process $(V_t, S_t)_{t \geq 0}$ with generator \mathcal{B}_w defined in (3.6) converges exponentially fast to its unique invariant measure π_w .

Proof. For all $T > 0$, we find a probability measure with support on $(1, \dots, 1) \times [0, T]^N$ such that the Doeblin condition can be derived for $P_{T,w}$.

For any sequence $t_1, \dots, t_N \in (0, T)$, for all $x = (v, s) \in E_1$ and $(V_0, S_0) = x$, (V_T, S_T) belongs to $(1, \dots, 1) \times \prod_{i=1}^N [0, t_i]$ as long as for all i :

- if $V_0^i = 1$, the neuron i jumps exactly twice in $[T - t_i, T]$ and does not jump in $[0, T - t_i)$,
- if $V_0^i = 0$, the neuron i jumps exactly once in $[T - t_i, T]$ and does not jump in $[0, T - t_i)$.

Note that $(V_T, S_T) \in (1, \dots, 1) \times \prod_{i=1}^N [0, t_i]$ can be obtained in many other ways but the one we give is sufficient to get what we need:

$$\mathbb{P}_x \left((V_T, S_T) \in (1, \dots, 1) \times \prod_{i=1}^N [0, t_i] \right) \geq \mathbb{P}_x \left(\bigcap_i \{N_{T-t_i, w}^i = 0, N_{T, w}^i - N_{T-t_i, w}^i = v^i + 1\} \right).$$

The right term is strictly positive as, by assumption (3.1), the jump rates are bounded by strictly positive constants. Moreover, it does not depend on the initial value $S_0 = s$ because the jump rates does not depend on S_t . Thus, we can fix $S_0 = (0, \dots, 0)$ for instance, and then define the measure $\tilde{\nu}$ on E_1 such that $\tilde{\nu}(E_1 \setminus (1, \dots, 1) \times [0, T]^N) = 0$ and

$$\tilde{\nu} \left((1, \dots, 1) \times \prod_{i=1}^N [0, t_i] \right) = \min_{v \in \{0, 1\}^N} \mathbb{P}_{(v, (0, \dots, 0))} \left(\bigcap_i \{N_{T-t_i, w}^i = 0, N_{T, w}^i - N_{T-t_i, w}^i = v^i + 1\} \right).$$

Thus, $\tilde{\nu}(E_1) \in (0, 1)$ and we conclude, denoting by ν the probability measure $\nu \stackrel{\text{def.}}{=} \frac{\tilde{\nu}}{\tilde{\nu}(E_1)}$, that for all $A \in \mathcal{B}(E_1)$, $x \in E_1$,

$$P_{T, w}(x, A) \geq \tilde{\nu}(E_1) \nu(A). \quad (3.8)$$

Applying Theorem 3.7 ends the proof. \square

3.3.2 The slow process limit dynamics

We first introduce all the notations indexed by ε that we need to then give the main result and its proof.

Notations indexed by ε

We recall the main assumption: the synaptic weight dynamics is slow compared to the network dynamics, in the sense that the process $(V_t, S_t)_{t \geq 0}$ changes fast compared to $(W_t)_{t \geq 0}$. Hence, in order to make a slow fast analysis, we replace the function ϕ_i defined by (3.4) by two families ϕ_i^ε of functions indexed by $\varepsilon \in (0, 1)$. With these notations, we can give a rigorous sense to (3.5). The first family is obtained by changing the functions p^+ and p^- in the definition (3.4) of ϕ_i .

Assumption 3.9. *Let us assume that the functions p^+ and p^- are of order ε . With a slight abuse of notation, we denote by εp^+ and εp^- these functions, that is we consider ϕ_i^ε defined by (3.4) where p^+ and p^- have been replaced by εp^+ and εp^- :*

$$\phi_i^\varepsilon(s, \tilde{w}, w) = \mathbb{1}_{A_w^i}(\tilde{w}) \prod_{j \neq i} \left([(\tilde{w}^{ij} - w^{ij}) \varepsilon p^+(s^j, w^{ij}) + (1 - (\tilde{w}^{ij} - w^{ij})) (1 - \varepsilon p^+(s^j, w^{ij}))] \right. \\ \left. [(w^{ji} - \tilde{w}^{ji}) \varepsilon p^-(s^j, w^{ji}) + (1 - (w^{ji} - \tilde{w}^{ji})) (1 - \varepsilon p^-(s^j, w^{ji}))] \right).$$

The second family is obtained from the model described in Section 3.2 by drawing a Bernoulli variable of parameter ε at each spiking time of a neuron. If this variable is equal to one, then the weights jump as described by the model, otherwise the weights do not jump.

Assumption 3.10. Let us assume that we keep the initial functions p^+ , p^- and define for all $\varepsilon \in (0, 1)$, for all $s \in \mathbb{R}_+^N$,

$$\forall \tilde{w} \neq w, \quad \phi_i^\varepsilon(s, \tilde{w}, w) = \varepsilon \phi_i(s, \tilde{w}, w) \quad \text{and} \quad \phi_i^\varepsilon(s, w, w) = 1 - \varepsilon(1 - \phi_i(s, w, w)).$$

The following Proposition holds either under Assumption 3.9 or Assumption 3.10. In what follows, we only use the results of this Proposition and thus, we just have to assume that one of these two Assumptions holds. In particular, this Proposition gives some general properties of the functions ϕ_i^ε .

Proposition 3.11. Under Assumption 3.9 or 3.10, ϕ_i^ε satisfies the following properties.

For all $s \in \mathbb{R}_+^N$, $\phi_i^\varepsilon(s, \cdot, \cdot)$ is a transition kernel on E_2 such that for all $\tilde{w}, w \in E_2$,

$$\sum_{\tilde{w} \neq w} \phi_i^\varepsilon(s, \tilde{w}, w) = R_i^\varepsilon(s, w) = 1 - \phi_i^\varepsilon(s, w, w), \quad (3.9)$$

where for all $(s, w) \in \mathbb{R}_+^N \times E_2$, $R_i^\varepsilon(s, w) = \mathcal{O}(\varepsilon)$. As for the functions ϕ_i , ϕ_i^ε is null except on the set A_w^i . Moreover, there exist functions φ_i and K_i^ε such that for all $\tilde{w} \neq w$, $s \in \mathbb{R}_+^N$,

$$\phi_i^\varepsilon(s, \tilde{w}, w) = \varepsilon \varphi_i(s, \tilde{w}, w) + K_i^\varepsilon(s, \tilde{w}, w), \quad (3.10)$$

where for all $(s, \tilde{w}, w) \in \mathbb{R}_+^N \times E_2 \times E_2$, $K_i^\varepsilon(s, \tilde{w}, w) = \mathcal{O}(\varepsilon)$. Finally, for all $j \neq i$,

$$\sum_{\substack{\tilde{w} \in A_w^i, \\ \tilde{w}^{ij} = w^{ij+1}}} \varphi_i(s, \tilde{w}, w) = p^+(s^j, w^{ij}) \quad \text{and} \quad \sum_{\substack{\tilde{w} \in A_w^i, \\ \tilde{w}^{ji} = w^{ji-1}}} \varphi_i(s, \tilde{w}, w) = p^-(s^j, w^{ji}). \quad (3.11)$$

Proof. First, under Assumption 3.9, denoting by $E_{ij} \stackrel{\text{def.}}{=} (\delta_{ik}\delta_{jl})_{1 \leq k, l \leq N}$, for all $(s, w) \in \mathbb{R}_+^N \times E_2$, we have

$$\phi_i^\varepsilon(s, w + E_{ij}, w) = \varepsilon p^+(s^j, w^{ij})(1 - \varepsilon p^-(s^j, w^{ji})) \prod_{k \neq i, j} (1 - \varepsilon p^+(s^k, w^{ik}))(1 - \varepsilon p^-(s^k, w^{ki})),$$

$$\phi_i^\varepsilon(s, w - E_{ji}, w) = (1 - \varepsilon p^+(s^j, w^{ij})) \varepsilon p^-(s^j, w^{ji}) \prod_{k \neq i, j} (1 - \varepsilon p^+(s^k, w^{ik}))(1 - \varepsilon p^-(s^k, w^{ki})).$$

Thus, $\phi_i^\varepsilon(s, w + E_{ij}, w) = \varepsilon p^+(s^j, w^{ij}) + \mathcal{O}(\varepsilon)$ and $\phi_i^\varepsilon(s, w - E_{ji}, w) = \varepsilon p^-(s^j, w^{ji}) + \mathcal{O}(\varepsilon)$. For all other $\tilde{w} \in A_w^i$ such that $\tilde{w} \neq w$, the jump from w to \tilde{w} corresponds to at least two changes in the synaptic weights and thus, $\phi_i^\varepsilon(s, \tilde{w}, w) = \mathcal{O}(\varepsilon)$.

Therefore, we define the functions K_i^ε by, for all $j \neq i$,

$$K_i^\varepsilon(s, w + E_{ij}, w) = \phi_i^\varepsilon(s, w + E_{ij}, w) - \varepsilon p^+(s^j, w^{ij}),$$

$$K_i^\varepsilon(s, w - E_{ji}, w) = \phi_i^\varepsilon(s, w - E_{ji}, w) - \varepsilon p^-(s^j, w^{ji}),$$

and $K_i^\varepsilon(s, \tilde{w}, w) = \phi_i^\varepsilon(s, \tilde{w}, w)$ for all other $\tilde{w} \neq w$.

Finally, we define the functions φ_i by: for all $j \neq i$, $s \in \mathbb{R}_+^N$, $w \in E_2$,

$$\varphi_i(s, w + E_{ij}, w) = p^+(s^j, w^{ij}), \quad \varphi_i(s, w - E_{ji}, w) = p^-(s^j, w^{ji}), \quad (3.12)$$

and $\varphi_i(s, \tilde{w}, w) = 0$ for all $\tilde{w} \in A_w^i \setminus \bigcup_j \{w + E_{ij}, w - E_{ji}\}$.

We conclude that, under Assumption 3.9, with such functions φ_i and K_i^ε , all the equations (3.9), (3.10) and (3.11) hold. The proposition is thus satisfied for Assumption 3.9.

Under Assumption 3.10, we have $\varphi_i = \phi_i$ and the proposition is trivially satisfied. \square

Remark 3.12. Under Assumption 3.9 or 3.10, there exists $\varphi^M \in \mathbb{R}^+$ such that for all i ,

$$\sum_{\tilde{w} \neq w} \varphi_i(s, \tilde{w}, w) \leq \varphi^M.$$

In what follows, we assume that either Assumption 3.9 or Assumption 3.10 holds.

We now highlight the difference of timescale in the new generator \mathcal{C}^ε which is the same as \mathcal{C} with ϕ_i^ε instead of ϕ_i . For all $f \in D(\mathcal{C}^\varepsilon)$ and $(v, s, w) \in E$,

$$\begin{aligned} \mathcal{C}^\varepsilon f(v, s, w) &= \sum_i \partial_{s^i} f(v, s, w) + \sum_i \delta_1(v^i) \beta [f(v - e_i, s, w) - f(v, s, w)] \\ &\quad + \sum_i \delta_0(v^i) \alpha_i(v, w) [f(v + e_i, s - s^i e_i, w) - f(v, s, w)] \underbrace{\phi_i^\varepsilon(s, w, w)}_{1 - R_i^\varepsilon(s, w)} \\ &\quad + \sum_i \delta_0(v^i) \alpha_i(v, w) \left(\sum_{\tilde{w} \neq w} [f(v + e_i, s - s^i e_i, \tilde{w}) - f(v, s, w)] \underbrace{\phi_i^\varepsilon(s, \tilde{w}, w)}_{\varepsilon \varphi_i(s, \tilde{w}, w) + K_i^\varepsilon(s, \tilde{w}, w)} \right). \end{aligned}$$

We denote by \mathcal{B}_{net} , \mathcal{B}_{syn} , $\mathcal{B}_r^\varepsilon$ and $\mathcal{B}_\uparrow^\varepsilon$ the following operators:

$$\begin{aligned} \mathcal{B}_{net} f(v, s, w) &= \sum_i \partial_{s^i} f(v, s, w) + \sum_i \delta_1(v^i) \beta [f(v - e_i, s, w) - f(v, s, w)] \\ &\quad + \sum_i \delta_0(v^i) \alpha_i(v, w) [f(v + e_i, s - s^i e_i, w) - f(v, s, w)] \\ \mathcal{B}_{syn} f(v, s, w) &= \sum_i \delta_0(v^i) \alpha_i(v, w) \sum_{\tilde{w} \neq w} [f(v + e_i, s - s^i e_i, \tilde{w}) - f(v, s, w)] \varphi_i(s, \tilde{w}, w), \\ \mathcal{B}_r^\varepsilon f(v, s, w) &= \sum_i \delta_0(v^i) \alpha_i(v, w) \sum_{\tilde{w} \neq w} [f(v + e_i, s - s^i e_i, \tilde{w}) - f(v, s, w)] K_i^\varepsilon(s, \tilde{w}, w), \\ \mathcal{B}_\uparrow^\varepsilon f(v, s, w) &= - \sum_i \delta_0(v^i) \alpha_i(v, w) [f(v + e_i, s - s^i e_i, w) - f(v, s, w)] R_i^\varepsilon(s, w). \end{aligned} \tag{3.13}$$

Thus, one has

$$\mathcal{C}^\varepsilon = \varepsilon \mathcal{B}_{syn} + \mathcal{B}_{net} + \mathcal{B}_\uparrow^\varepsilon + \mathcal{B}_r^\varepsilon.$$

On this timescale, the network evolves at “speed” 1 and the plasticity at “speed” ε . We introduce the accelerated system

$$(V_t^\varepsilon, S_t^\varepsilon, W_t^\varepsilon) = (V_{t/\varepsilon}, S_{t/\varepsilon}, W_{t/\varepsilon}).$$

Its generator is

$$\frac{1}{\varepsilon} \mathcal{C}^\varepsilon = \mathcal{B}_{syn} + \frac{1}{\varepsilon} (\mathcal{B}_{net} + \mathcal{B}_\uparrow^\varepsilon + \mathcal{B}_r^\varepsilon). \tag{3.14}$$

In the following, we shall use some operators acting only on the fast variables and others only on the slow variables. To do so, we introduce

Notation 3.13. We denote by $\mathcal{B}_{net}^1 : D(\mathcal{B}_{net}^1) \subset C_b(E_1) \rightarrow C_b(E)$ and $\mathcal{B}_{syn}^2 : D(\mathcal{B}_{syn}^2) \subset C_b(E_2) \rightarrow C_b(E)$ two operators such that for all $(v, s, w) \in E$, for all $f_1 \in D(\mathcal{B}_{net}^1)$ and $g_2 \in D(\mathcal{B}_{syn}^2)$,

$$\mathcal{B}_{net}^1 f_1(v, s, w) = \mathcal{B}_{net} f(v, s, w) \quad \text{and} \quad \mathcal{B}_{syn}^2 g_2(v, s, w) = \mathcal{B}_{syn} g(v, s, w),$$

where the functions f and g are defined by

$$\forall (v, s) \in E_1, \forall w \in E_2, \quad f(v, s, w) = f_1(v, s) \quad \text{and} \quad g(v, s, w) = g_2(w). \tag{3.15}$$

Remark 3.14. In particular, with these notations, we obtain that for $f_1 \in D(\mathcal{B}_{net}^1)$ and $g_2 \in D(\mathcal{B}_{syn}^2)$, f and g defined by (3.15),

$$\begin{aligned} \mathcal{C}^\varepsilon f(v, s, w) &= \sum_{i=1}^N \partial_{s^i} f_1(v, s) + \sum_i \delta_1(v^i) \beta [f_1(v - e_i, s) - f_1(v, s)] \\ &\quad + \sum_i \delta_0(v^i) \alpha_i(v, w) (f_1(v + e_i, s - s^i e_i) - f_1(v, s)) \phi_i^\varepsilon(s, w, w) \\ &\quad + \sum_i \delta_0(v^i) \alpha_i(v, w) (f_1(v + e_i, s - s^i e_i) - f_1(v, s)) \sum_{\tilde{w} \neq w} \phi_i^\varepsilon(s, \tilde{w}, w) \\ &= \mathcal{B}_{net}^1 f_1(v, s, w), \end{aligned}$$

and

$$\mathcal{C}^\varepsilon g(v, s, w) = (\varepsilon \mathcal{B}_{syn} + \mathcal{B}_r^\varepsilon) g(v, s, w) = \varepsilon \mathcal{B}_{syn}^2 g_2(v, s, w) + \mathcal{B}_r^\varepsilon g(v, s, w).$$

We introduce the notation ν_t^{ε, g_2} as the rest term

$$\nu_t^{\varepsilon, g_2} = \int_0^t \frac{1}{\varepsilon} \mathcal{B}_r^\varepsilon g(V_u^\varepsilon, S_u^\varepsilon, W_u^\varepsilon) du.$$

We prove with Theorem 3.16 that ν_t^{ε, g_2} tends to 0 with ε .

Using these notations, the two following processes

$$\begin{aligned} \varepsilon f(V_t^\varepsilon, S_t^\varepsilon, W_t^\varepsilon) - \int_0^t \mathcal{C}^\varepsilon f(V_u^\varepsilon, S_u^\varepsilon, W_u^\varepsilon) du &= \varepsilon f_1(V_t^\varepsilon, S_t^\varepsilon) - \int_0^t \mathcal{B}_{net}^1 f_1(V_u^\varepsilon, S_u^\varepsilon, W_u^\varepsilon) du \\ g(V_t^\varepsilon, S_t^\varepsilon, W_t^\varepsilon) - \int_0^t \frac{1}{\varepsilon} \mathcal{C}^\varepsilon g(V_u^\varepsilon, S_u^\varepsilon, W_u^\varepsilon) du &= g_2(W_t^\varepsilon) - \int_0^t \mathcal{B}_{syn}^2 g_2(V_u^\varepsilon, S_u^\varepsilon, W_u^\varepsilon) du + \nu_t^{\varepsilon, g_2} \end{aligned}$$

are martingales for the natural filtration of the process $(V_t^\varepsilon, S_t^\varepsilon, W_t^\varepsilon)_{t \geq 0}$.

Let us note that for all $w \in E_2$, $f_1 \in D(\mathcal{B}_w)$, the operator \mathcal{B}_w defined by (3.6) satisfies $\mathcal{B}_w f_1(v, s) = \mathcal{B}_{net}^1 f_1(v, s, w)$. We previously showed that $(V_t, S_t)_{t \geq 0}$ converges to its unique invariant measure π_w . Hence, we can expect that, as ε tends to 0, the fast part quickly reaches its stationary distribution (which depends on the current weights) and drives the weight jumps. As soon as the weights change, the network will reach a new stationary distribution instantaneously and so forth.

We end this section of notations by giving the same following definitions as in [81].

Definition 3.15. Let X be a complete separable metric space.

We denote by $\mathcal{L}_m(X)$ the space of measures η on $\mathbb{R}_+ \times X$ such that for all $t \geq 0$, we have $\eta([0, t] \times X) = t$.

We denote by $D_X[0, +\infty)$ the space of càdlàg functions (right continuous with left limits) from $[0, +\infty)$ to X .

We introduce the occupation measure associated to the process $(V_t^\varepsilon, S_t^\varepsilon)_{t \geq 0}$: for all $t \geq 0$, $A \in \mathcal{B}(E_1)$,

$$\Gamma^\varepsilon([0, t], A) \stackrel{def.}{=} \int_0^t \mathbb{1}_A(V_u^\varepsilon, S_u^\varepsilon) du.$$

Hence, for each ε , $\Gamma^\varepsilon \in \mathcal{L}_m(E_1)$.

Theorem 3.16. *Assume that for any $\varepsilon \in (0, 1)$, the initial condition $(V_0^\varepsilon, S_0^\varepsilon, W_0^\varepsilon)$ have the same law with compact support $K_0 \subset E$. Then, as ε tends to 0, $(\Gamma^\varepsilon, W_t^\varepsilon)_{t \geq 0}$ converges in law to $(\Gamma, \overline{W}_t)_{t \geq 0}$ in $\mathcal{L}_m(E_1) \times D_{E_2}[0, +\infty)$, where*

1. $(\overline{W}_t)_{t \geq 0}$ is a pure jump process on E_2 with generator \mathcal{C}_{av} defined for all $g_2 \in D(\mathcal{C}_{av})$ by

$$\mathcal{C}_{av}g_2(w) = \int_{E_1} \mathcal{B}_{syn}^2 g_2(v, s, w) \pi_w(dv, ds). \quad (3.16)$$

2. $\Gamma(du, dv, ds) = \pi_{\overline{W}_u}(dv, ds)du$.

The proof of this theorem is based on Theorem 2.1 of [81]. We recall it here for the sake of clarity.

Theorem 3.17. *Let M_1 and M_2 be complete, separable metric spaces, and set $M = M_1 \times M_2$. For each n , let $(X_t^n, Y_t^n)_{t \geq 0}$ be a stochastic process with sample paths in $D_M[0, +\infty)$ adapted to a filtration $\{\mathcal{F}_t^n\}$. Assume that $(Y_t^n)_{t \geq 0}$ satisfies the compact containment condition, that is, for each $\epsilon > 0$ and $T > 0$, there exists a compact $K \subset M_2$ such that*

$$\inf_n \mathbb{P}(Y_t^n \in K, t < T) > 1 - \epsilon,$$

and assume that $(X_t^n)_{t \geq 0, n \in \mathbb{N}}$ is relatively compact (as a collection of M_1 -valued random variables). Suppose there is an operator $\mathcal{A} : D(\mathcal{A}) \subset C_b(M_2) \rightarrow C_b(M_1 \times M_2)$ such that for $f \in D(\mathcal{A})$ there is a process $\epsilon_t^{n,f}$ for which,

$$f(Y_t^n) - \int_0^t \mathcal{A}f(X_u^n, Y_u^n) du + \epsilon_t^{n,f}$$

is a $\{\mathcal{F}_t^n\}$ -martingale. Let $D(\mathcal{A})$ be dense in $C_b(M_2)$ in the topology of uniform convergence on compact sets. Suppose that for each $f \in D(\mathcal{A})$ and each $T > 0$, there exists $p > 1$ such that

$$\sup_n \mathbb{E} \left[\int_0^T |\mathcal{A}f(X_t^n, Y_t^n)|^p dt \right] < +\infty \quad \text{and} \quad \lim_{n \rightarrow \infty} \mathbb{E} \left[\sup_{t \leq T} |\epsilon_t^{n,f}| \right] = 0.$$

Let Γ^n be the $\mathcal{L}_m(M_1)$ -valued random variable given by

$$\Gamma^n([0, t], B) = \int_0^t \mathbb{1}_B(X_u^n) du.$$

Then $(\Gamma^n, Y_t^n)_{t \geq 0}$ is relatively compact in $\mathcal{L}_m(M_1) \times D_{M_2}[0, +\infty)$, and for any limit point $(\Gamma, Y_t)_{t \geq 0}$ there exists a filtration $\{\mathcal{G}_t\}$ such that

$$f(Y_t) - \int_0^t \int_{M_1} \mathcal{A}f(x, Y_u) \Gamma(du, dx)$$

is a $\{\mathcal{G}_t\}$ -martingale for each $f \in D(\mathcal{A})$.

In particular, we apply this Theorem in the special case of Example 2.3 given in the same article.

Proof of Theorem 3.16. Let $(\varepsilon_n)_{n \geq 0}$ be a sequence in $(0, 1)$ converging to 0 with n . In STEP 1, we show that the sequence of processes $((\Gamma^{\varepsilon_n}, W_t^{\varepsilon_n})_{t \geq 0})_{n \in \mathbb{N}}$ is relatively compact in $\mathcal{L}_m(E_1) \times$

$D_{E_2}[0, +\infty)$. Moreover, as n tends to infinity, for any limit point $(\Gamma, \bar{W}_t)_{t \geq 0}$, and $g_2 \in D(\mathcal{B}_{syn}^2)$, the process

$$g_2(\bar{W}_t) - \int_0^t \int_{E_1} \mathcal{B}_{syn}^2 g_2(v, s, \bar{W}_u) \Gamma(du, dv, ds) \quad (3.17)$$

is a $\{\mathcal{G}_t\}$ -martingale where $\mathcal{G}_t = \sigma\{\Gamma([0, u], H), \bar{W}_u : u \leq t, H \in \mathcal{B}(E_1)\}$. In STEP 2, similarly as Example 2.3 in [81], we characterize this limit by showing that

$$\Gamma(du, dv, ds) = \pi_{\bar{W}_u}(dv, ds)du. \quad (3.18)$$

Hence, by uniqueness of π_w for all $w \in E_2$ (see Proposition 3.8), we deduce that the unique solution to the martingale problem (3.17) with initial condition $\mathcal{L}(W_0^{\varepsilon_0})$ is the process $(\bar{W}_t)_{t \geq 0}$ with initial law $\mathcal{L}(W_0^{\varepsilon_0})$ and generator \mathcal{C}_{av} given by (3.16). As this limit does not depend on the choice of the sequence $(\varepsilon_n)_{n \in \mathbb{N}}$, we conclude on the convergence of the process $((\Gamma^\varepsilon, W_t^\varepsilon)_{t \geq 0})_{0 < \varepsilon < 1}$ to $(\Gamma, \bar{W}_t)_{t \geq 0}$ when ε tends to 0.

STEP 1:

This step relies on the three main properties that we enumerate here and then show in the following.

1. The process $(W_t^{\varepsilon_n})_{t \geq 0}$ satisfies the compact containment condition that is for each $\eta > 0$ and $T > 0$, there exists a compact set $K_T^\eta \subset E_2$ such that:

$$\inf_n \mathbb{P}(\forall t \in [0, T], W_t^{\varepsilon_n} \in K_T^\eta) \geq 1 - \eta.$$

2. For each $t \geq 0$, the collection of E_1 -valued random variables $(V_t^{\varepsilon_n}, S_t^{\varepsilon_n})_{t \geq 0, n \in \mathbb{N}}$ is relatively compact. In fact, we show that for each $\eta > 0$, there exists a compact set C^η of E_1 such that:

$$\inf_{t \geq 0, n \in \mathbb{N}} \mathbb{P}((V_t^{\varepsilon_n}, S_t^{\varepsilon_n}) \in C^\eta) \geq 1 - \eta.$$

Relative compactness follows by Prohorov's Theorem.

3. For each $g_2 \in D(\mathcal{B}_{syn}^2)$, there exists a process $(\nu_t^{\varepsilon_n, g_2})_{t \geq 0}$ for which

$$g_2(W_t^{\varepsilon_n}) - \int_0^t \mathcal{B}_{syn}^2 g_2(V_u^{\varepsilon_n}, S_u^{\varepsilon_n}, W_u^{\varepsilon_n}) du + \nu_t^{\varepsilon_n, g_2}$$

is a $\{\mathcal{F}_t^{\varepsilon_n}\}$ -martingale where $(\mathcal{F}_t^{\varepsilon_n})_{t \geq 0}$ is the natural filtration of $(V_t^{\varepsilon_n}, S_t^{\varepsilon_n}, W_t^{\varepsilon_n})_{t \geq 0}$.

Moreover, for each $g_2 \in D(\mathcal{B}_{syn}^2)$ and each $T > 0$,

$$\sup_{n \in \mathbb{N}} \mathbb{E} \left[\int_0^T |\mathcal{B}_{syn}^2 g_2(V_u^{\varepsilon_n}, S_u^{\varepsilon_n}, W_u^{\varepsilon_n})|^2 du \right] < \infty \quad (3.19)$$

and

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\sup_{0 \leq t \leq T} |\nu_t^{\varepsilon_n, g_2}| \right] = 0. \quad (3.20)$$

We give some useful notations before dealing with the technical details of the proof. For all $w \in E_2$, we denote by $\|w\|_\infty = \max_{i,j} \{|w^{ij}|\}$. We denote by

$$w_M = \sup_{(v,s,w) \in K_0} \|w\|_\infty \quad \text{and} \quad s_M = \sup_{(v,s,w) \in K_0, 1 \leq i \leq N} s^i.$$

We show the first point. To do so, we use the fact that the weights jump a finite number of times within a finite time interval, and this for all ε_n . For all $m \in \mathbb{N}$, we denote by R^m the subset of E_2 such that $R^m \stackrel{\text{def.}}{=} \{w \in E_2 : \|w\|_\infty \leq m + w_M\}$. Therefore, we want to show that for each $\eta > 0$ and $T > 0$, there exists m large enough to have for all $n \in \mathbb{N}$:

$$\mathbb{P}(\forall t \in [0, T], W_t^{\varepsilon_n} \in R^m) \geq 1 - \eta. \quad (3.21)$$

We denote by $(N_{t,slow}^{\varepsilon_n})_{t \geq 0}$ the number of jumps on $[0, t]$ of the process $(W_u^{\varepsilon_n})_{0 \leq u \leq t}$. From the definition of $\frac{1}{\varepsilon_n} \mathcal{C}^{\varepsilon_n}$ and Remark 3.12, we deduce that the jump rate of the process $(W_t^{\varepsilon_n})_{t \geq 0}$ is bounded by $r_M \stackrel{\text{def.}}{=} N \alpha_M \varphi^M > 0$. Then, we denote by $(\bar{N}_{t,slow})_{t \geq 0}$ the counting process with constant jump rate equals to r_M and with the same randomness as $(N_{t,slow}^{\varepsilon_n})_{t \geq 0}$. Thus, we have for all t, n ,

$$\forall \omega \in \Omega, \quad N_{t,slow}^{\varepsilon_n}(\omega) \leq \bar{N}_{t,slow}(\omega).$$

Hence,

$$\mathbb{P}(\forall t \in [0, T], W_t^{\varepsilon_n} \in R^m) \geq \mathbb{P}(N_{T,slow}^{\varepsilon_n} < m) \geq \mathbb{P}(\bar{N}_{T,slow} < m) \xrightarrow{m \rightarrow +\infty} 1.$$

Taking m_η such that $\mathbb{P}(\bar{N}_{T,slow} < m_\eta) \geq 1 - \eta$ and $K_T^\eta \stackrel{\text{def.}}{=} R^{m_\eta}$ ends the proof of the first point.

We show the second point. To do so, we use the reset of the time from the last spike and the fact that the probability of having no spike in $[0, T]$ decreases with T . For all $T > 0$, $n \in \mathbb{N}$, for each $t \leq T$,

$$\mathbb{P}((V_t^{\varepsilon_n}, S_t^{\varepsilon_n}) \in \{0, 1\}^N \times [0, s_M + T]^N) = 1$$

and for each $t > T$,

$$\mathbb{P}((V_t^{\varepsilon_n}, S_t^{\varepsilon_n}) \in \{0, 1\}^N \times [0, s_M + T]^N) = 1 - \mathbb{P}(\exists i, S_t^{i,\varepsilon_n} > s_M + T).$$

We denote by $(N_{t,fast}^{i,\varepsilon_n})_{t \geq 0}$ the number of jumps of the process $(V_u^{i,\varepsilon_n}, S_u^{i,\varepsilon_n})_{0 \leq u \leq t}$. We can compare these counting processes to the family $((\underline{N}_{t,fast}^i)_{t \geq 0})_{1 \leq i \leq N}$ which are independent counting processes with jump rate α_m and the same randomness as $(N_{t,fast}^{i,\varepsilon_n})_{t \geq 0}$. Thus, by assumption (3.1), for all i, t, n ,

$$\forall \omega \in \Omega, \quad N_{t,fast}^{i,\varepsilon_n}(\omega) \geq \underline{N}_{t,fast}^i(\omega).$$

Therefore, as S_t^{i,ε_n} is reset to 0 when the neuron i spikes (jump of V_t^{i,ε_n} from 0 to 1), the first reset of S_t^{i,ε_n} occurs only after $V_0^{i,\varepsilon_n} + 1$ jumps of V_t^{i,ε_n} . Hence,

$$\begin{aligned} & \mathbb{P}(\exists i, S_t^{i,\varepsilon_n} > s_M + T) \\ &= \mathbb{P}\left(\exists i, \mathbb{1}_{\{t > T + s_M\}} (N_{t,fast}^{i,\varepsilon_n} - N_{t-s_M-T,fast}^{i,\varepsilon_n}) \right. \\ & \quad \left. + \mathbb{1}_{\{T + s_M - S_0^{i,\varepsilon_n} < t \leq T + s_M\}} N_{t,fast}^{i,\varepsilon_n} \leq V_{\max(0, t-s_M-T)}^{i,\varepsilon_n}\right) \\ &\leq \mathbb{P}\left(\exists i, \mathbb{1}_{\{t > T + s_M\}} (N_{t,fast}^{i,\varepsilon_n} - N_{t-s_M-T,fast}^{i,\varepsilon_n}) + \mathbb{1}_{\{T + s_M - S_0^{i,\varepsilon_n} < t \leq T + s_M\}} N_{t,fast}^{i,\varepsilon_n} \leq 1\right) \\ &\leq \mathbb{P}(\exists i, N_{T,fast}^{i,\varepsilon_n} \leq 1) \leq \mathbb{P}(\exists i, \underline{N}_{T,fast}^i \leq 1) \leq N \mathbb{P}(\underline{N}_{T,fast}^1 \leq 1) \xrightarrow{T \rightarrow +\infty} 0. \end{aligned}$$

We denote by $T^\eta = \inf\{t \geq 0 : \mathbb{P}(\underline{N}_{t,fast}^1 \leq 1) \leq \frac{\eta}{N}\}$. Hence, defining $C^\eta = \{0, 1\}^N \times [0, s_M + T^\eta]$ ends the proof of the second point.

We show the third point. In the following, for all $g_2 \in D(\mathcal{B}_{syn}^2)$, we define $g \in D(\mathcal{C}^{\varepsilon_n})$ such that for all $(v, s, w) \in E$, $g(v, s, w) = g_2(w)$ and we denote by

$$\nu_t^{\varepsilon_n, g_2} = \int_0^t \frac{1}{\varepsilon_n} \mathcal{B}_r^{\varepsilon_n} g(V_u^{\varepsilon_n}, S_u^{\varepsilon_n}, W_u^{\varepsilon_n}) du.$$

Thus, from the definition of the process $(V_t^{\varepsilon_n}, S_t^{\varepsilon_n}, W_t^{\varepsilon_n})_{t \geq 0}$, we have that

$$\begin{aligned} g(V_t^{\varepsilon_n}, S_t^{\varepsilon_n}, W_t^{\varepsilon_n}) - \int_0^t \frac{1}{\varepsilon_n} \mathcal{C}^{\varepsilon_n} g(V_u^{\varepsilon_n}, S_u^{\varepsilon_n}, W_u^{\varepsilon_n}) du \\ = g_2(W_t^{\varepsilon_n}) - \int_0^t \mathcal{B}_{syn}^2 g_2(V_u^{\varepsilon_n}, S_u^{\varepsilon_n}, W_u^{\varepsilon_n}) du + \nu_t^{\varepsilon_n, g_2} \end{aligned}$$

is a $\{\mathcal{F}_t^{\varepsilon_n}\}$ -martingale, see Remark 3.14.

Using the bound given in Remark 3.12 and the boundedness of g_2 , we obtain that for all $(v, s, w) \in E$,

$$\begin{aligned} |\mathcal{B}_{syn}^2 g_2(v, s, w)| &= \left| \sum_i \delta_0(v^i) \alpha_i(v, w) \sum_{\tilde{w} \neq w} (g_2(\tilde{w}) - g_2(w)) \varphi_i(s, \tilde{w}, w) \right| \\ &\leq 2 \|g_2\|_{\infty} \left| \sum_i \delta_0(v^i) \alpha_i(v, w) \sum_{\tilde{w} \neq w} \varphi_i(s, \tilde{w}, w) \right| \\ &\leq 2 \|g_2\|_{\infty} \varphi^M \left| \sum_i \delta_0(v^i) \alpha_i(v, w) \right| \leq 2 \|g_2\|_{\infty} \varphi^M N \alpha_M. \end{aligned}$$

Hence, we deduce that (3.19) holds true.

We obtain (3.20) similarly. Indeed, using the fact that only a finite number of states $\tilde{w} \in E_2$ are reachable from any $w \in E_2$, we have for all $(v, s, w) \in E$,

$$\begin{aligned} \left| \frac{1}{\varepsilon_n} \mathcal{B}_r^{\varepsilon_n} g(v, s, w) \right| &= \left| \sum_i \delta_0(v^i) \alpha_i(v, w) \sum_{\tilde{w} \neq w} (g(\tilde{w}) - g(w)) \frac{1}{\varepsilon_n} K_i^{\varepsilon_n}(s, \tilde{w}, w) \right| \\ &\leq 2 \|g\|_{\infty} \alpha_M \sum_i \sum_{\tilde{w} \neq w} \frac{1}{\varepsilon_n} |K_i^{\varepsilon_n}(s, \tilde{w}, w)| = \mathcal{O}(1). \end{aligned}$$

Thereby, $\nu_t^{\varepsilon_n, g_2} = \mathcal{O}(1)$.

From these three points, we can apply Theorem 3.17 to the sequence $((\Gamma^{\varepsilon_n}, W_t^{\varepsilon_n})_{t \geq 0})_{n \geq 0}$ and we obtain (3.17) and (3.18). It ends the STEP 1.

STEP 2:

Similarly as before and as in Remark 3.14, for all $f_1 \in C_b^1(E_1) \subset D(\mathcal{B}_{net}^1)$,

$$\varepsilon_n f_1(V_t^{\varepsilon_n}, S_t^{\varepsilon_n}) - \int_0^t \mathcal{B}_{net}^1 f_1(V_u^{\varepsilon_n}, S_u^{\varepsilon_n}, W_u^{\varepsilon_n}) du \quad (3.22)$$

is a $\{\mathcal{F}_t^{\varepsilon_n}\}$ -martingale. When n tends to infinity, the first term tends to 0. For the second term, we apply Lemma 1.5 a) of [81]. It gives that along an appropriate subsequence, that we still denote by $(\varepsilon_n)_n$ for simplicity, we have

$$\begin{aligned} \int_0^t \mathcal{B}_{net}^1 f_1(V_u^{\varepsilon_n}, S_u^{\varepsilon_n}, W_u^{\varepsilon_n}) du &= \int_0^t \int_{E_1} \mathcal{B}_{net}^1 f_1(v, s, W_u^{\varepsilon_n}) \Gamma^{\varepsilon_n}(du, dv, ds) \\ &\xrightarrow{n \rightarrow \infty} \int_0^t \int_{E_1} \mathcal{B}_{net}^1 f_1(v, s, \overline{W}_u) \Gamma(du, dv, ds). \end{aligned}$$

Thus, by equation (3.22), the limit term of the last equation is a martingale.

Since it is continuous and of bounded variation, it must be constant, see Proposition 1.2 [113]. Hence, as the limit is null when t tends to 0, we get that for all t almost surely

$$\int_0^t \int_{E_1} \mathcal{B}_{\overline{W}_u}^1 f_1(v, s, \overline{W}_u) \Gamma(du, dv, ds) = \int_0^t \int_{E_1} \mathcal{B}_{\overline{W}_u} f_1(v, s) \Gamma(du, dv, ds) = 0.$$

Since for all $n \in \mathbb{N}$ we have that $\Gamma^{\varepsilon_n} \in \mathcal{L}_m(E_1)$, then by Lemma 1.4 in [81] we deduce that there exists a process γ_u on the set of probability measure on E_1 such that for every function $f_1 \in C_b^1(E_1)$,

$$\int_0^t \int_{E_1} \mathcal{B}_{\overline{W}_u} f_1(v, s) \Gamma(du, dv, ds) = \int_0^t \int_{E_1} \mathcal{B}_{\overline{W}_u} f_1(v, s) \gamma_u(dv, ds) du = 0.$$

Thus, we have for dt -almost every $u \in \mathbb{R}^+$,

$$\int_{E_1} \mathcal{B}_{\overline{W}_u} f_1(v, s) \gamma_u(dv, ds) = 0.$$

From Proposition 3.8, for each $u \geq 0$, the unique probability measure satisfying the previous equation is $\pi_{\overline{W}_u}$ and hence, for all $u \geq 0$,

$$\gamma_u = \pi_{\overline{W}_u}.$$

We deduce that, as n tends to infinity, for any limit point $(\Gamma, \overline{W}_t)_{t \geq 0}$, we have $\Gamma(du, dv, ds) = \pi_{\overline{W}_u}(dv, ds) du$ and

$$g_2(\overline{W}_t) - \int_0^t \int_{E_1} \mathcal{B}_{\overline{W}_u}^2 g_2(v, s, \overline{W}_u) \pi_{\overline{W}_u}(dv, ds) du \quad (3.23)$$

is a $\{\mathcal{G}_t\}$ -martingale for any $g_2 \in D(\mathcal{B}_{syn}^2)$ where $\mathcal{G}_t = \sigma\{\pi_{\overline{W}_u}(H), \overline{W}_u : u \leq t, H \in \mathcal{B}(E_1)\}$. Therefore, $(\overline{W}_t)_{t \geq 0}$ is uniquely defined as the process with generator $(\mathcal{C}_{av}, D(\mathcal{C}_{av}))$ such that for all $g_2 \in D(\mathcal{C}_{av})$,

$$\mathcal{C}_{av} g_2(w) = \int_{E_1} \mathcal{B}_{syn}^2 g_2(v, s, w) \pi_w(dv, ds)$$

and initial condition $\overline{W}_0 \stackrel{\mathcal{L}}{=} W_0^{\varepsilon_0}$. The uniqueness of the limit point $(\Gamma, \overline{W}_t)_{t \geq 0}$ enables us to conclude on the convergence of $((\Gamma^{\varepsilon_n}, \overline{W}_t^{\varepsilon_n})_{t \geq 0})_{n \in \mathbb{N}}$ to $(\Gamma, \overline{W}_t)_{t \geq 0}$ when n tends to ∞ . □

Remark 3.18. *All the previous results hold if we take into account inhibitory neurons. Indeed, we never use the positivity of the weights in this subsection. However, for inhibitory neurons to be realistic, their postsynaptic weights should follow the opposite plasticity rule: p^+ and p^- should be multiplied by the sign of the weight. With such a model of inhibitory neurons, similar results as the following one (holding on the long time behaviour of the weight limit dynamics) can be obtained, see Remark 3.41.*

3.4 Long time behaviour: conditions for recurrence and transience

Plasticity models have evolved interacting with neurobiologists' discoveries. For instance, models based on STDP confirmed the need of homeostasis in order to regulate the evolution of the weights: preventing from their divergence or extinction, see [137] for instance. Indeed, Hebbian learning suffers from a positive feedback instability and leads to all neurons wiring together, see for example [136]. Synaptic scaling and metaplasticity are the main homeostatic mechanisms used in models through different ways, see [134]. In our model, we do not have such mechanisms, like hard or soft bounds. Nevertheless, we can show that the weights stabilize under appropriate conditions. We propose to compare the process $(\bar{W}_t)_{t \geq 0}$ to slightly different ones in order to bound its jump rate. Then, we show how to use the Laplace transform of π_w to study the long time behaviour of $(\bar{W}_t)_{t \geq 0}$.

3.4.1 General condition on the rates and parameters

In our case, the process $(\bar{W}_t)_{t \geq 0}$ takes values in the countable state E_2 , is a Markov process and is non-homogeneous in space and homogeneous in time. As underlined in [96], Lyapunov techniques are well adapted to analyse the long time behaviour of such processes. In particular, we use here the Foster's Theorem, see Theorem A.10 in the appendix [96, Thm 2.6.4].

From Theorem 3.16, we are interested in the dynamics of the process $(\bar{W}_t)_{t \geq 0}$ with generator \mathcal{C}_{av} such that for all $f \in D(\mathcal{C}_{av})$,

$$\begin{aligned} \mathcal{C}_{av}f(w) &= \int_{E_1} \sum_k \delta_0(v^k) \alpha_k(v, w) \left(\sum_{\tilde{w} \in E_2} (f(\tilde{w}) - f(w)) \varphi_k(s, \tilde{w}, w) \right) \pi_w(dv, ds) \\ &= \sum_{\tilde{w} \in E_2} (f(\tilde{w}) - f(w)) r(\tilde{w}, w), \end{aligned}$$

where r is the Q -matrix (see [62, Rk 5.3]) of $(\bar{W}_t)_{t \geq 0}$, $r(w, w) \stackrel{def.}{=} -\sum_{\tilde{w} \neq w} r(\tilde{w}, w)$ and

$$\forall \tilde{w} \neq w, \quad r(\tilde{w}, w) \stackrel{def.}{=} \int_{E_1} \sum_k \delta_0(v^k) \alpha_k(v, w) \varphi_k(s, \tilde{w}, w) \pi_w(dv, ds).$$

Notation 3.19. *In the following, we use the notations:*

$$\forall i \neq j, \forall w \in E_2, \quad r_{ij}^+(w) \stackrel{def.}{=} \sum_{\substack{\tilde{w} \in A_w^i \\ \tilde{w}^{ij} = w^{ij} + 1}} r(\tilde{w}, w) \quad \text{and} \quad r_{ij}^-(w) \stackrel{def.}{=} \sum_{\substack{\tilde{w} \in A_w^j \\ \tilde{w}^{ij} = w^{ij} - 1}} r(\tilde{w}, w).$$

Moreover, we denote by R^+ and R^- the square matrices of size N , null on the diagonal and such that $R^+ = (r_{ij}^+)_{1 \leq i \neq j \leq N}$ and $R^- = (r_{ij}^-)_{1 \leq i \neq j \leq N}$. Finally, we denote by \cdot the scalar product between two square matrices A and B :

$$A \cdot B = \sum_{1 \leq i, j \leq N} A^{ij} B^{ij}.$$

Remark 3.20. From Proposition 3.11,

$$r_{ij}^+(w) = \int_{E_1} \delta_0(v^i) \alpha_i(v, w) p^+(s^j, w^{ij}) \pi_w(dv, ds) = \mathbb{E} \left[\mathbb{1}_{\{V_w^i=0\}} \alpha_i(V_w, w) p^+(S_w^j, w^{ij}) \right], \quad (3.24)$$

$$r_{ij}^-(w) = \mathbb{E} \left[\mathbb{1}_{\{V_w^j=0\}} \alpha_j(V_w, w) p^-(S_w^i, w^{ij}) \right] \quad (3.25)$$

where $(V_w, S_w) \stackrel{\mathcal{L}}{=} \pi_w$.

Moreover, under Assumption 3.9, the jumps of the process $(\bar{W}_t)_{t \geq 0}$ are restricted to $\pm E_{ij}$ at rate $r_{ij}^\pm(w)$ for all $i \neq j$.

A natural assumption in order to ensure the irreducibility and aperiodicity of the embedded chain of the process $(\bar{W}_t)_{t \geq 0}$ is

Assumption 3.21. Assume that p^+ (resp. p^-) is strictly positive on $\mathbb{R}^+ \times \mathbb{N}^*$ (resp. $\mathbb{R}^+ \times \mathbb{N}^* \setminus \{1\}$).

In particular, this assumption gives

Lemma 3.22. Assume that Assumption 3.21 holds. Then, there exist $0 < r_m \leq r_M < \infty$ such that

$$\forall w \in E_2, \quad r_m \leq |r(w, w)| \leq r_M.$$

Proof. First, for all $w \in E_2$, $|r(w, w)| \leq N \alpha_M \varphi^M \stackrel{\text{def.}}{=} r_M < \infty$.

Second,

$$\begin{aligned} |r(w, w)| &= \int_{E_1} \sum_k \delta_0(v^k) \alpha_k(v, w) \left(\sum_{\tilde{w} \neq w} \varphi_k(s, \tilde{w}, w) \right) \pi_w(dv, ds) \\ &\geq \alpha_m \max_l \left\{ \sum_k \int_{E_1} \delta_0(v^k) (p^+(s^l, w^{kl}) + p^-(s^l, w^{lk})) \pi_w(dv, ds) \right\} \stackrel{\text{def.}}{=} r_m > 0 \end{aligned}$$

where strict positivity is obtained from Assumption 3.21. \square

Definition 3.23. Let $(X_t)_{t \geq 0}$ be a stochastic process with jumps. The embedded chain of $(X_t)_{t \geq 0}$ is the chain $(X_{\tau_n})_{n \in \mathbb{N}}$ where $(\tau_n)_{n \in \mathbb{N}^*}$ is the sequence of the jumping times of $(X_t)_{t \geq 0}$ and $\tau_0 = 0$.

Proposition 3.24. Assume that Assumption 3.21 holds. Then, we have the two following results.

1. Using Notation 3.19, we assume that

$$\lim_{\substack{w \in E_2, \\ \|w\| \rightarrow +\infty}} (R^+(w) - R^-(w)) \cdot w = -\infty. \quad (3.26)$$

Then, the embedded Markov chain associated to the process $(\bar{W}_t)_{t \geq 0}$ is positive recurrent. Moreover, the process $(\bar{W}_t)_{t \geq 0}$ converges to a unique invariant measure.

2. Assume that

$$\lim_{\substack{w \in E_2, \\ \|w\| \rightarrow +\infty}} (R^+(w) - R^-(w)) \cdot w = +\infty.$$

Then, the embedded Markov chain associated to the process $(\bar{W}_t)_{t \geq 0}$ is transient. Moreover, the process $(\bar{W}_t)_{t \geq 0}$ is also transient.

Proof. First, from the assumption of strict positivity of p^+ and p^- we have that for all $i \neq j$ and $w \in E_2$, $r(w + E_{ij}, w) > 0$ and if $w^{ij} > 1$, $r(w - E_{ij}, w) > 0$. Thus, the embedded chain of $(\bar{W}_t)_{t \geq 0}$ is irreducible and aperiodic.

Then, we use the Foster's Theorem, see Theorem A.10 in the appendix. We apply this theorem to the embedded chain of $(\bar{W}_t)_{t \geq 0}$: $(\bar{W}_{\tau_n})_{n \geq 0}$ where $\tau_0 = 0$ and $(\tau_n)_{n \in \mathbb{N}^*}$ is the sequence of the jumping times of $(\bar{W}_t)_{t \geq 0}$. By Lemma 3.22, we have for all $w \in E_2$,

$$0 < r_m < |r(w, w)| < r_M < \infty.$$

Thus, the transition kernel Q of the embedded chain is given by $Q(\tilde{w}, w) = \frac{r(\tilde{w}, w)}{|r(w, w)|}$ for all $\tilde{w} \neq w \in E_2$ and $Q(w, w) = 0$. Then, for all function f on E_2 , for all $w \in E_2$,

$$\mathbb{E} [f(\bar{W}_{\tau_{n+1}}) - f(\bar{W}_{\tau_n}) \mid \bar{W}_{\tau_n} = w] = Qf(w) - f(w) = \frac{\mathcal{C}_{av}f(w)}{|r(w, w)|}.$$

Let $f(w) = \sum_{i,j} (w^{ij})^2$ for $w \in E_2$, then,

$$\begin{aligned} \mathcal{C}_{av}f(w) &= \sum_{\tilde{w}} \sum_{i,j} ((\tilde{w}^{ij})^2 - (w^{ij})^2) r(\tilde{w}, w) = \sum_{i,j: i \neq j} \sum_{\tilde{w}} ((\tilde{w}^{ij})^2 - (w^{ij})^2) r(\tilde{w}, w) \\ &= \sum_{i,j: i \neq j} \left(\sum_{\tilde{w} \in A_w^i, \tilde{w}^{ij} = w^{ij} + 1} (2w^{ij} + 1) r(\tilde{w}, w) + \sum_{\tilde{w} \in A_w^j, \tilde{w}^{ij} = w^{ij} - 1} (-2w^{ij} + 1) r(\tilde{w}, w) \right) \\ &\leq N(N-1)r_M + \sum_{i,j: i \neq j} 2w^{ij} (r_{ij}^+(w) - r_{ij}^-(w)). \end{aligned}$$

For $R \in \mathbb{R}^+$, we denote $B(R) = \{w \in E_2 : \|w\|_\infty \leq R\}$. From (3.26), the last term converges to $-\infty$ when $\|w\|_\infty$ tends to infinity. Hence, there exists R_0 and $\gamma < 0$ such that for all $w \in E_2 \setminus B(R_0)$, $\mathcal{C}_{av}f(w) \leq \gamma$ and thus

$$Qf(w) - f(w) = \frac{\mathcal{C}_{av}f(w)}{|r(w, w)|} \leq \frac{\gamma}{r_M} < 0.$$

Moreover, for all $\bar{W}_{\tau_n} \in B(R_0)$, we know that $\bar{W}_{\tau_{n+1}} \in B(R_0 + 1)$ so that $f(\bar{W}_{\tau_{n+1}}) \leq N^2(R_0 + 1)^2$. Thus,

$$\begin{aligned} \mathbb{E} [f(\bar{W}_{\tau_{n+1}}) \mid \bar{W}_{\tau_n} = w] &= \sum_{\tilde{w} \in B(R_0+1)} f(\tilde{w}) Q(\tilde{w}, w) \\ &\leq N^2(R_0 + 1)^2 \sum_{\tilde{w} \in B(R_0+1)} Q(\tilde{w}, w) = N^2(R_0 + 1)^2 < \infty. \end{aligned}$$

Hence, by Foster's Theorem A.10, the Markov chain $(\bar{W}_{\tau_n})_{n \in \mathbb{N}^*}$ is positive recurrent. From aperiodicity and irreducibility of the chain, we have by Theorem 2.1.6 of [96] that this chain converges to its unique invariant measure.

Transience is obtained in the same way using Theorem 2.5.15 of [96].

Finally, passing from the embedded chain to the process is obvious since the process is non explosive, see Theorem 3.5.1 in [105] for instance. \square

Corollary 3.25. *Assume that Assumption 3.21 holds, then we have the two following results.*

1. Assume that there exists $\bar{\gamma} < 0$ such that for all i, j ,

$$\lim_{w_{lim} \rightarrow +\infty} \left(\sup_{\substack{w \in E_2, \\ w^{ij} = w_{lim}}} r_{ij}^+(w) - r_{ij}^-(w) \right) = \bar{\gamma}.$$

Then, the embedded Markov chain associated to the process $(\bar{W}_t)_{t \geq 0}$ is positive recurrent. Moreover, the process $(\bar{W}_t)_{t \geq 0}$ converges to a unique invariant measure.

2. Assume that there exists $\underline{\gamma} > 0$ such that for all i, j ,

$$\lim_{w_{lim} \rightarrow +\infty} \left(\inf_{\substack{w \in E_2, \\ w^{ij} = w_{lim}}} r_{ij}^+(w) - r_{ij}^-(w) \right) = \underline{\gamma}.$$

Then, the embedded Markov chain associated to the process $(\bar{W}_t)_{t \geq 0}$ is transient. Moreover, the process $(\bar{W}_t)_{t \geq 0}$ is also transient.

This corollary incites us to study of r_{ij}^+ and r_{ij}^- in order to get information on the long time behaviour of the process $(\bar{W}_t)_{t \geq 0}$.

Remark 3.26. From the previous proposition and Remark 3.20, we note that it is not obvious that the condition $p^+(\hat{s}, \hat{w}) < p^-(\hat{s}, \hat{w})$ for all $(\hat{s}, \hat{w}) \in \mathbb{R}^+ \times \mathbb{N}^*$ is sufficient for the recurrence. Indeed, we observe the divergence of some weights under that condition in the simulations, see Figure 3.4.

We now bound the functions r_{ij}^+ and r_{ij}^- . To this aim, we use a classical coupling method and compare the number of jumps of $(V_t, \underline{S}_t)_{t \geq 0}$ to the associated processes where the rate α_i are replaced by the minimal rate α_m or the maximal rate α_M . In what follows, we define, for all j , the processes $(\underline{V}_{t,j}, \underline{S}_{t,j})_{t \geq 0}$ and $(\bar{V}_{t,j}, \bar{S}_{t,j})_{t \geq 0}$ in E_1 . Their N coordinates are independent. The dynamics of these coordinates are driven by similar generators that we denote by \mathcal{B}_α for $\alpha \in \{\alpha_m, \alpha_M\}$ and such that for all $f \in D(\mathcal{B}_\alpha)$, $(\hat{v}, \hat{s}) \in \{0, 1\} \times \mathbb{R}^+$,

$$\mathcal{B}_\alpha f(\hat{v}, \hat{s}) = \partial_{\hat{s}} f(\hat{v}, \hat{s}) + \beta \delta_1(\hat{v}) [f(0, \hat{s}) - f(1, \hat{s})] + \alpha \delta_0(\hat{v}) [f(1, 0) - f(0, \hat{s})]. \quad (3.27)$$

The index j means that the dynamics of their j th coordinate is different from the others. The j th component $(\underline{V}_{t,j}^j, \underline{S}_{t,j}^j)_{t \geq 0}$ is driven by the generator \mathcal{B}_{α_m} whereas the other components are driven by the generator \mathcal{B}_{α_M} . The j th component $(\bar{V}_{t,j}^j, \bar{S}_{t,j}^j)_{t \geq 0}$ is driven by the generator \mathcal{B}_{α_M} whereas the other components are driven by the generator \mathcal{B}_{α_m} .

We compute the invariant measure of these processes.

Proposition 3.27. The unique invariant probability measure μ_α associated to the generator \mathcal{B}_α is

$$\mu_\alpha(d\hat{v}, d\hat{s}) = \frac{\beta}{\alpha + \beta} \delta_0(\hat{v}) \frac{\alpha \beta}{\alpha - \beta} (e^{-\beta \hat{s}} - e^{-\alpha \hat{s}}) d\hat{s} + \frac{\alpha}{\alpha + \beta} \delta_1(\hat{v}) \beta e^{-\beta \hat{s}} d\hat{s}. \quad (3.28)$$

Moreover, for all j , the processes $(\underline{V}_{t,j}, \underline{S}_{t,j})_{t \geq 0}$ and $(\bar{V}_{t,j}, \bar{S}_{t,j})_{t \geq 0}$ converge respectively to their unique invariant measure that we denote by $\underline{\pi}_j$ and $\bar{\pi}_j$:

$$\bar{\pi}_j = \mu_{\alpha_m}^{\otimes(j-1)} \otimes \mu_{\alpha_M} \otimes \mu_{\alpha_m}^{\otimes(N-j)} \quad \text{and} \quad \underline{\pi}_j = \mu_{\alpha_M}^{\otimes(j-1)} \otimes \mu_{\alpha_m} \otimes \mu_{\alpha_M}^{\otimes(N-j)}. \quad (3.29)$$

Proof. First, according to assumption (3.1), for $\alpha \in \{\alpha_m, \alpha_M\}$ we have $\alpha \neq \beta$. From Proposition 3.8, a process with generator \mathcal{B}_α converges to its unique invariant measure μ_α . Moreover, as the dynamics of discrete part of the process does not depend on the continuous one, we easily get that $\mu_\alpha(\{0\}, \mathbb{R}^+) = \frac{\beta}{\alpha+\beta}$ and $\mu_\alpha(\{1\}, \mathbb{R}^+) = \frac{\alpha}{\alpha+\beta}$. Thus, the measure μ_α can be written as:

$$\mu_\alpha(\hat{v}, d\hat{s}) = \frac{\beta}{\alpha+\beta} \delta_0(\hat{v}) \mu_\alpha^0(d\hat{s}) + \frac{\alpha}{\alpha+\beta} \delta_1(\hat{v}) \mu_\alpha^1(d\hat{s})$$

with $\int_{\mathbb{R}^+} \mu_\alpha^0(d\hat{s}) = \int_{\mathbb{R}^+} \mu_\alpha^1(d\hat{s}) = 1$. Moreover, it is an invariant measure if and only if

$$\forall f \in D(\mathcal{B}_\alpha), \quad \langle \mathcal{B}_\alpha f, \mu_\alpha \rangle = 0.$$

Thanks to functions f well-chosen, we get a closed system of equations on the Laplace transforms of μ_α^0 and μ_α^1 . For $y \in \{0, 1\}$, we denote by $e_\lambda^y(\hat{v}, \hat{s}) = e^{-\lambda \hat{s}} \delta_{\hat{v}}(y)$. We thus obtain:

$$\langle \mathcal{B}_\alpha e_\lambda^y, \mu_\alpha \rangle = \frac{\beta}{\alpha+\beta} \int_{\mathbb{R}^+} \mathcal{B}_\alpha e_\lambda^y(0, \hat{s}) \mu_\alpha^0(d\hat{s}) + \frac{\alpha}{\alpha+\beta} \int_{\mathbb{R}^+} \mathcal{B}_\alpha e_\lambda^y(1, \hat{s}) \mu_\alpha^1(d\hat{s}) = 0.$$

Using the definition of \mathcal{B}_α given by (3.27), we get

$$\begin{cases} \frac{\beta}{\alpha+\beta} \int_{\mathbb{R}^+} (\alpha + \lambda) e^{-\lambda \hat{s}} \mu_\alpha^0(d\hat{s}) = \frac{\alpha}{\alpha+\beta} \int_{\mathbb{R}^+} \beta e^{-\lambda \hat{s}} \mu_\alpha^1(d\hat{s}) \\ \beta \int_{\mathbb{R}^+} \alpha \mu_\alpha^0(d\hat{s}) = \alpha \int_{\mathbb{R}^+} (\beta + \lambda) e^{-\lambda \hat{s}} \mu_\alpha^1(d\hat{s}). \end{cases}$$

But $\int_{\mathbb{R}^+} \mu_\alpha^0(d\hat{s}) = \int_{\mathbb{R}^+} \mu_\alpha^1(d\hat{s}) = 1$. Therefore,

$$\begin{cases} \int_{\mathbb{R}^+} (\alpha + \lambda) e^{-\lambda \hat{s}} \mu_\alpha^0(d\hat{s}) = \alpha \int_{\mathbb{R}^+} e^{-\lambda \hat{s}} \mu_\alpha^1(d\hat{s}) \\ \int_{\mathbb{R}^+} e^{-\lambda \hat{s}} \mu_\alpha^1(d\hat{s}) = \frac{\beta}{(\beta+\lambda)} \Leftrightarrow \mu_\alpha^1(\hat{s}) = \beta e^{-\beta \hat{s}}. \end{cases}$$

Hence,

$$\begin{cases} \int_{\mathbb{R}^+} e^{-\lambda \hat{s}} \mu_\alpha^0(d\hat{s}) = \frac{\alpha\beta}{(\alpha+\lambda)(\beta+\lambda)} \Leftrightarrow \mu_\alpha^0(s) = \frac{\alpha\beta}{\alpha-\beta} (e^{-\beta s} - e^{-\alpha s}) \\ \mu_\alpha^1(\hat{s}) = \beta e^{-\beta \hat{s}}. \end{cases}$$

We thus obtain a probability measure candidate to be invariant for the generator \mathcal{B}_α . We finally check this last point showing that for all $f \in D(\mathcal{B}_\alpha)$ we have $\langle \mathcal{B}_\alpha f, \mu_\alpha \rangle = 0$. To do so, we compute the two right hand side terms of

$$\langle \mathcal{B}_\alpha f, \mu_\alpha \rangle = \frac{\beta}{\alpha+\beta} \int_{\mathbb{R}^+} \mathcal{B}_\alpha f(0, \hat{s}) \mu_\alpha^0(d\hat{s}) + \frac{\alpha}{\alpha+\beta} \int_{\mathbb{R}^+} \mathcal{B}_\alpha f(1, \hat{s}) \mu_\alpha^1(d\hat{s}). \quad (3.30)$$

By integration by parts we obtain that

$$\begin{aligned} \int_{\mathbb{R}^+} \mathcal{B}_\alpha f(0, \hat{s}) \mu_\alpha^0(d\hat{s}) &= \frac{\alpha\beta}{\alpha-\beta} \int_{\mathbb{R}^+} (\alpha[f(1, 0) - f(0, \hat{s})] + \partial_s f(0, \hat{s})) (e^{-\beta \hat{s}} - e^{-\alpha \hat{s}}) d\hat{s} \\ &= \alpha f(1, 0) - \alpha\beta \int_{\mathbb{R}^+} f(0, \hat{s}) e^{-\beta \hat{s}} \end{aligned}$$

and

$$\begin{aligned} \int_{\mathbb{R}^+} \mathcal{B}_\alpha f(1, \hat{s}) \mu_\alpha^1(d\hat{s}) &= \beta \int_{\mathbb{R}^+} (\beta[f(0, \hat{s}) - f(1, \hat{s})] + \partial_s f(1, \hat{s})) e^{-\beta \hat{s}} d\hat{s} \\ &= \beta^2 \int_{\mathbb{R}^+} f(0, \hat{s}) e^{-\beta \hat{s}} d\hat{s} - \beta f(1, 0). \end{aligned}$$

We deduce with equation (3.30) that $\mathbb{E}_{\mu_\alpha}[\mathcal{B}_\alpha f] = 0$ which completes the proof of the first point.

The second point of the proposition is a consequence of the first one as the processes $(\underline{V}_{t,j}, \underline{S}_{t,j})_{t \geq 0}$ and $(\overline{V}_{t,j}, \overline{S}_{t,j})_{t \geq 0}$ are vectors with independent components driven by generators \mathcal{B}_α with $\alpha \in \{\alpha_m, \alpha_M\}$. \square

From these notations, we can show the

Proposition 3.28. *Assume that for all $\hat{w} \in \mathbb{N}^*$, $p^+(\cdot, \hat{w})$ and $p^-(\cdot, \hat{w})$ are decreasing and differentiable functions. Thus, for all i, j with $i \neq j$, we have*

$$\begin{aligned} r_{ij}^+(w) - r_{ij}^-(w) &\leq \mathbb{E} \left[\mathbb{1}_{\{\overline{V}_j^i=0\}} \alpha_i(\overline{V}_j, w) p^+(\overline{S}_j^i, w^{ij}) \right] - \mathbb{E} \left[\mathbb{1}_{\{\overline{V}_j^j=0\}} \alpha_j(\overline{V}_j, w) p^-(\overline{S}_j^j, w^{jj}) \right], \\ r_{ij}^+(w) - r_{ij}^-(w) &\geq \mathbb{E} \left[\mathbb{1}_{\{\underline{V}_i^i=0\}} \alpha_i(\underline{V}_i, w) p^+(\underline{S}_i^i, w^{ii}) \right] - \mathbb{E} \left[\mathbb{1}_{\{\underline{V}_i^j=0\}} \alpha_j(\underline{V}_i, w) p^-(\underline{S}_i^j, w^{ij}) \right], \end{aligned} \quad (3.31)$$

where for all k , $(\underline{V}_k, \underline{S}_k) \stackrel{\mathcal{L}}{=} \underline{\pi}_k$ and $(\overline{V}_k, \overline{S}_k) \stackrel{\mathcal{L}}{=} \overline{\pi}_k$.

Proof. The idea is to show that for all $w \in E_2$, the difference $r_{ij}^+(w) - r_{ij}^-(w)$ can be bounded using the processes $(\overline{V}_{t,j}, \overline{S}_{t,j})_{t \geq 0}$ and $(\underline{V}_{t,j}, \underline{S}_{t,j})_{t \geq 0}$. We only present the proof of the upper bound of $r_{ij}^+ - r_{ij}^-$. The lower bound can be obtained exactly in the same way.

We denote by $\overline{P}_{t,j}$ the transition kernel of $(\overline{V}_{t,j}, \overline{S}_{t,j})_{t \geq 0}$ and $\overline{\mathcal{B}}_j$ its generator. Then, we have for all $f \in D(\overline{\mathcal{B}}_j)$,

$$\pi_w \overline{P}_{t,j} f - \pi_w f = \int_0^t \pi_w \overline{\mathcal{B}}_j \overline{P}_{u,j} f du = \int_0^t \pi_w \overline{P}_{u,j} \overline{\mathcal{B}}_j f du.$$

Moreover, as π_w is invariant for the process with generator \mathcal{B}_w , we have $\pi_w \mathcal{B}_w f = 0$ for all $f \in D(\mathcal{B}_w)$. Noting that $D(\mathcal{B}_w) = D(\overline{\mathcal{B}}_j)$ we have for all $f \in D(\mathcal{B}_w)$,

$$\pi_w \overline{P}_{t,j} f - \pi_w f = \int_0^t \pi_w (\overline{\mathcal{B}}_j - \mathcal{B}_w) \overline{P}_{u,j} f du = \int_0^t \pi_w \overline{P}_{u,j} (\overline{\mathcal{B}}_j - \mathcal{B}_w) f du. \quad (3.32)$$

Now, for all j and $i \neq j$, we define the two functions f_{ij}^+ and f_{ij}^- such that for all $(v, s) \in E_1$,

$$f_{ij}^+(v, s) \stackrel{def.}{=} \delta_0(v^i) \alpha_i(v, w) p^+(s^j, w^{ij}) \quad \text{and} \quad f_{ij}^-(v, s) \stackrel{def.}{=} \delta_0(v^j) \alpha_j(v, w) p^-(s^i, w^{ij}),$$

and we show that $(\overline{\mathcal{B}}_j - \mathcal{B}_w)(f_{ij}^+ - f_{ij}^-) \geq 0$. Indeed, we compute

$$\begin{aligned} \mathcal{B}_w f_{ij}^+(v, s) &= \sum_{k=1}^N \partial_{s^k} f_{ij}^+(v, s) + \sum_k \delta_1(v^k) \beta [f_{ij}^+(v - e_k, s) - f_{ij}^+(v, s)] \\ &\quad + \sum_k \delta_0(v^k) \alpha_k(v, w) [f_{ij}^+(v + e_k, s - s^k e_k) - f_{ij}^+(v, s)] \\ &= \delta_0(v^i) \alpha_i(v, w) p^{+'}(s^j, w^{ij}) + \delta_1(v^i) \beta p^+(s^j, w^{ij}) - \delta_0(v^i) \alpha_i(v, w)^2 p^+(s^j, w^{ij}) \\ &\quad + \delta_0(v^j) \alpha_j(v, w) \delta_0(v^i) \alpha_i(v, w) (p^+(0, w^{ij}) - p^+(s^j, w^{ij})) \end{aligned}$$

and similarly

$$\begin{aligned} \mathcal{B}_w f_{ij}^-(v, s) &= \delta_0(v^j) \alpha_j(v, w) p^{-l}(s^i, w^{ij}) + \delta_1(v^j) \beta p^-(s^i, w^{ij}) - \delta_0(v^j) \alpha_j(v, w)^2 p^-(s^i, w^{ij}) \\ &\quad + \delta_0(v^i) \alpha_i(v, w) \delta_0(v^j) \alpha_j(v, w) (p^-(0, w^{ij}) - p^-(s^i, w^{ij})). \end{aligned}$$

$$\begin{aligned} \bar{\mathcal{B}}_j f_{ij}^+(v, s) &= \delta_0(v^i) \alpha_i(v, w) p^{+l}(s^j, w^{ij}) + \delta_1(v^i) \beta p^+(s^j, w^{ij}) - \alpha_m \delta_0(v^i) \alpha_i(v, w) p^+(s^j, w^{ij}) \\ &\quad + \delta_0(v^j) \alpha_m \delta_0(v^i) \alpha_i(v, w) (p^+(0, w^{ij}) - p^+(s^j, w^{ij})). \end{aligned}$$

$$\begin{aligned} \bar{\mathcal{B}}_j f_{ij}^-(v, s) &= \delta_0(v^j) \alpha_j(v, w) p^{-l}(s^i, w^{ij}) + \delta_1(v^j) \beta p^-(s^i, w^{ij}) - \alpha_m \delta_0(v^j) \alpha_j(v, w) p^-(s^i, w^{ij}) \\ &\quad + \delta_0(v^i) \alpha_m \delta_0(v^j) \alpha_j(v, w) (p^-(0, w^{ij}) - p^-(s^i, w^{ij})). \end{aligned}$$

Thus, we note that the term of transport and the one of jump of the potential from 1 to 0 cancel when computing $(\bar{\mathcal{B}}_j - \mathcal{B}_w)(f_{ij}^+ - f_{ij}^-)$. We finally obtain that

$$\begin{aligned} &(\bar{\mathcal{B}}_j - \mathcal{B}_w)(f_{ij}^+ - f_{ij}^-)(v, s) \\ &= \delta_0(v^j) (\alpha_m - \alpha_j(v, w)) \left(\delta_0(v^i) \alpha_i(v, w) (p^+(0, w^{ij}) - p^+(s^j, w^{ij})) + \alpha_j(v, w) p^-(s^i, w^{ij}) \right) \\ &\quad - \delta_0(v^i) (\alpha_m - \alpha_i(v, w)) \left(\delta_0(v^j) \alpha_j(v, w) (p^-(0, w^{ij}) - p^-(s^i, w^{ij})) + \alpha_i(v, w) p^+(s^j, w^{ij}) \right). \end{aligned}$$

Under the assumption that p^+ and p^- are decreasing functions with respect to their first variable, we obtain that for all $(v, s) \in E_1$, $w \in E_2$,

$$(\bar{\mathcal{B}}_j - \mathcal{B}_w)(f_{ij}^+ - f_{ij}^-)(v, s) \geq 0.$$

We deduce from equation (3.32) that for all $t \geq 0$,

$$\pi_w \bar{P}_{t,j}(f_{ij}^+ - f_{ij}^-) - \pi_w(f_{ij}^+ - f_{ij}^-) \geq 0.$$

Taking the large time limit in this inequality, we deduce from Proposition 3.27 the upper bound given in the first line of (3.31).

We obtain the lower bound by similar computations with $(\underline{V}_{t,j}, \underline{S}_{t,j})_{t \geq 0}$ instead of $(\bar{V}_{t,j}, \bar{S}_{t,j})_{t \geq 0}$. It ends the proof. \square

Corollary 3.29. *Assume that Assumption 3.21 holds. Then, the process $(\bar{W}_t)_{t \geq 0}$ is positive recurrent if for all i, j*

$$\lim_{w_{lim} \rightarrow +\infty} \sup_{\substack{w \in E_2, \\ w^{ij} = w_{lim}}} \left(\mathbb{E} \left[\delta_0(\bar{V}_j^i) \alpha_i(\bar{V}_j, w) p^+(\bar{S}_j^j, w_{lim}) \right] - \mathbb{E} \left[\delta_0(\bar{V}_j^j) \alpha_j(\bar{V}_j, w) p^-(\bar{S}_j^i, w_{lim}) \right] \right) < 0,$$

and it is transient if

$$\lim_{w_{lim} \rightarrow +\infty} \inf_{\substack{w \in E_2, \\ w^{ij} = w_{lim}}} \left(\mathbb{E} \left[\delta_0(\underline{V}_i^i) \alpha_i(\underline{V}_i, w) p^+(\underline{S}_i^j, w_{lim}) \right] - \mathbb{E} \left[\delta_0(\underline{V}_i^j) \alpha_j(\underline{V}_i, w) p^-(\underline{S}_i^i, w_{lim}) \right] \right) > 0.$$

We now consider a special case where the functions p^+ and p^- have factorized expressions:

Assumption 3.30. *There exist parameters $A_+, A_- \in (0, 1)$, $\tau_+, \tau_- \in \mathbb{R}_*^+$, and functions $g^+ : \mathbb{N}^* \rightarrow \frac{1}{A_+}$ and $g^- : \mathbb{N}^* \rightarrow \frac{1}{A_-}$ with $g^-(1) = 0$ and for any $w > 1$, $g^-(w) > 0$, such that*

$$p^+(\hat{s}, \hat{w}) = A_+ e^{\frac{-\hat{s}}{\tau_+}} g^+(\hat{w}) \quad \text{and} \quad p^-(\hat{s}, \hat{w}) = A_- e^{\frac{-\hat{s}}{\tau_-}} g^-(\hat{w}). \quad (3.33)$$

This assumption is satisfied by the classical functions used in STDP modelling. They match very well with the experimental results obtained in [17].

Proposition 3.31. *Grant Assumption 3.30. Then, the process $(\bar{W}_t)_{t \geq 0}$ converges to its unique invariant measure if*

$$\lim_{w_{lim} \rightarrow +\infty} \frac{\alpha_M^2 (\beta - \alpha_m) g^+(w_{lim}) A_+ \tau_+ (\alpha_M (\alpha_M \tau_+ + 1) - \beta (\beta \tau_+ + 1)) (\alpha_m \tau_- + 1) (\beta \tau_- + 1)}{\alpha_m^2 (\alpha_M - \beta) g^-(w_{lim}) A_- \tau_- (\beta (\beta \tau_- + 1) - \alpha_m (\alpha_m \tau_- + 1)) (\alpha_M \tau_+ + 1) (\beta \tau_+ + 1)} < 1. \quad (3.34)$$

Moreover, it is transient if

$$\lim_{w_{lim} \rightarrow +\infty} \frac{\alpha_m^2 (\alpha_M - \beta) g^+(w_{lim}) A_+ \tau_+ (\alpha_m (\alpha_m \tau_+ + 1) - \beta (\beta \tau_+ + 1)) (\alpha_M \tau_- + 1) (\beta \tau_- + 1)}{\alpha_M^2 (\alpha_m - \beta) g^-(w_{lim}) A_- \tau_- (\alpha_M (\alpha_M \tau_- + 1) - \beta (\beta \tau_- + 1)) (\alpha_m \tau_+ + 1) (\beta \tau_+ + 1)} > 1.$$

Proof. It is an application of Corollary 3.29. We present the computations only for the positive recurrent case. We find that for all $i \neq j$, $w \in E_2$,

$$\begin{aligned} r_{ij}^+(w) - r_{ij}^-(w) &\leq \alpha_M \mathbb{E} \left[\delta_0(\bar{V}_j^i) p^+(\bar{S}_j^i, w^{ij}) \right] - \alpha_m \mathbb{E} \left[\delta_0(\bar{V}_j^j) p^-(\bar{S}_j^j, w^{ij}) \right] \\ &= \alpha_M \mathbb{P}(\bar{V}_j^i = 0) \mathbb{E} \left[p^+(\bar{S}_j^i, w^{ij}) \right] - \alpha_m \mathbb{P}(\bar{V}_j^j = 0) \mathbb{E} \left[p^-(\bar{S}_j^j, w^{ij}) \right] \\ &= \alpha_M \frac{\beta}{\alpha_m + \beta} \int_{\mathbb{R}^+} p^+(\hat{s}, w^{ij}) \mu_{\alpha_M}(\{0, 1\}, d\hat{s}) \\ &\quad - \alpha_m \frac{\beta}{\alpha_M + \beta} \int_{\mathbb{R}^+} p^-(\hat{s}, w^{ij}) \mu_{\alpha_m}(\{0, 1\}, d\hat{s}), \end{aligned}$$

where μ_α is defined in Proposition 3.27 and satisfies:

$$\mu_\alpha(\{0, 1\}, d\hat{s}) = \frac{\alpha\beta}{\alpha^2 - \beta^2} \left(\alpha e^{-\beta\hat{s}} - \beta e^{-\alpha\hat{s}} \right) d\hat{s}.$$

We conclude with the following computations:

$$\begin{aligned} r_{ij}^+(w) - r_{ij}^-(w) &\leq \alpha_M \frac{\beta}{\alpha_m + \beta} \frac{\alpha_M \beta}{\alpha_M^2 - \beta^2} \int_{\mathbb{R}^+} A_+ e^{\frac{-\hat{s}}{\tau_+}} g^+(w^{ij}) \left(\alpha_M e^{-\beta\hat{s}} - \beta e^{-\alpha_M \hat{s}} \right) d\hat{s} \\ &\quad - \alpha_m \frac{\beta}{\alpha_M + \beta} \frac{\alpha_m \beta}{\alpha_m^2 - \beta^2} \int_{\mathbb{R}^+} A_- e^{\frac{-\hat{s}}{\tau_-}} g^-(w^{ij}) \left(\alpha_m e^{-\beta\hat{s}} - \beta e^{-\alpha_m \hat{s}} \right) d\hat{s} \\ &\leq \frac{\alpha_M^2 \beta^2}{(\alpha_m + \beta)(\alpha_M^2 - \beta^2)} A_+ g^+(w^{ij}) \left(\frac{\alpha_M}{\beta + \frac{1}{\tau_+}} - \frac{\beta}{\alpha_M + \frac{1}{\tau_+}} \right) \\ &\quad - \frac{\alpha_m^2 \beta^2}{(\alpha_M + \beta)(\alpha_m^2 - \beta^2)} A_- g^-(w^{ij}) \left(\frac{\alpha_m}{\beta + \frac{1}{\tau_-}} - \frac{\beta}{\alpha_m + \frac{1}{\tau_-}} \right). \end{aligned}$$

Then, the limit (3.34) gives the result by applying Corollary 3.25. \square

In this section, we propose a method to verify the conditions for the convergence of the limit process $(\overline{W}_t)_{t \geq 0}$ given by Proposition 3.24, see Proposition 3.28 and Corollary 3.29. We applied these results in Proposition 3.31 where we give explicit conditions in function of the parameters of the model on an example. This is the first time, to our knowledge, that such conditions can be given without adding any homeostatic mechanism and without any inhibitory neurons. With the latter, our model becomes ready to use, being aware of the criticizes that we present in the Discussion, see Section 3.6. In the following, we propose another method to test the conditions given in Proposition 3.24 using the Laplace transform of the invariant measure of the fast process π_w .

3.4.2 Condition using the Laplace transform

We begin with two notations.

Notation 3.32. We denote by \mathbb{C}_+^N the elements of \mathbb{C}^N with positive real part. We denote by $\mathcal{L}\{\rho\}$ the Laplace transform of the measure ρ on \mathbb{R}_+^N . It is defined by:

$$\forall \lambda = (\lambda^1, \dots, \lambda^N) \in \mathbb{C}_+^N, \quad \mathcal{L}\{\rho\}(\lambda) = \int_{\mathbb{R}_+^N} e^{-\lambda \cdot s} \rho(ds),$$

where $\lambda \cdot s = \sum_{k=1}^N \lambda^k s^k$.

Remark 3.33. Although the general definition of the Laplace transform lies on $\lambda \in \mathbb{C}_+^N$, in the following, we only use $\lambda \in \mathbb{R}_+^N$.

Notation 3.34. We denote by ν_w^v the marginal in v of the distribution π_w

$$\forall w \in E_2, \forall v \in \{0, 1\}^N, \quad \nu_w^v \stackrel{\text{def.}}{=} \pi_w(\{v\}, \mathbb{R}_+^N).$$

In Proposition 3.39, we give a way to compute ν_w^v and we show that $\nu_w^v > 0$ for all v, w .

Moreover, we denote by π_w^v the law of π_w conditionally to v

$$\forall w \in E_2, \forall v \in \{0, 1\}^N, \quad \pi_w^v(ds) \stackrel{\text{def.}}{=} \frac{\pi_w(\{v\}, ds)}{\nu_w^v}.$$

Under Assumption 3.30, we can easily simplify the expressions of r_{ij}^+ and r_{ij}^- using the Laplace transforms of π_w^v . Using (3.24) and (3.25), one has

$$\begin{aligned} r_{ij}^+(w) &= \sum_v \delta_0(v^i) \alpha_i(v, w) \int_{\mathbb{R}_+^N} p^+(s^j, w^{ij}) \nu_w^v \pi_w^v(ds) \\ &= \sum_v \delta_0(v^i) \alpha_i(v, w) \int_{\mathbb{R}_+^N} A_+ e^{-\frac{s^j}{\tau_+}} g^+(w^{ij}) \nu_w^v \pi_w^v(ds) \\ &= \sum_v \delta_0(v^i) \alpha_i(v, w) g^+(w^{ij}) A_+ \nu_w^v \underbrace{\mathcal{L}\{\pi_w^v\}}_j(0, \dots, 0, \underbrace{\frac{1}{\tau_+}}_j, 0, \dots, 0), \\ r_{ij}^-(w) &= \sum_v \delta_0(v^j) \alpha_j(v, w) g^-(w^{ij}) A_- \nu_w^v \underbrace{\mathcal{L}\{\pi_w^v\}}_i(0, \dots, 0, \underbrace{\frac{1}{\tau_-}}_i, 0, \dots, 0). \end{aligned} \quad (3.35)$$

Assumption 3.30 makes appear immediately the Laplace transform of the measures π_w^v . In fact, it can also be used under more general assumptions.

Remark 3.35. *More generally, under some simple assumptions, we can get explicitly the dynamics of the weights which is a Markov process on E_2 with jump rates depending on the Laplace transform of π_w . Suppose that for all i and $\tilde{w}, w \in E_2$, there exists $\Phi_{(\tilde{w}, w)}^i$ such that $\varphi_i(s, \tilde{w}, w) = \mathcal{L}\{\Phi_{(\tilde{w}, w)}^i\}(s) = \int_{\mathbb{R}_+^N} e^{-\lambda \cdot s} \Phi_{(\tilde{w}, w)}^i(\lambda) d\lambda$. Then, from Fubini's theorem*

$$\begin{aligned} \mathcal{C}_{av}f(w) &= \sum_{\tilde{w} \in E_2} (f(\tilde{w}) - f(w)) \left(\sum_{v \in I} \nu_w^v \sum_i \delta_0(v^i) \alpha_i(v, w) \int_{\mathbb{R}_+^N} \mathcal{L}\{\Phi_{(\tilde{w}, w)}^i\}(s) \pi_w^v(s) (ds) \right) \\ &= \sum_{\tilde{w} \in E_2} (f(\tilde{w}) - f(w)) \left(\sum_{v \in I} \nu_w^v \sum_i \delta_0(v^i) \alpha_i(v, w) \int_{\mathbb{R}_+^N} \Phi_{(\tilde{w}, w)}^i(\lambda) \mathcal{L}\{\pi_w^v\}(\lambda) (d\lambda) \right). \end{aligned}$$

In the same spirit, suppose that for all $\hat{w} \in \mathbb{N}_*$ there exist $P_{\hat{w}}^+$ and $P_{\hat{w}}^-$ such that $p^+(\cdot, \hat{w}) = \mathcal{L}\{P_{\hat{w}}^+\}(\cdot)$ and $p^-(\cdot, \hat{w}) = \mathcal{L}\{P_{\hat{w}}^-\}(\cdot)$, then

$$\begin{aligned} r_{ij}^+(w) &= \sum_v \delta_0(v^i) \alpha_i(v, w) \nu_w^v \int_{\mathbb{R}_+} P_{w^{ij}}^+(\lambda) \mathcal{L}\{\pi_w^v\}(0, \dots, 0, \underbrace{\lambda}_j, 0, \dots, 0) d\lambda, \\ r_{ij}^-(w) &= \sum_v \delta_0(v^j) \alpha_j(v, w) \nu_w^v \int_{\mathbb{R}_+} P_{w^{ij}}^-(\lambda) \mathcal{L}\{\pi_w^v\}(0, \dots, 0, \underbrace{\lambda}_i, 0, \dots, 0) d\lambda. \end{aligned}$$

All these previous special cases show that the Laplace transform of both π_w and its marginals is a tool adapted to compute the $(r_{ij}^\pm)_{1 \leq i \neq j \leq N}$. By this way, we obtain information on the long term behaviour of $(\overline{W}_t)_{t \geq 0}$ through Corollary 3.25. In the following, we give a method to compute these Laplace transforms and in particular, we apply this method in the toy model with two neurons.

An example with two neurons

In this paragraph, we consider a toy model with two neurons. With this very simple setting, we can describe in details a method efficient to obtain the Laplace transforms of the four probability measures on \mathbb{R}_+^2 : $(\pi_w^v)_{v \in \{0,1\}^2}$. Method that we extend to the general case (N neurons) in Proposition 3.39.

The basic idea of the method is to use the definition of the invariant measure π_w

$$\forall f \in D(\mathcal{B}_w), \quad \sum_{v \in \{0,1\}^2} \int_{\mathbb{R}_+^2} \mathcal{B}_w f(v, s) \nu_w^v \pi_w^v(ds) = 0. \quad (3.36)$$

with some well-chosen functions f making appear the Laplace transforms.

In the following, for all $x, y \in \{0, 1\}$, the shorter notation xy will be used instead of (x, y) . Using the previous decomposition of the measure π_w , see Notation 3.34, we denote by Υ_{π_w} the vector of the Laplace transforms of $\nu_w^v \pi_w^v$

$$\forall \lambda^1, \lambda^2 \in \mathbb{R}^+, \quad \Upsilon_{\pi_w}(\lambda^1, \lambda^2) = \begin{bmatrix} \nu_w^{00} \mathcal{L}\{\pi_w^{00}\}(\lambda^1, \lambda^2) \\ \nu_w^{01} \mathcal{L}\{\pi_w^{01}\}(\lambda^1, \lambda^2) \\ \nu_w^{10} \mathcal{L}\{\pi_w^{10}\}(\lambda^1, \lambda^2) \\ \nu_w^{11} \mathcal{L}\{\pi_w^{11}\}(\lambda^1, \lambda^2) \end{bmatrix}.$$

After computations for all $\tilde{v} \in \{0, 1\}^2$, we get equation (3.37) with

$$M(\lambda) = \begin{bmatrix} \alpha_2(00, w) + \alpha_1(00, w) & -\beta & -\beta & 0 \\ 0 & \alpha_1(01, w) + \beta & 0 & -\beta \\ 0 & 0 & \alpha_2(10, w) + \beta & -\beta \\ 0 & 0 & 0 & 2\beta \end{bmatrix} + (\lambda^1 + \lambda^2)I_4.$$

We obtain the matrices M_1 and M_2 evaluating equation (3.37) on $\lambda = (\lambda^1, 0)$ and $\lambda = (0, \lambda^2)$ and using the fact that for all $v \in \{0, 1\}^2$, $\mathcal{L}\{\pi_w^v\}(0, 0) = 1$:

$$M_1(\lambda^1) = \begin{bmatrix} \alpha_2(00, w) + \alpha_1(00, w) & -\beta & -\beta & 0 \\ -\alpha_2(00, w) & \alpha_1(01, w) + \beta & 0 & -\beta \\ 0 & 0 & \alpha_2(10, w) + \beta & -\beta \\ 0 & 0 & -\alpha_2(10, w) & 2\beta \end{bmatrix} + \lambda^1 I_4$$

and

$$M_2(\lambda^2) = \begin{bmatrix} \alpha_2(00, w) + \alpha_1(00, w) & -\beta & -\beta & 0 \\ 0 & \alpha_1(01, w) + \beta & 0 & -\beta \\ -\alpha_1(00, w) & 0 & \alpha_2(10, w) + \beta & -\beta \\ 0 & -\alpha_1(01, w) & 0 & 2\beta \end{bmatrix} + \lambda^2 I_4.$$

Finally, from assumption (3.1), for all $\lambda^1, \lambda^2 \in \mathbb{R}^+$, $M(\lambda)$ is an upper triangular matrix with strictly positive diagonal elements and for $\lambda^1 > 0$, $\lambda^2 > 0$, the matrices $M_1(\lambda_1)$ and $M_2(\lambda_2)$ are strictly diagonally dominant matrices. Thereby, these three matrices are invertible. In $\lambda = (0, 0)$ we have $\Upsilon_{\pi_w}(0, 0) = \nu_w$. \square

Remark 3.37. We note that using equation (3.37), we can obtain the Laplace transform of π_w from the knowledge of $\Upsilon_{\pi_w}(\lambda^1, 0)$ and $\Upsilon_{\pi_w}(0, \lambda^2)$. We use this idea in the proof of Proposition 3.39.

Under Assumption 3.30 (p^+ and p^- are exponential), this lemma enables us to compute for $i, j \neq i \in \{1, 2\}$, r_{ij}^+ and r_{ij}^- following these steps. First, we find the kernel of Q which gives us the vector ν_w . Then, we compute the inverse of the matrices $M_1(\lambda^1)$ and $M_2(\lambda^2)$ to obtain the vectors $\Upsilon_{\pi_w}(0, \lambda^2)$ and $\Upsilon_{\pi_w}(\lambda^1, 0)$. We finally deduce the r_{ij}^\pm from equation (3.35),

$$\begin{aligned} r_{12}^+(w) &= A_+ \left[\alpha_1(00, w) \nu_w^{00} \mathcal{L}\{\pi_w^{00}\} \left(0, \frac{1}{\tau_+}\right) + \alpha_1(01, w) \nu_w^{01} \mathcal{L}\{\pi_w^{01}\} \left(0, \frac{1}{\tau_+}\right) \right] \\ &= A_+ \left[\alpha_1(00, w) \Upsilon_{\pi_w}^1 \left(0, \frac{1}{\tau_+}\right) + \alpha_1(01, w) \Upsilon_{\pi_w}^2 \left(0, \frac{1}{\tau_+}\right) \right] \\ r_{12}^-(w) &= \mathbb{1}_{[2, +\infty[(w^{12})} A_- \left[\alpha_2(00, w) \Upsilon_{\pi_w}^1 \left(\frac{1}{\tau_-}, 0\right) + \alpha_2(10, w) \Upsilon_{\pi_w}^3 \left(\frac{1}{\tau_-}, 0\right) \right] \\ r_{21}^+(w) &= A_+ \left[\alpha_2(00, w) \Upsilon_{\pi_w}^1 \left(\frac{1}{\tau_+}, 0\right) + \alpha_2(10, w) \Upsilon_{\pi_w}^3 \left(\frac{1}{\tau_+}, 0\right) \right] \\ r_{21}^-(w) &= \mathbb{1}_{[2, +\infty[(w^{21})} A_- \left[\alpha_1(00, w) \Upsilon_{\pi_w}^1 \left(0, \frac{1}{\tau_-}\right) + \alpha_1(01, w) \Upsilon_{\pi_w}^2 \left(0, \frac{1}{\tau_-}\right) \right]. \end{aligned}$$

From the $(r_{ij}^\pm)_{i, j \neq i \in \{1, 2\}}$ and applying Corollary 3.25, we can determine the long time behaviour of the process $(\overline{W}_t)_{t \geq 0}$ in function of the parameters. We used these computations for numerical

applications in Section 3.5.1. Moreover, under Assumption 3.9, the $(r_{ij}^{\pm})_{i,j \neq i \in \{1,2\}}$ are exactly the jumping rates of $(\bar{W}_t)_{t \geq 0}$, see Remark 3.20.

A last application of this lemma is to obtain explicitly π_w . After inverting the matrix $M(\lambda)$, we obtain the Laplace transform of π_w . Unfortunately, inverting it gives too long formulas to be shown here. We give more details on the form of π_w in Appendix A.3.

General results on the Laplace transform of the invariant measure (fast process)

We perform similar computations as in Lemma 3.36 but now with a general N : for all $\lambda \in \mathbb{R}_+^N$, we compute the matrices $M^{(N)}(\lambda)$ and $(M_l^{(N)}(\lambda^l))_{1 \leq l \leq N}$. The linear systems obtained are however more complicated in this generalisation. Therefore, we need to introduce some new notations.

First, let (v_1, \dots, v_{2N}) be an enumeration of $\{0, 1\}^N$ such that

$$k \geq l \quad \Rightarrow \quad \sum_{i=1}^N v_k^i \geq \sum_{i=1}^N v_l^i. \quad (3.39)$$

We will see that such an enumeration makes the matrix $M^{(N)}(\lambda)$ upper triangular. As an example, this property is satisfied in the case $N = 2$, see the proof of Lemma 3.36. Then, for all $\lambda \in \mathbb{R}_+^N$, we define $|\lambda| \stackrel{\text{def.}}{=} \sum_{i=1}^N \lambda^i$.

Notation 3.38. For all $\lambda \in \mathbb{R}_+^N$, for any sequence $l_1, l_2, \dots, l_d \in \llbracket 1, N \rrbracket$, $d \leq N$, we define $\hat{\lambda}_{l_1, \dots, l_d} \in \mathbb{R}_+^N$ such that

$$\forall i \in \{l_1, \dots, l_d\}, \quad \hat{\lambda}_{l_1, \dots, l_d}^i = 0 \quad \text{and} \quad \forall i \in \llbracket 1, N \rrbracket \setminus \{l_1, \dots, l_d\}, \quad \hat{\lambda}_{l_1, \dots, l_d}^i = \lambda^i.$$

Finally, we denote by

$$\check{\lambda}_l \stackrel{\text{def.}}{=} \hat{\lambda}_{1, \dots, l-1, l+1, \dots, N} = (0, \dots, 0, \lambda^l, 0, \dots, 0).$$

Following Lemma 3.36 in two dimension, we obtain the following one.

Proposition 3.39. For all $\lambda \in \mathbb{R}_+^N$, the Laplace transform $\Upsilon_{\pi_w}(\lambda)$ is a linear combination of the finite family $(\mathcal{L}\{\pi_w^{v_k}\}(\check{\lambda}_l))_{\substack{1 \leq k \leq 2^N \\ 1 \leq l \leq N}}$.

Moreover, the elements of this family satisfy

$$\forall 1 \leq l \leq N, \quad \Upsilon_{\pi_w}(\check{\lambda}_l) = M_l^{(N)}(\lambda^l)^{-1} \Theta_l,$$

where the matrices $M_l^{(N)}$ and the vectors Θ_l are defined as follows: for all $1 \leq j, k \leq 2^N$,

$$\begin{cases} M_l^{jk, (N)}(\lambda^l) = \delta_0(v_k^l) \alpha_l(v_k, w) \delta_{v_j}(v_k + e_l) & \forall k < j \\ M_l^{jj, (N)}(\lambda^l) = \left(\sum_{i=1}^N \delta_1(v_j^i) \beta + \delta_0(v_j^i) \alpha_i(v_j, w) \right) + \lambda^l \\ M_l^{jk, (N)}(\lambda^l) = - \sum_{i=1}^N \delta_1(v_k^i) \beta \delta_{v_j}(v_k - e_i) & \forall k > j \end{cases}$$

and for all $1 \leq j \leq 2^N$,

$$\Theta_l^j = \sum_{k=1}^{2^N} \delta_0(v_k^l) \alpha_l(v_k, w) \delta_{v_j}(v_k + e_l) \nu_w^{v_k}.$$

Moreover, $\nu_w \stackrel{\text{def.}}{=} (\nu_w^{ij})_{i,j \in \{0,1\}}$ is the invariant probability measure of the Markov process $(V_t)_{t \geq 0}$ with infinitesimal generator

$$Q^{jk} = \sum_{i=1}^N \delta_1(v_k^i) \beta \delta_{v_j}(v_k - e_i) + \delta_0(v_k^i) \alpha_i(v_k, w) \delta_{v_j}(v_k + e_i) \quad \text{and} \quad Q^{jj} = - \sum_{k \neq j} Q^{jk}. \quad (3.40)$$

In addition, for all $1 \leq k \leq 2^N$, $\nu_w^{v_k} \neq 0$.

Proof. We prove the first assertion showing that for all $l_1, l_2, \dots, l_d \in \llbracket 1, N \rrbracket$, $d \leq N - 1$, $v \in \{0, 1\}^N$, $\mathcal{L}\{\pi_w^v\}(\tilde{\lambda}_{l_1, \dots, l_d})$ is a linear combination of $\left(\mathcal{L}\{\pi_w^{v_l}\}(\tilde{\lambda}_{l_1, \dots, l_d, l_{d+1}}) \right)_{\substack{1 \leq l \leq 2^N \\ l_{d+1} \in \llbracket 1, 2^N \rrbracket \setminus \{l_1, \dots, l_d\}}}$.

We conclude the first point by recurrence. We then show how to compute the elements of the family $(\mathcal{L}\{\pi_w^{v_k}\}(\tilde{\lambda}_l))_{\substack{1 \leq k \leq 2^N \\ 1 \leq l \leq N}}$. Finally, we find the matrix Q satisfying (3.40).

By definition of π_w , we have

$$\forall f \in D(\mathcal{B}_w), \quad \pi_w \mathcal{B}_w f = \sum_{k=1}^{2^N} \int_{\mathbb{R}_+^N} \mathcal{B}_w f(v_k, s) \nu_w^{v_k} \pi_w^{v_k}(ds) = 0, \quad (3.41)$$

where $(\mathcal{B}_w, D(\mathcal{B}_w))$ is the generator of the process $(V_t, S_t, W_t \equiv w)_{t \geq 0}$, see (3.6). Let us take $f(v, s) = e^{-\lambda \cdot s} g(v)$ which gives

$$\begin{aligned} \mathcal{B}_w f(v, s) &= \sum_{i=1}^N \delta_0(v^i) \alpha_i(v, w) \left[e^{-\lambda \cdot (s - s^i e_i)} g(v + e_i) - e^{-\lambda \cdot s} g(v) \right] \\ &\quad + \sum_{i=1}^N \delta_1(v^i) \beta \left[e^{-\lambda \cdot s} g(v - e_i) - e^{-\lambda \cdot s} g(v) \right] \\ &\quad - \underbrace{\left(\sum_{i=1}^N \lambda^i \right)}_{|\lambda|} e^{-\lambda \cdot s} g(v). \end{aligned}$$

Hence, by equation (3.41) we obtain that

$$\begin{aligned} \pi_w \mathcal{B}_w f &= \sum_{k=1}^{2^N} \sum_{i=1}^N \delta_0(v_k^i) \alpha_i(v_k, w) g(v_k + e_i) \underbrace{\int_{\mathbb{R}_+^N} e^{-\lambda \cdot (s - s^i e_i)} \nu_w^{v_k} \pi_w^{v_k}(ds)}_{\nu_w^{v_k} \mathcal{L}\{\pi_w^{v_k}\}(\tilde{\lambda}_i)} \\ &\quad + \sum_{k=1}^{2^N} \underbrace{\int_{\mathbb{R}_+^N} e^{-\lambda \cdot s} \nu_w^{v_k} \pi_w^{v_k}(ds)}_{\nu_w^{v_k} \mathcal{L}\{\pi_w^{v_k}\}(\lambda)} \left[\left(\sum_{i=1}^N \delta_1(v_k^i) \beta [g(v_k - e_i) - g(v_k)] - \delta_0(v_k^i) \alpha_i(v_k, w) g(v_k) \right) \right. \\ &\quad \left. - |\lambda| g(v_k) \right] \\ &= 0. \end{aligned}$$

Taking $g(v) = -\delta_{v_j}(v)$ in the previous equation, we find that

$$\begin{aligned} \sum_{k=1}^{2^N} \nu_w^{v_k} \mathcal{L}\{\pi_w^{v_k}\}(\lambda) & \left[\underbrace{|\lambda| \delta_{v_j}(v_k) + \sum_{i=1}^N \delta_1(v_k^i) \beta [\delta_{v_j}(v_k) - \delta_{v_j}(v_k - e_i)] + \delta_0(v_k^i) \alpha_i(v_k, w) \delta_{v_j}(v_k)}_{\stackrel{\text{def.}}{=} M^{jk, (N)}(\lambda)} \right] \\ & = \underbrace{\sum_{k=1}^{2^N} \sum_{i=1}^N \alpha_i(v_k, w) \delta_0(v_k^i) \delta_{v_j}(v_k + e_i) \nu_w^{v_k} \mathcal{L}\{\pi_w^{v_k}\}(\widehat{\lambda}_i)}_{\stackrel{\text{def.}}{=} \Lambda^j(\lambda)}. \end{aligned} \quad (3.42)$$

Thus, denoting by $M^{(N)} = (M^{jk, (N)})_{1 \leq j, k \leq 2^N}$ and $\Lambda = (\Lambda^j)_{1 \leq j \leq 2^N}$, we have

$$\forall \lambda \in \mathbb{R}_+^N, \quad M^{(N)}(\lambda) \Upsilon_{\pi_w}(\lambda) = \Lambda(\lambda). \quad (3.43)$$

In particular, we note that

$$\begin{cases} M^{jk, (N)}(\lambda) = 0 & \forall k < j \\ M^{jj, (N)}(\lambda) = \left(\sum_{i=1}^N \delta_1(v_j^i) \beta + \delta_0(v_j^i) \alpha_i(v_j, w) \right) + |\lambda| \\ M^{jk, (N)}(\lambda) = - \sum_{i=1}^N \delta_1(v_k^i) \beta \delta_{v_j}(v_k - e_i) & \forall k > j. \end{cases}$$

Hence, thanks to the enumeration property (3.39), $M^{(N)}$ is upper triangular with strictly positive terms on the diagonal, it is thus invertible.

Now, take a sequence $l_1, l_2, \dots, l_d \in \llbracket 1, N \rrbracket$, $d \leq N - 1$. We have from equation (3.43)

$$M^{(N)}(\widehat{\lambda}_{l_1, \dots, l_d}) \Upsilon_{\pi_w}(\widehat{\lambda}_{l_1, \dots, l_d}) = \Lambda(\widehat{\lambda}_{l_1, \dots, l_d}).$$

Using equation (3.42), we get

$$\begin{aligned} \Lambda^j(\widehat{\lambda}_{l_1, \dots, l_d}) & = \sum_{k=1}^{2^N} \left[\sum_{i \in \llbracket 1, N \rrbracket \setminus \{l_1, \dots, l_d\}} \delta_0(v_k^i) \alpha_i(v_k, w) \delta_{v_j}(v_k + e_i) \nu_w^{v_k} \mathcal{L}\{\pi_w^{v_k}\}(\widehat{\lambda}_{l_1, \dots, l_d, i}) \right. \\ & \quad \left. + \underbrace{\sum_{i \in \{l_1, \dots, l_d\}} \delta_0(v_k^i) \alpha_i(v_k, w) \delta_{v_j}(v_k + e_i) \nu_w^{v_k} \mathcal{L}\{\pi_w^{v_k}\}(\widehat{\lambda}_{l_1, \dots, l_d})}_{M'^{jk, (N)}_{l_1, \dots, l_d}} \right] \end{aligned} \quad (3.44)$$

Hence, we can decompose $\Lambda(\widehat{\lambda}_{l_1, \dots, l_d})$ as follows:

$$\Lambda(\widehat{\lambda}_{l_1, \dots, l_d}) = \Lambda_{l_1, \dots, l_d}(\lambda) + M'^{(N)}_{l_1, \dots, l_d} \Upsilon_{\pi_w}(\widehat{\lambda}_{l_1, \dots, l_d}),$$

where $\Lambda_{l_1, \dots, l_d}(\lambda)$ depends on λ only through $(\mathcal{L}\{\pi_w^{v_k}\}(\widehat{\lambda}_{l_1, \dots, l_d, i}))_{i \in \llbracket 1, N \rrbracket \setminus \{l_1, \dots, l_d\}, 1 \leq k \leq 2^N}$. Note that thanks to the enumeration property (3.39) of the set $\{0, 1\}^N$ and the Dirac measures $\delta_{v_j}(v_k + e_i)$, we have that $M'^{jk, (N)}_{l_1, \dots, l_d} = 0$ for all $k \geq j$.

Thus, we obtain that

$$\left[M^{(N)}(\widehat{\lambda}_{l_1, \dots, l_d}) - M'_{l_1, \dots, l_d} \right] \Upsilon_{\pi_w}(\widehat{\lambda}_{l_1, \dots, l_d}) = \Lambda_{l_1, \dots, l_d}(\lambda).$$

Eventually, we show that $M^{(N)}(\widehat{\lambda}_{l_1, \dots, l_d}) - M'_{l_1, \dots, l_d}$ is invertible for all λ . Denoting by $K = \{l_1, \dots, l_d\}$, we have for all k ,

$$\begin{aligned} & \sum_{j \neq k} |M^{jk, (N)}(\widehat{\lambda}_{l_1, \dots, l_d}) - M'^{jk, (N)}_{l_1, \dots, l_d}| \\ &= \sum_{j > k} |M'^{jk, (N)}_{l_1, \dots, l_d}| + \sum_{j < k} |M^{jk, (N)}(\widehat{\lambda}_{l_1, \dots, l_d})| \\ &= \sum_{j \neq k} \left(\sum_{i \in K} \delta_0(v_k^i) \alpha_i(v_k, w) \delta_{v_j}(v_k + e_i) + \sum_{i=1}^N \delta_1(v_k^i) \beta \delta_{v_j}(v_k - e_i) \right) \\ &\leq \sum_{j \neq k} \sum_{i=1}^N (\delta_0(v_k^i) \alpha_i(v_k, w) \delta_{v_j}(v_k + e_i) + \delta_1(v_k^i) \beta \delta_{v_j}(v_k - e_i)) \\ &= \sum_{i=1}^N \delta_0(v_k^i) \alpha_i(v_k, w) + \delta_1(v_k^i) \beta = M^{kk, (N)}(\widehat{\lambda}_{l_1, \dots, l_d}) - |\widehat{\lambda}_{l_1, \dots, l_d}| \\ &< \underbrace{|M^{kk, (N)}(\widehat{\lambda}_{l_1, \dots, l_d})|}_{>0} - \underbrace{|M'^{kk, (N)}_{l_1, \dots, l_d}|}_{=0}. \end{aligned}$$

Hence, for all $|\widehat{\lambda}_{l_1, \dots, l_d}| > 0$, $M^{(N)}(\widehat{\lambda}_{l_1, \dots, l_d}) - M'_{l_1, \dots, l_d}$ is invertible as a strictly dominant diagonal matrix. Using the fact that for all $v \in \{0, 1\}^N$, $\mathcal{L}\{\pi_w^v\}(0, \dots, 0) = 1$, we conclude by recurrence that there is a unique way to express each $\mathcal{L}\{\pi_w^v\}(\lambda)$, $v \in \{0, 1\}^N$, as a linear combination of the terms of the family $(\mathcal{L}\{\pi_w^{v_k}\}(\check{\lambda}_l))_{1 \leq l \leq N, 1 \leq k \leq 2^N}$. It ends the proof of the first point.

Then, the second point is obtained by evaluating equation (3.43) in $\check{\lambda}_l$. For all $2^N \geq k \geq j \geq 0$, $M_l^{jk, (N)}(\lambda^l) = M^{jk, (N)}(\check{\lambda}_l)$. For all $0 \leq k \leq j \leq 2^N$, $M_l^{jk, (N)}(\lambda^l)$ corresponds to the second term of the right hand side of equation (3.44) where we replace i by l . Thus,

$$\left\{ \begin{array}{ll} M_l^{jk, (N)}(\lambda^l) = \delta_0(v_k^l) \alpha_l(v_k, w) \delta_{v_j}(v_k + e_l) & \forall k < j \\ M_l^{jj, (N)}(\lambda^l) = M^{jj, (N)}(\check{\lambda}_l) = \left(\sum_{i=1}^N \delta_1(v_j^i) \beta + \delta_0(v_j^i) \alpha_i(v_j, w) \right) + \lambda^l \\ M_l^{jk, (N)}(\lambda^l) = M^{jk, (N)}(\check{\lambda}_l) = - \sum_{i=1}^N \delta_1(v_k^i) \beta \delta_{v_j}(v_k - e_i) & \forall k > j. \end{array} \right.$$

Then, Θ_l is obtained from the first term of the right hand side of equation (3.44) by replacing i by l . As previously, for all l , $M_l^{(N)}$ is invertible as a strictly dominant diagonal matrix.

We finally show the last point. The dynamics of the process $(V_t)_{t \geq 0}$ does not depend on $(S_t)_{t \geq 0}$. Moreover, thanks to assumption (3.1), $(V_t)_{t \geq 0}$ is irreducible and aperiodic on the finite state space $\{0, 1\}^N$ so it converges in law to its unique invariant measure $\nu_w \stackrel{def.}{=} (\nu_w^{v_k})_{1 \leq k \leq 2^N} \in (0, 1)^{2^N}$. For j, k in $\{1, \dots, 2^N\}$, by evaluating the relation (3.6) in $v = v_k$ with the specific function $f(\cdot, s) = \delta_{v_j}(\cdot)$, we obtain that for all $k \neq j$,

$$\mathcal{B}_w f(v_k, s) = \sum_i \delta_1(v_k^i) \beta \delta_{v_j}(v_k - e_i) + \delta_0(v_k^i) \alpha_i(v_k, w) \delta_{v_j}(v_k + e_i) = Q^{jk},$$

and for $k = j$,

$$\mathcal{B}_w f(v_j, s) = \sum_i -\delta_1(v_j^i) \beta - \delta_0(v_j^i) \alpha_i(v_j, w) = Q^{jj}.$$

□

Remark 3.40. *As seen in Lemma 3.36, the $\nu_w^{v^k}$ can be found by computing the null space of the transition matrix of the process $(V_t)_{t \geq 0}$ alone.*

Remark 3.41. *Adding inhibitory neurons modelled as in Remark 3.18 will not fundamentally change the previous results. Indeed, with such a model, the dynamics of the weights of inhibitory neurons is the exact opposite to the excitatory ones, and as they live on \mathbb{Z}^- instead of \mathbb{Z}^+ , both dynamics are equivalent: they will diverge or be positive recurrent under the same conditions. However, the parameter range for which the chain is positive recurrent should be larger thanks to the inhibition that reduces the spiking rates.*

3.5 Simulations

As shown in Proposition 3.39, we can find the Laplace transform of π_w , the invariant measure of the fast process. However, inverting it analytically for a network of N neurons, with a large N , requires heavy computations: we have to deal with square matrices of size 2^N . Hence, we apply our results in a network of 2 neurons and then simulate a bigger network.

3.5.1 Application of our results

First, we give the main assumptions of this subsection on the functions p^+ , p^- and α_i as well as some biologically coherent parameters. Then, we apply our results to a network of two neurons.

Biologically coherent parameters:

We suppose that Assumption 3.30 holds with $g^+ = g^- \equiv 1$. Then, there exists $A_+, A_- \in [0, 1]$ and $\tau_-, \tau_+ \in \mathbb{R}_*^+$ such that

$$\forall (\hat{s}, \hat{w}) \in \mathbb{R}^+ \times \mathbb{N}^*, \quad p^+(\hat{s}, \hat{w}) = A_+ e^{-\frac{\hat{s}}{\tau_+}} \quad \text{and} \quad p^-(\hat{s}, \hat{w}) = A_- e^{-\frac{\hat{s}}{\tau_-}} \mathbb{1}_{[2, +\infty[}(\hat{w}).$$

Here, we also assume that for all i , α_i is a sigmoid function α of the synaptic current onto the neuron i and with parameters $\sigma \in \mathbb{R}^+$, $\theta \in \mathbb{R}$: for all $(v, w) \in \{0, 1\}^N \times E_2$,

$$\alpha_i(v, w) = \alpha \left(\sum_j w^{ij} v^j \right) \quad \text{where for all } x \in \mathbb{R}^+, \quad \alpha(x) = \frac{\alpha_M - \alpha_m}{1 + e^{-\sigma(x-\theta)}} + \alpha_m.$$

The network is fully described by parameters: $\varepsilon, A_+, A_-, \tau_-, \tau_+, \sigma, \theta, \beta, \alpha_m$ and α_M . The time of influence of a spike is of the order of 10 milliseconds whereas the time of a spike is of the order of 1 millisecond (\sim means *of the order of*), so we take $\beta \sim 0.1 \text{ ms}^{-1}$. The firing rates of the neurons are bounded by $\alpha_m \sim 0.01 \text{ ms}^{-1}$ and $\alpha_M \sim 1 \text{ ms}^{-1}$. The form of the sigmoid is given by σ determining the slope and θ giving the antecedent of $\frac{\alpha_M + \alpha_m}{2}$, see Figure 3.2b. Then, the plasticity parameters are in the following range: $\tau_{+/-} \in [1, 50]$, $A_{+/-} \in [0, 1]$, see Figure 3.2a for an example. Finally, the scaling parameter $\varepsilon \sim 0.01$ because experimentally, the STDP curve

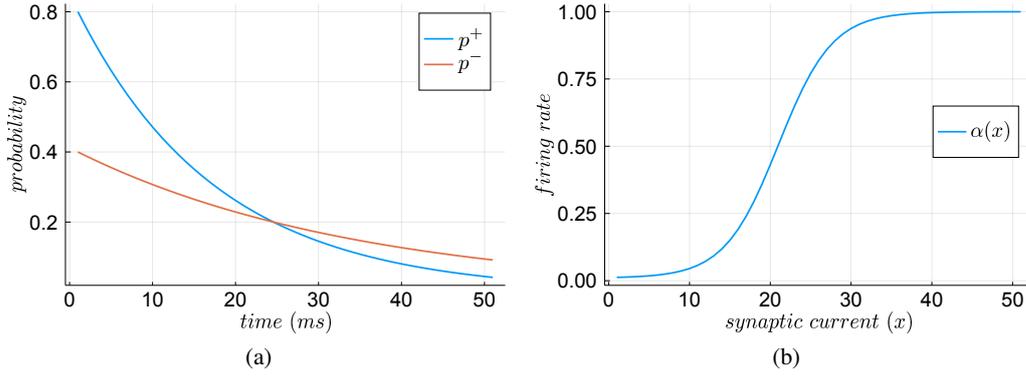


Figure 3.2: (3.2a): Functions $p^+(\cdot, \hat{w})$ and $p^-(\cdot, \hat{w})$ with parameters $A_+ = 0.8$, $A_- = 0.4$, $\tau_- = 2\tau_+ = 34$ ms. (3.2b): Sigmoid function α with parameters $\alpha_m = 0.01 \text{ ms}^{-1}$, $\alpha_M = 1 \text{ ms}^{-1}$, $\sigma = 0.3$, $\theta = 20$.

used here is obtained after around one hundred repetitions of spike pair, see Figure 1.5c for an approximation (STDP curve) of the experimental results obtained in [17].

An example with two neurons

We consider here the case $N = 2$. We assume that Assumption 3.9 holds. The functions φ_i are then given by (3.12) in the proof of Proposition 3.11. In this case, for all $i, j \neq i \in \{1, 2\}$, r_{ij}^+ (resp. r_{ij}^-) is the jump rate of w^{ij} to $w^{ij} + 1$ (resp. $w^{ij} - 1$). These jumps are the only one possible for the process $(\bar{W}_t)_{t \geq 0}$, see Remark 3.20. We compute the jumping rates as detailed just after Lemma 3.36.

We first show the validity of our timescale separation by comparing in Figure 3.3a the weight distributions obtained with a Monte Carlo method on

1. the trajectories of the process $(V_t^\varepsilon, S_t^\varepsilon, W_t^\varepsilon)_{t \geq 0}$,
2. the trajectories of the process $(\bar{W}_t)_{t \geq 0}$ (simulated using the rates obtained from the Laplace transform of π_w).

We observe quite a good match for simulations of 10^4 trajectories ($NMC = 10^4$), see Figure 3.3b.

We next study the divergence or not of the process. To do so, we find two sets of parameters: one set of parameters for which theoretically, from Proposition 3.24, the weight chain should be positive recurrent, and another one for which the chain should be transient. We describe these two sets of parameters now. Figures 3.4 and 3.5 are presented as two columns. The parameters for both columns are identical except that $A_+ = 0.3$ for the first column and $A_+ = 0.2$ for the second column. The other parameters are:

$$\alpha_m = 0.01 \text{ ms}^{-1}, \beta = 0.5 \text{ ms}^{-1}, \alpha_M = 1 \text{ ms}^{-1}, A_- = 0.6, \\ \tau_- = 34 \text{ ms}, \tau_+ = 17 \text{ ms}, \sigma = 0.3, \theta = 20 \text{ and } \varepsilon = 0.01. \quad (3.45)$$

We plot the functions p^+ and p^- in Figure 3.4. In order to study the divergence or not of the process, we plot the difference $r_{12}^+ - r_{12}^-$, see Figures (3.5a) and (3.5b). We surprisingly find that

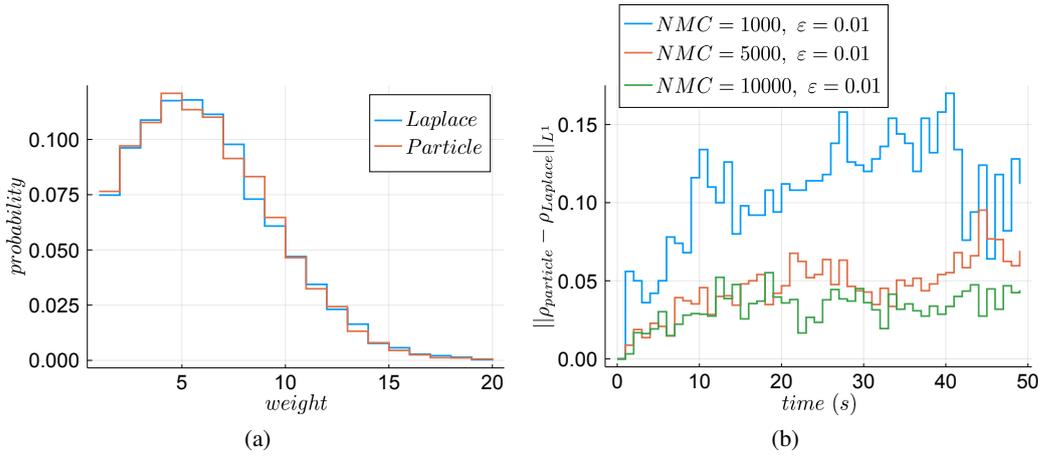


Figure 3.3: (3.3a) In blue: distribution of the process $(\overline{W}_t)_{t \geq 0}$ at time $t = 3s$ obtained from 10^4 trajectories simulated using the Laplace transform. In red: distribution of the process $(W_t^\varepsilon)_{t \geq 0}$ at time $t = 3s$ obtained from 10^4 trajectories of the process $(V_t^\varepsilon, S_t^\varepsilon, W_t^\varepsilon)_{t \geq 0}$. (3.3b) Distance in norm L_1 between the two distributions through time. Parameters are the same as Figure 3.2 with $\beta = \alpha_M = 1 \text{ ms}^{-1}$, $\alpha_m = 0.05 \text{ ms}^{-1}$ and $\varepsilon = 0.01$.



Figure 3.4: (3.4a), (3.4b): plots of the functions $p^+(\cdot, w)$ and $p^-(\cdot, w > 1)$. Parameters are given in (3.45) plus on the left, $A_+ = 0.3$ and on the right, $A_+ = 0.2$.

for some set of parameters, p^- has to be “significantly” above p^+ for the process $(\overline{W}_t)_{t \geq 0}$ to be positive recurrent, see Figure 3.4. Moreover, in the case of transience, the more probable way of weight divergence is when its symmetric weight is small enough. For example, if both weights are large enough, they both decrease until the moment their trajectories separate: one continues its descent and the other starts to increase. We illustrate this phenomenon on Figure 3.5g.

According to the simulations, we find a good agreement between the long time behaviour (positive recurrence or transience) predicted by our criterion given by Proposition 3.24. Indeed, Figure 3.5 shows the agreement between the theory (using the Laplace transform) and the initial neural network. In Figures (3.5a) and (3.5b), the heatmaps give the areas where the process is called back to 0 (negative areas) and the one where it diverges (positive areas). We observe that the maximal values of the function $w \mapsto (R^+(w) - R^-(w)) \cdot w$ are obtained for $w = (1, w^{12})$ and $w = (w^{21}, 1)$. Therefore, we plotted the function $w^{21} \mapsto (R^+(1, w^{21}) - R^-(1, w^{21})) \cdot (1, w^{21})$ for both set of parameters in Figures 3.5c and 3.5d. We deduce from Proposition 3.24 that the process should be positive recurrent (resp. transient) for the set of parameters corresponding to the right (resp. left) figures. This is effectively the case when we simulate the trajectories of the

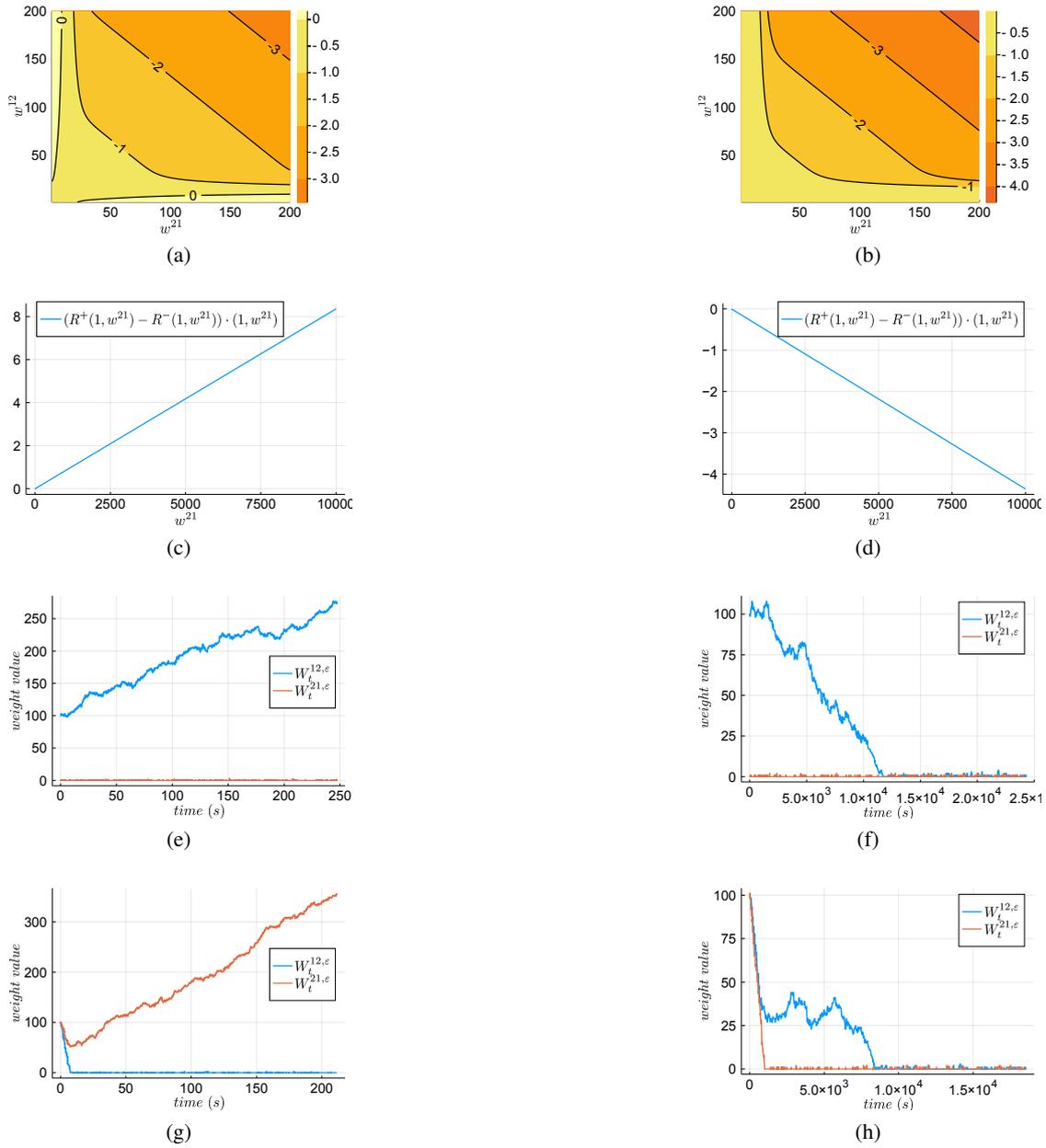


Figure 3.5: Parameters are given in (3.45) plus on the left, $A_+ = 0.3$ and on the right, $A_+ = 0.2$. (3.5a), (3.5b): heatmaps of the function $w = (w^{12}, w^{21}) \mapsto w^{12}(r_{12}^+(w) - r_{12}^-(w)) + w^{21}(r_{21}^+(w) - r_{21}^-(w))$. (3.5c), (3.5d): plots of the function $w^{21} \mapsto (R^+(1, w^{21}) - R^-(1, w^{21})) \cdot (1, w^{21})$. (3.5e), (3.5f): typical trajectories of $(W_t^\epsilon)_{t \geq 0}$ for $(W_0^{12, \epsilon}, W_0^{21, \epsilon}) = (100, 0)$. (3.5g), (3.5h): typical trajectory of $(W_t^\epsilon)_{t \geq 0}$ for $(W_0^{12, \epsilon}, W_0^{21, \epsilon}) = (100, 100)$.

process $(V_t^\epsilon, S_t^\epsilon, W_t^\epsilon)_{t \geq 0}$ with these two different sets of parameters, see Figures (3.5e) to (3.5h).

3.5.2 Beyond our results

In the previous subsection, we used a method, based on the Laplace transform, to check positive recurrence or transience for a network with 2 neurons. Here, we wonder whether the parameters

implying positive recurrence for a network with 2 neurons would imply positive recurrence for a network with more neurons. We tested with 100 neurons and realized that the network seems to prevent the weight escape when the number of neuron increases. If it is the case, it would mean that the computations in order to ensure positive recurrence could be done on a simple 2 neuron network, which is really an easy task.

We use the previous parameters (3.45) plus different values of A_+ , in order to see whether the weights diverge or not in a network of N equals 100 neurons. We note that contrary to Figure 3.5, for $A_+ = 0.3$ the mean of the weights stays bounded and converges to a value near 0.8, see Figure 3.6a. In order to test divergence, we tried different initial weight matrices. The time to reach the limit value is the longest in the case where there is one line of big weights and the others are in state one, see Figure 3.6a. Changing only the parameter A_+ , we observe on Figure 3.6 that the weights seem to diverge for parameters $A_+ \geq 0.6$. Hence, in the case of $N = 100$ neurons, divergence requires a stronger potentiation probability than in the case of $N = 2$ neurons. This is counter intuitive as we would expect the weights to be more pushed by the other neurons when the number of neurons becomes larger, thus diverging more easily.

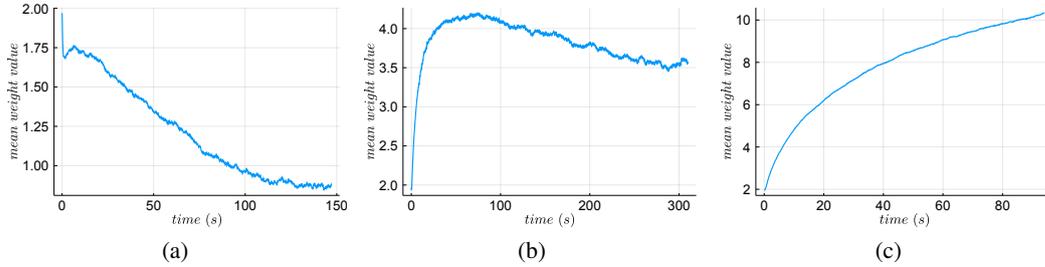


Figure 3.6: Mean of the weights through time. The parameters for all figures are identical except that $A_+ = 0.3$ in (3.6a), $A_+ = 0.5$ in (3.6b) and $A_+ = 0.6$ in (3.6c). We stopped the simulations after $2 \cdot 10^6$ jumps of the weights in (3.6a) and (3.6c) and after $6 \cdot 10^6$ jumps to (3.6b). The other parameters are: $\alpha_m = 0.01ms^{-1}$, $\beta = 0.5ms^{-1}$, $\alpha_M = 1ms^{-1}$, $A_- = 0.6$, $\tau_- = 34ms$, $\tau_+ = 17ms$, $\sigma = 0.3$, $\theta = 20$ and $\varepsilon = 0.01$.

Remark 3.42. *We have chosen 100 neurons for plotting constraints. It is easy to simulate the process with thousands of them.*

3.6 Discussion

Results

Based on a well known neural network model, we added plasticity in order to get insight on the combined neurons-weights dynamics. We could analyse plasticity on the slow timescale of the weight dynamics compared to the neuron dynamics, thus providing a simplified model. The latter gives the weight dynamics under the stationary distribution of the fast process and is a continuous time Markov jump process on the infinite state space of weights with non homogeneous in space jump rates. Such processes are hard to deal with and current results are given in [96]. Moreover, even if we proved the existence and uniqueness of the invariant measure of the fast process, we were not able to express it explicitly. Nevertheless, we managed to study the long time behaviour

of the limit dynamics of the weights. First, we give a general condition on positive recurrence and transience of the limit chain. This condition is obtained by comparing (using a coupling method) the fast process to simpler processes using the bounds on the neuron jump rates. Second, we overcome the difficulty of the unknown invariant measure using its Laplace transform instead. Indeed, we give a method to compute it and more importantly, we give a simple way to use it in order to determine the jump rates of the limit dynamics of the weights. We used this method in an example with 2 neurons which works pretty well even for not too big difference of timescale between the neural network and the weights ($\varepsilon = 0.01$). The problem is nevertheless quickly harder when the number of neurons becomes larger. Indeed, the computations of the Laplace transform consists in inverting a 2^N square matrix for a given weight matrix and as soon as the latter change, we need to invert a new 2^N square matrix again. Such computations risk to become too heavy to be finished in reasonable time delays. A hope would be that the study in the $N = 2$ neurons case be sufficient for larger N . That seems the case according to our simulations, but we did not prove it yet. Thus, our study gives more insights in dealing with the long time behaviour of the weights in STDP models, especially in recurrent excitatory neural networks.

Limitations of our model and future work

We are aware that the individual neuron model considered in this work is far from the biological reality. It is simple in order to perform a rigorous and mathematical study of the plasticity. Some questions raise when we try to understand it from a biological point of view. For instance, what does β represent? In our model, β is linked to two things: the time over which one neuron influences the others and, because the neuron cannot spike while being in state 1, $\frac{1}{\beta}$ is the mean duration of a spike. Thus, we should add another variable in order to take into account these two phenomena separately. Moreover, the binary model of the neurons' potential is a drastic simplification. Indeed, neurons are usually described through their membrane potential. Replacing the binary potentials by continuous ones would make the mathematical analysis harder. In particular, the study of the invariant measure (if it exists and is unique) of the fast process is complex. Finally, testing (with simulations) whether the network can reproduce some well known biological phenomena has not been done in this chapter. For example, the learning capacity of the network, the dependence on the postsynaptic potential of plasticity, the presence of bidirectional and unidirectional connectivity could be tested at least numerically.

Conclusion

We propose a new view on STDP models. In contrast with usual tiny deterministic weight jumps, in our model, the weights have some weak probability to make a “big” jump. Thus, instead of being continuous, our weights are discrete. Associated to the time since the last spikes and the neural network state, we get a Markov process. We simplified it thanks to a separation of timescale and found simple conditions of positive recurrence. This work opens up a new framework to study plasticity which we hope will give rise to more mathematical results on plasticity in the future.

Chapter 4

Mean Field Analysis of the Stochastic STDP model

Abstract

The large number of elements (ionic channels, synapses, neurons, etc.) makes any attempt to obtain an exact model of the brain or even parts of it doomed to failure. Here, using a mean-field approach on a neural network model with plastic interaction, we take into account the large number of neurons in each cortical zone to propose a local simplification of the activity of a *typical* neuron. We study a well known phenomenon called Spike-Timing-Dependent Plasticity (STDP) which is implemented within a probabilistic Wilson-Cowan neural network model. The neural network is then described by N triplets composed of the neuron potential, the time since its last spike and its N incoming synaptic weights. Based on this initial description, we describe the neural network with *new* variables that ease performing a mean field approximation. In the asymptotic of a large neuron number, the *typical* neuron is described by a stochastic process with three components: the first two ones are the same as in the initial description and the last one is the **probability distribution** of the state of the presynaptic neurons (triplet detailed just above) replacing the N weights by the weight incoming onto the postsynaptic neuron concerned. Therefore, we conjecture the limit system which is still quite complex. To the best of our knowledge, this is the first time that a mean field analysis is performed on a plastic neural network model with STDP.

4.1 Introduction

During the last few decades, synaptic plasticity has been widely studied and its interplay with neural networks is still full of mysteries with many fascinating questions on the network structure formation, its stability and its function, see [100, 106, 24]. On the modelling point of view, one of the most studied phenomena of plasticity is the so-called Spike-Timing-Dependent Plasticity (STDP), see Section 1.1.2. The implementation of STDP in a neural network model has already been done in many previous studies, see for example [94, 95, 59, 107]. However, rigorous mathematical analysis of such complex interaction has not yet been done. One natural mathematical method to use is the well known mean field approximation which has already been used to analyse some interacting neural networks without plasticity, see for instance [35, 39, 44, 40, 30]. In this chapter, we present the mean field limit of a probabilistic Wilson-Cowan neural network model in which neurons interact through a stochastic STDP rule.

The first difficulty is to find what is the neural network description suitable for such a mean field analysis. In recent studies, a short term memory model has been analysed with a mean field method in [53]. In their model, the neural interaction is homogeneous which is not the case in our model. This is the main reason of the study difficulty: the heterogeneity of the interaction combined with their dynamics. Indeed, usually McKean-Vlasov limit equations are naturally derived in networks where the interaction is assumed to be homogeneous, see [123]. However, when it is not the case, the first difficulty is to determine this limit. The second difficulty is to study this limit and the third one is to show the convergence to this limit. In this work, we do not deal with these two difficulties and focus only on the first one by assuming the convergence in order to conjecture the limit dynamics that matches with our analytical computations. We perform numerical simulations to illustrate our results.

This chapter is organised as follows. In Section 4.2, we first present the microscopic model (initial neural network description) before presenting the first hints on the macroscopic model. In particular, we introduce *new* variables describing the neural network from which we perform a mean field analysis in Section 4.3. Hence, the latter is devoted to the derivation of the limit equation from the analysis of the empirical measure of the *new* neural network description. In the first part, we study the simpler case of interaction without plasticity before adding the plasticity rule in a second part. Thereby, a McKean-Vlasov equation is conjectured on a typical neuron which is composed of: the neuron state, the time from its last spike and the distribution of the triplets composed of the presynaptic neuron states, their time since their last spike and their synaptic weights onto the typical neuron. We finally perform numerical simulations on these last equations in Section 4.4.

4.2 Presentation of the model

The model is a special case of the one studied in Chapter 3. The main *restriction* here is the special form of the spiking rate: in this chapter, we assume that $\alpha_i(V_t^N, W_t^N) = \alpha(I_t^{i,N})$ where $I_t^{i,N} = \frac{1}{N} \sum_j W_t^{ij} V_t^j$ is the global incoming synaptic current. We start with a presentation of the microscopic model involving Poisson measures (see Definition A.5 in the appendix). Then, we introduce the corresponding macroscopic model.

4.2.1 The microscopic model

We recall the main features of the model. We study a network of N binary neurons, $V_t^{i,N} \in \{0, 1\}$, all-to-all connected. The interactions are characterized by the synaptic weights matrix $W_t^N \in \mathbb{Z}^{N^2}$. Note that we no more assume in this chapter that the diagonal elements $W_t^{ii,N}$ are null. In particular, we assume that the dynamics of the diagonal elements is similar to the one of $W_t^{ij,N}$ for $i \neq j$. This last assumption simplifies the mean field analysis. In any case, the macroscopic behaviour of the network is not affected by the choice of one weight when the number N of presynaptic neurons tends to infinity. The weight $\frac{1}{N} W_t^{ij,N}$ represents the effect on the neuron i of a spike emitted by the neuron j at time t . At time t , we denote $S_t^{i,N} \in \mathbb{R}^+$ the time spent since the last spike of the neuron i . The process we are interested in is a Piecewise Deterministic Markov Process (PDMP). The deterministic part between the jumps is very simple, the variables $(S_t^{i,N})_{1 \leq i \leq N}$ increase linearly with time: $dS_t^{i,N} = dt$. The **jump** of the neuron i from 0 to 1 is called a **spike** of the neuron i and occurs at rate $\alpha(I_t^{i,N})$ where

$$I_t^{i,N} \stackrel{\text{def.}}{=} \frac{1}{N} \sum_j W_t^{ij,N} V_t^{j,N} \quad (4.1)$$

and $\alpha : \mathbb{R} \rightarrow [\alpha_m, \alpha_M]$ is a strictly positive and bounded function. At the time t of a spike of neuron i , $S_t^{i,N}$ is reset to 0 and the weights jump as follows:

- $W_{t^-}^{ij} \rightarrow W_{t^-}^{ij} + 1$ with probability $p^+(S_{t^-}^j, W_{t^-}^{ij})$,
- $W_{t^-}^{ji} \rightarrow W_{t^-}^{ji} - 1$ with probability $p^-(S_{t^-}^j, W_{t^-}^{ji})$,

where p^+ and p^- are functions from $\mathbb{R}^+ \times \mathbb{Z}$ to $[0, 1]$. Finally, the potentials $V_t^{i,N}$ of the neurons jump from 1 to 0 at constant rate $\beta > 0$; at these jumping times, neither the synaptic weights nor the times $S_t^{i,N}$ are modified. We also call this jump: return to the resting potential. Without loss of generality and as in Chapter 3, we assume that $0 < \alpha_m < \beta < \alpha_M < \infty$.

In order to represent the jumps of the dynamics, we introduce $(\zeta^i(du, dz, d\theta))_{i=1, \dots, N}$ a family of N independent Poisson measures on $\mathbb{R}_+^2 \times [0, 1]^{2N}$ with intensity $dudzd\theta$ where

$$d\theta \stackrel{\text{def.}}{=} \prod_{k=1}^N d\theta_k^+ \prod_{k=1}^N d\theta_k^-.$$

Thus, this model can be written in the following form: for any $i, j \in \llbracket 1, N \rrbracket$,

$$\begin{aligned} V_t^{i,N} &= V_0^{i,N} + \int_0^t \int_{E(N)} \mathbb{1}_{\{z \leq \alpha(I_{u^-}^{i,N}) \mathbb{1}_{\{V_{u^-}^{i,N}=0\}}\}} - \mathbb{1}_{\{z \leq \beta \mathbb{1}_{\{V_{u^-}^{i,N}=1\}}\}} \zeta^i(du, dz, d\theta) \\ S_t^{i,N} &= S_0^{i,N} + t - \int_0^t \int_{E(N)} S_{u^-}^{i,N} \mathbb{1}_{\{z \leq \alpha(I_{u^-}^{i,N}) \mathbb{1}_{\{V_{u^-}^{i,N}=0\}}\}} \zeta^i(du, dz, d\theta) \\ W_t^{ij,N} &= W_0^{ij,N} + \int_0^t \int_{E(N)} \mathbb{1}_{\{\theta_j^+ \leq p^+(S_{u^-}^{j,N}, W_{u^-}^{ij,N})\}} \mathbb{1}_{\{z \leq \alpha(I_{u^-}^{i,N}) \mathbb{1}_{\{V_{u^-}^{i,N}=0\}}\}} \zeta^i(du, dz, d\theta) \\ &\quad - \int_0^t \int_{E(N)} \mathbb{1}_{\{\theta_i^- \leq p^-(S_{u^-}^{i,N}, W_{u^-}^{ij,N})\}} \mathbb{1}_{\{z \leq \alpha(I_{u^-}^{j,N}) \mathbb{1}_{\{V_{u^-}^{j,N}=0\}}\}} \zeta^j(du, dz, d\theta), \end{aligned} \quad (4.2)$$

where we use the notation $E^{(N)} \stackrel{\text{def.}}{=} \mathbb{R}^+ \times [0, 1]^{2N}$.

We can now detail the probability space we use thereafter. We denote $Y_t^N = (Y_t^{1,N}, \dots, Y_t^{N,N})$ such that for all i , $Y_t^{i,N} = (V_t^{i,N}, S_t^{i,N}, (W_t^{ij,N})_{1 \leq j \leq N})$. In the following, we consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with the filtration

$$\mathcal{F}_t^N \stackrel{\text{def.}}{=} \sigma(\{Y_0^N, \zeta^1([0, a_1] \times A^1), \dots, \zeta^N([0, a_N] \times A^N) : \forall i a_i \in [0, t], A^i \in \mathcal{A}_N\})$$

where $\mathcal{A}_N = \mathcal{B}(E^{(N)})$.

We end the description of the microscopic model by giving the assumptions on the initial conditions.

Assumption 4.1. *The triplets $(V_0^{1,N}, S_0^{1,N}, (W_0^{1j,N})_{1 \leq j \leq N}), \dots, (V_0^{N,N}, S_0^{N,N}, (W_0^{Nj,N})_{1 \leq j \leq N})$ are i.i.d. at time $t = 0$ with law κ_0^N . Moreover, the law ρ_0 of $S_0^{i,N}$ is absolutely continuous with respect to the Lebesgue measure λ . The density is assumed to be bounded.*

Under this assumption, the neural network possesses two important properties that are extensively used throughout this chapter. First, the times from the last spike are almost surely distinct.

Lemma 4.2. *Grant Assumption 4.1. Consider the process Y_t^N solution of (4.2). Then, for any $t \geq 0$ and for all $i \neq j \in \llbracket 1, N \rrbracket$, we have almost surely $S_t^{i,N} \neq S_t^{j,N}$.*

Proof. By assumption, for all i , the law of $S_0^{i,N}$ admits a density and hence, the $(S_0^{i,N})_{0 \leq i \leq N}$ are almost surely distinct. Between the jumps of $(V_t^{i,N})_{0 \leq i \leq N}$ from 0 to 1, we have $d(S_t^{i,N} - S_t^{j,N}) = 0$. Finally, the probability that two different neurons spike at the same time is zero. \square

Second, their laws stay absolutely continuous with respect to the Lebesgue measure.

Lemma 4.3. *We denote $\|\rho_0\|_\infty$ the upper bound of ρ_0 and we assume that Assumption 4.1 holds. Then, for any $t \geq 0$, the law ρ_t of $S_t^{i,N}$ also admits a density. In particular, for any $A \in \mathcal{B}(\mathbb{R}^+)$,*

$$\rho_t(A) \leq (\alpha_M + \|\rho_0\|_\infty)\lambda(A).$$

Proof. We have for all $A \in \mathcal{B}(\mathbb{R}^+)$, $t > 0$,

$$\mathbb{P}(S_t^{i,N} \in A) = \mathbb{P}(S_t^{i,N} \in A \cap [0, t]) + \mathbb{P}(S_t^{i,N} \in A \cap [t, +\infty[).$$

First, denoting $A - t = \{x \in \mathbb{R}^+, x + t \in A\}$, one has

$$\left\{ S_t^{i,N} \in A \cap [t, +\infty[\right\} \subset \left\{ S_0^{i,N} \in A - t \right\}.$$

Thus,

$$\mathbb{P}\left(S_t^{i,N} \in A \cap [t, +\infty[\right) \leq \mathbb{P}\left(S_0^{i,N} \in A - t\right) \leq \|\rho_0\|_\infty \lambda(A - t) = \|\rho_0\|_\infty \lambda(A).$$

Second, the event $\left\{ S_t^{i,N} \in A \cap [0, t] \right\}$ means that the last spike of the neuron i occurred at a time in $t - A$. The probability of this event is less than the probability that there is at least one spike of the neuron i in $t - A$. So, one has,

$$\begin{aligned} \mathbb{P}(S_t^{i,N} \in A \cap [0, t]) &\leq \mathbb{E} \left[1 - \exp \left(- \int_{t-A}^t \mathbb{1}_{\{V_u^{i,N} = 0\}} \alpha(I_u^{i,N}) du \right) \right] \leq \mathbb{E} \left[\int_{t-A}^t \alpha(I_u^{i,N}) du \right] \\ &\leq \alpha_M \lambda(t - A) = \alpha_M \lambda(A). \end{aligned}$$

\square

We end this section with notations used for the jumping times and for the probability measure spaces.

Notation 4.4. We denote by

$$T = \{T_1, T_2, \dots\} \quad (4.3)$$

the sequence of the jumping times of $(\mu_t^N)_{t \geq 0}$. They correspond to the instants at which one of the $V_t^{i,N}$ changes (from 0 to 1 or from 1 to 0).

We denote by

$$\tau = \{\tau_1, \tau_2, \dots\} \subset T \quad (4.4)$$

the sequence of the spiking times, that is the instants at which one of the $V_t^{i,N}$ jumps from 0 to 1.

Notation 4.5. Let (G, \mathcal{G}) be a measurable space. We denote by $\mathcal{P}(G)$ the set of probability measures on this space. We equip this space with the total variation distance:

$$\forall \mu, \nu \in \mathcal{P}(G), \quad d_{TV}(\mu, \nu) \stackrel{def.}{=} \sup\{|\mu(A) - \nu(A)|, A \in \mathcal{G}\}.$$

We denote by $\mathcal{B}(\mathcal{P}(G))$ the collection of Borel sets of the metric space $(\mathcal{P}(G), d_{TV})$.

Moreover, we call “set of empirical measures of the order N on (G, \mathcal{G}) ” the set of atomic probability measures on G with N distinct atoms of weight $\frac{1}{N}$. We denote by $\mathcal{P}_N(G)$ this set.

4.2.2 Towards the macroscopic model

In the following, with a slight abuse of notation, we also call *neuron* the process describing the neuron. Starting from the microscopic model, we wish to describe the macroscopic model obtained when the number of neurons tends to infinity. To do so, we are going to identify the neurons $X_t^{1,N}, \dots, X_t^{N,N}$ allowing us to obtain *effectively* the limit dynamics. The neuron i , $X_t^{i,N}$, includes at least the potential of the neuron $V_t^{i,N}$. We wish that the family of neurons $X_t^{1,N}, \dots, X_t^{N,N}$ has the following property: $X_t^N = (X_t^{i,N})_{1 \leq i \leq N}$ is a Markov process. This requires that the incoming synaptic current $I_t^{i,N}$ on the neuron i , defined by (4.1), can be computed easily from the state of the neuron i . Moreover, the description of the dynamics of $I_t^{i,N}$ requires the knowledge of the distribution of the couples $(V_t^{j,N}, S_t^{j,N})_{1 \leq j \leq N}$. Thus, for each neuron i , we consider the empirical distribution $\xi_t^{i,N}$ of the triplets $(V_t^{j,N}, S_t^{j,N}, W_t^{ij,N})_{1 \leq j \leq N}$.

Definition 4.6. The empirical measure

$$\forall i = 1, \dots, N, \quad \xi_t^{i,N} \stackrel{def.}{=} \frac{1}{N} \sum_j \delta_{(V_t^{j,N}, S_t^{j,N}, W_t^{ij,N})}.$$

is a random variable on the space of probability measures on

$$E_m \stackrel{def.}{=} \{0, 1\} \times \mathbb{R}^+ \times \mathbb{Z}. \quad (4.5)$$

In particular, with Notation 4.5, we have that for all $\omega \in \Omega$, $\xi_t^{i,N}(\omega) \in \mathcal{P}_N(E_m)$.

Defining the function $I : \mathcal{P}(E_m) \rightarrow \mathbb{R}$ by

$$\forall \xi \in \mathcal{P}(E_m), \quad I(\xi) \stackrel{def.}{=} \mathbb{E}^\xi(WV) = \int_{E_m} wv \xi(dv, ds, dw), \quad (4.6)$$

we note that the knowledge of $\xi_t^{i,N}$ is enough to obtain the incoming synaptic current $I_t^{i,N} = I(\xi_t^{i,N})$.

Remark 4.7. We want to emphasize the following point: the knowledge of $(\xi_t^{j,N})_{1 \leq j \leq N}$ is sufficient to evaluate the spiking rates of the neurons. However, the measures $(\xi_t^{j,N})_{1 \leq j \leq N}$ are empirical distributions and do not contain the labels of the presynaptic neurons. Precisely, as soon as one neuron, say neuron i_0 , has a jump of $V_t^{i_0,N}$, we have to propagate this jump (and its potential consequences, such that a reset of $S_t^{i_0,N}$ and changes in the synaptic weights) to every $(\xi_t^{j,N})$. For any j , we know that one atom of $\xi_t^{j,N}$ corresponds to the neuron i_0 but we have to find it. Thanks to Lemma 4.2, we recognize this atom with the value $S_t^{i_0,N}$.

Therefore, we can now give the new description of the neural network.

Definition 4.8. The neurons are described by the following triplets

$$\forall i = 1, \dots, N, \quad X_t^{i,N} \stackrel{\text{def.}}{=} (V_t^{i,N}, S_t^{i,N}, \xi_t^{i,N}).$$

We show in Proposition 4.12 that $X_t^N \stackrel{\text{def.}}{=} (X_t^{1,N}, \dots, X_t^{N,N})$ is a Markov process.

Remark 4.9. By Assumption 4.1, the measures $\xi_0^{1,N}, \dots, \xi_0^{N,N}$ have the same law. Also, this assumption added to the fact that the variables $V_t^{i,N}, S_t^{i,N}, W_t^{ij,N}$ have for all i, j , the same dynamics, implies that for all i and $t \geq 0$, the $\xi_t^{i,N}$ are equal in law. We conclude this remark with the crucial following point: using Assumption 4.1, we deduce that the law of X_t^N is exchangeable (or symmetric), which means that for any permutation σ of $\{1, \dots, N\}$, $X_t^{\sigma,N} = (X_t^{\sigma(1),N}, \dots, X_t^{\sigma(N),N})$ has the same law as X_t^N .

Under Assumption 4.1, we obtain $X_0^{1,N}, \dots, X_0^{N,N}$ from N independent draws of the κ_0^N law. We already have the dynamics of $(V_t^{i,N})_{1 \leq i \leq N}$ and $(S_t^{i,N})_{1 \leq i \leq N}$. In order to describe the dynamics of the neural network composed of $X_t^{1,N}, \dots, X_t^{N,N}$, the only dynamics left is the one of the $(\xi_t^{i,N})_{1 \leq i \leq N}$: for all $i = 1, \dots, N$,

$$\begin{aligned} \xi_t^{i,N} = & \xi_0^{i,N} + \int_0^t \int_{E^{(N)}} \frac{1}{N} \sum_{(\tilde{V}, \tilde{S}, \tilde{W}) \in \text{supp}(\xi_u^{i,N})} \left\{ \right. \\ & \left(\mathbb{1}_{\{\tilde{S}=S_u^{i,N}\}} \left[\delta_{(1,0,\tilde{W}+1)} \mathbb{1}_{\{\theta_i^+ \leq p^+(S_u^{i,N}, \tilde{W})\}} - \mathbb{1}_{\{\theta_i^- \leq p^-(S_u^{i,N}, \tilde{W})\}} \right] - \delta_{(0,\tilde{S},\tilde{W})} \right) \\ & + \sum_{j \neq i} \mathbb{1}_{\{\tilde{S}=S_u^{j,N}\}} \left[\delta_{(\tilde{V}, \tilde{S}, \tilde{W}+1)} \mathbb{1}_{\{\theta_j^+ \leq p^+(S_u^{j,N}, \tilde{W})\}} - \delta_{(\tilde{V}, \tilde{S}, \tilde{W})} \right] \mathbb{1}_{\{z \leq \alpha(I_u^{i,N}) \mathbb{1}_{\{V_u^{i,N}=0\}}\}} \zeta^i(du, dz, d\theta) \\ & + \sum_{j \neq i} \mathbb{1}_{\{\tilde{S}=S_u^{j,N}\}} \left[\delta_{(1,0,\tilde{W}-1)} \mathbb{1}_{\{\theta_i^- \leq p^-(S_u^{i,N}, \tilde{W})\}} - \delta_{(0,\tilde{S},\tilde{W})} \right] \mathbb{1}_{\{z \leq \alpha(I_u^{j,N}) \mathbb{1}_{\{V_u^{j,N}=0\}}\}} \zeta^j(du, dz, d\theta) \\ & \left. + \sum_j \mathbb{1}_{\{\tilde{S}=S_u^{j,N}\}} \left[\delta_{(0,\tilde{S},\tilde{W})} - \delta_{(1,\tilde{S},\tilde{W})} \right] \mathbb{1}_{\{z \leq \beta \mathbb{1}_{\{V_u^{j,N}=1\}}\}} \zeta^j(du, dz, d\theta) \right\}. \end{aligned}$$

As explained in Remark 4.7, as soon as one neuron spikes, say neuron i , we recognize it in the other $\xi_t^{j,N}$ with the value of $S_t^{i,N}$.

4.3 Study of the empirical measure and its limit

Now that we have defined the new variables describing the neural network, we analyse its asymptotic dynamics when the number of neurons tends to infinity. To do so, we study the empirical

measure associated to the *new* neural network description which is defined on a space that does not depend on N . We first show that, as a time dependent variable, this probability measure is a Markov process. Then, we give the main result of this chapter: our conjecture on the limit dynamics of a *typical* neuron. We end by performing computations that explain why our conjecture is reasonable, first in the case without plasticity and then with plasticity.

4.3.1 The empirical measure

Definition 4.10. *The empirical measure*

$$\mu_t^N \stackrel{\text{def.}}{=} \frac{1}{N} \sum_i \delta_{X_t^{i,N}} = \frac{1}{N} \sum_i \delta_{(V_t^{i,N}, S_t^{i,N}, \xi_t^{i,N})} \quad (4.7)$$

is a random probability measure over the space of probability measures on

$$E \stackrel{\text{def.}}{=} \{0, 1\} \times \mathbb{R}^+ \times \mathcal{P}(E_m). \quad (4.8)$$

Hence, for all $\omega \in \Omega$, $\mu_t^N(\omega) \in \mathcal{P}_N(E) \subset \mathcal{P}(E)$.

In the following, we keep in mind the randomness of μ_t^N and $(\xi_t^{i,N})_{1 \leq i \leq N}$ and we alleviate the notations by referring to the outcomes $\omega \in \Omega$ only when it helps understanding. From the definition of $(\mu_t^N)_{t \geq 0}$ and the ones of $(\xi_t^{1,N})_{t \geq 0}, \dots, (\xi_t^{N,N})_{t \geq 0}$, we note that they have the same joint law on V and S ,

Remark 4.11. For all $i \in \llbracket 1, N \rrbracket$ and $t \geq 0$, we have by Definitions 4.6 and 4.10, for all $(A, B) \in \mathcal{B}(\{0, 1\} \times \mathbb{R}^+)$,

$$\mu_t^N(A, B, \mathcal{P}(E_m)) = \xi_t^{i,N}(A, B, \mathbb{Z}).$$

In what follows, we denote this property by: $\mu_t^N(\cdot, \cdot, \mathcal{P}(E_m)) = \xi_t^{i,N}(\cdot, \cdot, \mathbb{Z})$.

Proposition 4.12. *The random process $(\mu_t^N)_{t \geq 0}$ is a Markov process on $\mathcal{P}_N(E)$.*

Proof. The process $(\mu_t^N)_{t \geq 0}$ is a PDMP on $\mathcal{P}_N(E)$ which is described below.

We first give the law of the jumping times $T = \{T_1, T_2, \dots\}$ of $(\mu_t^N)_{t \geq 0}$. They correspond to the times at which one of the $V^{i,N}$ changes (see Notation 4.4). Then, we give the flow between these times and finally, we give the transition kernel.

Law of the components of T and τ

We denote by $\lambda^J : \mathcal{P}_N(E) \rightarrow [\alpha_m N, \alpha_M N]$ the jump rate of $(\mu_t^N)_{t \geq 0}$

$$\forall t \geq 0 \quad \lambda^J(\mu_t^N) \stackrel{\text{def.}}{=} \sum_i \alpha(I_t^{i,N}) \mathbb{1}_{\{V_t^{i,N}=0\}} + \beta \mathbb{1}_{\{V_t^{i,N}=1\}}.$$

Note that $\lambda^J(\mu_t^N)$ only depends on V and W and is constant between the jumps.

Thus, the time $T_{k+1} - T_k$ between two jumps is exponentially distributed with parameter $\lambda^J(\mu_{T_k}^N)$. At time T_{k+1} , we denote j_{k+1} the label of the neuron such that

$$V_{T_k}^{j_{k+1},N} \neq V_{T_{k+1}}^{j_{k+1},N}.$$

We have

$$\mathbb{P}(j_{k+1} = i) = \frac{\alpha(I_{T_k}^{i,N}) \mathbb{1}_{\{V_{T_k}^{i,N}=0\}} + \beta \mathbb{1}_{\{V_{T_k}^{i,N}=1\}}}{\lambda^J(\mu_{T_k}^N)}.$$

If $V_{T_k}^{j_k,N} = 0$, the time T_{k+1} is a spiking time and belongs to τ .

To alleviate the notations, if a jumping time T_{k+1} is a spiking time, say τ_ℓ , we simply denote by q_ℓ the label of the **spiking neuron**.

Flow between jumps

Between the jumps of $(\mu_t^N)_{t \geq 0}$, the flow is very simple and is given by

$$\forall T_k \leq t < T_{k+1}, \quad \mu_t^N = \frac{1}{N} \sum_i \delta \left(V_{T_k}^{i,N}, S_{T_k}^{i,N} + t - T_k, \frac{1}{N} \sum_{(\tilde{V}, \tilde{S}, \tilde{W}) \in \text{supp}(\xi_{T_k}^{i,N})} \delta_{(\tilde{V}, \tilde{S} + t - T_k, \tilde{W})} \right).$$

Transition kernel

For all $t \in \{T_1, T_2, \dots\}$, we first detail the jumps of the variables in the support of the empirical measures $(\xi_t^{j,N})_{1 \leq j \leq N}$ and then give the jumps of $(\mu_t^N)_{t \geq 0}$. As noted in Remark 4.7, attention is drawn to the fact that we cannot label the support of the $(\xi_t^{j,N})_{1 \leq j \leq N}$ with the same labels as the support of μ_t^N . Therefore, we have to find a way to access the right triplets to change when μ_t^N jumps. For instance, when the neuron i jumps, we have to change all the $\delta_{(V_t^{i,N}, \dots)}$ by $\delta_{(1 - V_t^{i,N}, \dots)}$ in all the empirical measures $(\xi_t^{j,N})_j$. In order to access to the triplet $(V_t^{i,N}, S_t^{i,N}, W_t^{i,N}) \in \text{supp}(\xi_t^{j,N})$, we go through the (V, S, W) in the support of $\xi_t^{j,N}$ and keep the only one (see Lemma 4.2) with $S = S_t^{i,N}$.

Each time a neuron spikes, we draw $2N$ *i.i.d.* random variables with uniform law on $[0, 1]$. At time $t = \tau_k$, we denote them by $(U_+^{l,k})_{1 \leq l \leq N}$ and $(U_-^{l,k})_{1 \leq l \leq N}$. At this time, we have the following jumps of the empirical measures.

- Let $(\tilde{V}, \tilde{S}, \tilde{W}) \in \text{supp}(\xi_{\tau_k^-}^{q_k,N})$. If $\tilde{S} = S_{\tau_k^-}^{q_k,N}$, by Lemma 4.2 $\tilde{V} = 0$ and then

$$(\tilde{V}, \tilde{S}, \tilde{W}) \rightarrow \left(1, 0, \tilde{W} + \mathbb{1}_{\{U_+^{q_k,k} \leq p^+(S_{\tau_k^-}^{q_k,N}, \tilde{W})\}} - \mathbb{1}_{\{U_-^{q_k,k} \leq p^-(S_{\tau_k^-}^{q_k,N}, \tilde{W})\}} \right).$$

If $\tilde{S} = S_{\tau_k^-}^{l,N}$, $l \neq q_k$,

$$(\tilde{V}, \tilde{S}, \tilde{W}) \rightarrow \left(\tilde{V}, \tilde{S}, \tilde{W} + \mathbb{1}_{\{U_+^{l,k} \leq p^+(S_{\tau_k^-}^{l,N}, \tilde{W})\}} \right).$$

- For all $j \neq q_k$, for all $(\tilde{V}, \tilde{S}, \tilde{W}) \in \text{supp}(\xi_{\tau_k^-}^{j,N})$, if $\tilde{S} = S_{\tau_k^-}^{q_k,N}$, by Lemma 4.2 $\tilde{V} = 0$ and then

$$(\tilde{V}, \tilde{S}, \tilde{W}) \rightarrow \left(1, 0, \tilde{W} - \mathbb{1}_{\{U_-^{j,k} \leq p^-(S_{\tau_k^-}^{j,N}, \tilde{W})\}} \right).$$

Now, let $t \in \{T_1, T_2, \dots\} \setminus \tau$ and assume that the neuron returning to its resting potential is the neuron j_0 . For all j , let $(\tilde{V}, \tilde{S}, \tilde{W}) \in \text{supp}(\xi_{t^-}^{j,N})$ such that $\tilde{S} = S_{t^-}^{j_0,N}$, by Lemma 4.2 $\tilde{V} = 1$ and then

$$(\tilde{V}, \tilde{S}, \tilde{W}) \rightarrow (0, \tilde{S}, \tilde{W}).$$

We now give the transition kernel of $(\mu_t^N)_{t \geq 0}$. For any $t \in \{T_1, T_2, \dots\}$,

- if $t \notin \{\tau_1, \tau_2, \dots\}$, with probability $\frac{\beta}{\sum_i \mathbb{1}_{\{V_{t^-}^i = 1\}}} \mathbb{1}_{\{V_{t^-}^{j_0,N} = 1\}}$ the measure leaves the state $\mu_{t^-}^N$ to jump to

$$\mu_t^N = \mu_{t^-}^N + \frac{1}{N} \left[\delta_{(0, S_{t^-}^{j_0,N}, \xi_{t^-, j_0}^{j_0,N})} - \delta_{(1, S_{t^-}^{j_0,N}, \xi_{t^-, j_0}^{j_0,N})} \right] + \frac{1}{N} \sum_{j \neq j_0} \left[\delta_{(V_{t^-}^{j,N}, S_{t^-}^{j,N}, \xi_{t^-, j_0}^{j,N})} - \delta_{(V_{t^-}^{j,N}, S_{t^-}^{j,N}, \xi_{t^-, j_0}^{j,N})} \right],$$

where

$$\xi_{t^-, j_0}^{j,N} \stackrel{\text{def.}}{=} \xi_{t^-}^{j,N} + \frac{1}{N} \sum_{(1, S_{t^-}^{j_0,N}, \tilde{W}) \in \text{supp}(\xi_{t^-}^{j,N})} \left(\delta_{(0, S_{t^-}^{j_0,N}, \tilde{W})} - \delta_{(1, S_{t^-}^{j_0,N}, \tilde{W})} \right).$$

- if $t \in \{\tau_1, \tau_2, \dots\}$, with probability $\frac{\alpha(I_{t^-}^{j_0,N})}{\sum_i \alpha(I_{t^-}^i) \mathbb{1}_{\{V_{t^-}^i = 0\}}} \mathbb{1}_{\{V_{t^-}^{j_0,N} = 0\}}$, the measure leaves the state $\mu_{t^-}^N$ to jump to

$$\mu_t^N = \mu_{t^-}^N + \frac{1}{N} \left[\delta_{(1, 0, \xi_{t^-, j_0}^{j_0,N})} - \delta_{(0, S_{t^-}^{j_0,N}, \xi_{t^-, j_0}^{j_0,N})} \right] + \frac{1}{N} \sum_{j \neq j_0} \left[\delta_{(V_{t^-}^{j,N}, S_{t^-}^{j,N}, \xi_{t^-, j_0}^{j,N})} - \delta_{(V_{t^-}^{j,N}, S_{t^-}^{j,N}, \xi_{t^-, j_0}^{j,N})} \right]$$

where

$$\begin{aligned} \xi_{t^-, j_0}^{j_0,N} &\stackrel{\text{def.}}{=} \xi_{t^-}^{j_0,N} + \sum_{l \neq j_0} \frac{1}{N} \sum_{(\tilde{V}, S_{t^-}^{l,N}, \tilde{W}) \in \text{supp}(\xi_{t^-}^{j_0,N})} \left(\delta_{(\tilde{V}, S_{t^-}^{l,N}, \tilde{W} + \mathbb{1}_{\{U_+^{l,k} \leq p^+(S_{t^-}^{l,N}, \tilde{W})\}})} \right. \\ &\quad \left. - \delta_{(\tilde{V}, S_{t^-}^{l,N}, \tilde{W})} \right) \\ &+ \frac{1}{N} \sum_{(0, S_{t^-}^{j_0,N}, \tilde{W}) \in \text{supp}(\xi_{t^-}^{j_0,N})} \left(\delta_{(1, 0, \tilde{W} + \mathbb{1}_{\{U_+^{j_0,k} \leq p^+(S_{t^-}^{j_0,N}, \tilde{W})\}} - \mathbb{1}_{\{U_-^{j_0,k} \leq p^-(S_{t^-}^{j_0,N}, \tilde{W})\}})} \right. \\ &\quad \left. - \delta_{(0, S_{t^-}^{j_0,N}, \tilde{W})} \right), \\ \xi_{t^-, j_0}^{j,N} &\stackrel{\text{def.}}{=} \xi_{t^-}^{j,N} + \frac{1}{N} \sum_{(0, S_{t^-}^{j_0,N}, \tilde{W}) \in \text{supp}(\xi_{t^-}^{j,N})} \left(\delta_{(1, 0, \tilde{W} - \mathbb{1}_{\{U_-^{j,k} \leq p^-(S_{t^-}^{j,N}, \tilde{W})\}})} - \delta_{(0, S_{t^-}^{j_0,N}, \tilde{W})} \right). \end{aligned}$$

□

4.3.2 Conjecture on the limit process dynamics

We are interested in finding the possible deterministic limit processes $(\mu_t^*)_{t \geq 0}$ of $(\mu_t^N)_{t \geq 0}$ when N tends to infinity, for the weak topology. To do so, for $T > 0$, we consider the empirical

measures on $D_E[0, T]$, the space of càdlàg functions from $[0, T]$ to E , see the Definition 3.15 in Chapter 3. Defining $X^{i, N} = (X_t^{i, N})_{0 \leq t \leq T}$ and $X^* = (X_t^*)_{0 \leq t \leq T}$ with $\mathcal{L}(X_t^*) = \mu_t^*$, we denote $\mu^N = \frac{1}{N} \sum_{i=1}^N \delta_{X^{i, N}}$ and $\mu^* = \mathcal{L}(X^*)$. Thereby, we consider the convergence in law of $(\mu^N) \in \mathcal{P}(D_E[0, T])$ where we recall that

$$E = \{0, 1\} \times \mathbb{R}^+ \times \mathcal{P}(\{0, 1\} \times \mathbb{R}^+ \times \mathbb{Z}).$$

The space $D_E[0, T]$ is equipped with the usual Skorohod J_1 topology. We also consider a distance d on $D_E[0, T]$ with the two properties: first d induces the Skorohod J_1 topology; second, the space $(D_E[0, T], d)$ is a separable and complete space (i.e a Polish space). The existence of such a distance is proved in Billingsley [19, Sec 12]. We also use the property: if a space G is Polish, then $\mathcal{P}(G)$ equipped with the associated weak convergence, is also Polish [76, Thm 17.23].

Assuming that a limit $\mu^* \in \mathcal{P}(D_E[0, T])$ is deterministic, i.e. $\mathcal{L}(\mu^N) \rightarrow_{N \rightarrow \infty} \delta_{\mu^*}$, the aim of this chapter is to find the system satisfied by any limit point which satisfies

$$\forall \Psi \in C_b(E),$$

$$\mathbb{E} \left[\int_0^T \langle \mu_u^N, \Psi \rangle du \right] \rightarrow_{N \rightarrow \infty} \mathbb{E} \left[\int_0^T \langle \mu_u^*, \Psi \rangle du \right] = \int_0^T \langle \mu_u^*, \Psi \rangle du. \quad (4.9)$$

Indeed, knowing $\langle \mu_t^*, \Psi \rangle$ for all $\Psi \in C_b(E)$ is sufficient to determine the measure μ_t^* . Hence, we restrict ourselves to **the study of $\mathbb{E} \langle \mu_t^N, \Psi \rangle$** for $\Psi \in C_b(E)$. Thereby, we are looking for the limit equation of a *typical neuron* that we denote $X_t^* = (V_t^*, S_t^*, \xi_t^*)$ with law μ_t^* .

Assumption 4.13. *Assume that for all $T > 0$, the empirical measures of the system $(\mu^N) \in \mathcal{P}(D_E[0, T])$ converges in law to a (deterministic) probability measure, $\mu^* \in \mathcal{P}(D_E[0, T])$, when N tends to infinity. In particular, the limit (4.9) holds.*

From this assumption, we can link the laws ξ^* and μ^* using the

Notation 4.14. *Any probability measure ξ in $\mathcal{P}(E_m)$ can be split into the following form*

$$\forall (v, A, w) \in \{0, 1\} \times \mathcal{B}(\mathbb{R}^+) \times \mathbb{Z}, \quad \xi(\{v\}, A, \{w\}) = \delta_0(v) \xi^0(A, \{w\}) + \delta_1(v) \xi^1(A, \{w\}),$$

where ξ^0 and ξ^1 are positive measures on $\mathbb{R}^+ \times \mathbb{Z}$ having respectively the masses

$$\sum_{w \in \mathbb{Z}} \int_{\mathbb{R}^+} \xi(\{0\}, ds, \{w\}) \quad \text{and} \quad \sum_{w \in \mathbb{Z}} \int_{\mathbb{R}^+} \xi(\{1\}, ds, \{w\}).$$

Moreover, for all $x, y \in \{0, 1\}$, we denote by $\delta_x \otimes \xi^y$ the measure such that

$$\forall (v, A, w) \in \{0, 1\} \times \mathcal{B}(\mathbb{R}^+) \times \mathbb{Z}, \quad \delta_x \otimes \xi^y(\{v\}, A, \{w\}) \stackrel{\text{def.}}{=} \delta_x(v) \xi^y(A, \{w\}).$$

Finally, when μ_t^* and ξ_t^* admit densities in s , we use the abuse of notation

$$\mu_t^*(\{v\}, ds, d\xi) = \mu_t^*(\{v\}, s, d\xi) ds \quad \text{and} \quad \xi_t^{*v}(ds, \{w\}) = \xi_t^{*v}(s, \{w\}) ds.$$

Remark 4.15. *By application of the strong law of large numbers, we get the two following results. First, Assumptions 4.1 and 4.13 imply that S_0^* has law ρ_0 (which admits a density). Second, Assumptions 4.1 and 4.13 associated to the Remark 4.11 implies that for μ_0^* almost all ξ_0^**

$$\xi_0^*(\cdot, \cdot, \mathbb{Z}) = \mu_0^*(\cdot, \cdot, \mathcal{P}(E_m)) = \mathcal{L}(V_0^*, S_0^*).$$

Moreover, the evolution equations on X_t^* , that are detailed in the next conjecture, imply that we expect the previous equation still hold for all $t \geq 0$ (see Consequence 4.26), for μ_t^* almost all ξ_t^* :

$$\xi_t^*(\cdot, \cdot, \mathbb{Z}) = \mu_t^*(\cdot, \cdot, \mathcal{P}(E_m)) = \mathcal{L}(V_t^*, S_t^*).$$

Some regularity is needed on the functions α and p^\pm for the process to converge. For instance, in order to prove the uniqueness of the limit process, we expect the function α to be Lipschitz. As the study of the convergence goes beyond the scope of this chapter, we only assume the weaker assumptions we need for our computations to hold:

Assumption 4.16. *The functions $\xi \mapsto \alpha(I(\xi))$ and for all $w \in \mathbb{Z}$, $s \mapsto p^\pm(s, w)$, are bounded continuous functions.*

Our computations concerning the dynamics of $\mathbb{E}\langle \mu_t^N, \Psi \rangle$ enables us to conjecture the limit measure dynamics. In order to make this conjecture a theorem, we still need to prove the convergence as N tends to infinity of some terms of this dynamics. It will be explained in further details just after Proposition 4.23 for one term and just before Conjecture 4.25 for another one. Our conjecture concerns the dynamics of the typical neuron $(X_t^*)_{t \geq 0}$. In particular, it exposes the PDE that we expect to be satisfied by the density of ξ_t^* in s and requires the compatibility

Assumption 4.17. *Assume that for all $t \geq 0$, ξ_t^* admits a density of class C^1 in s and in particular at time $t = 0$, for all $w \in \mathbb{Z}$, the following densities satisfy the boundary conditions for μ_0^* almost all ξ_0^* ,*

$$\begin{aligned} \xi_0^{*0}(0, \{w\}) &= 0 \\ \xi_0^{*1}(0, \{w\}) &= \int_{\mathbb{R}^+ \times \mathcal{P}(E_m)} \alpha(I(\xi')) \frac{\xi_0^{*0}(s', \{w\})}{\xi_0^{*0}(s', \mathbb{Z})} \mu_0^*(\{0\}, s', d\xi') ds', \end{aligned}$$

We are now in position to give the conjecture on the limit process. We denote $\mu_t^* = \mathcal{L}(X_t^*) = \mathcal{L}(V_t^*, S_t^*, \xi_t^*)$.

Conjecture 4.18. *Assume that Assumptions 4.1, 4.13, 4.16 and 4.17 hold.*

Then, the limit process $(V_t^, S_t^*, \xi_t^*)_{t \geq 0}$ satisfies the following McKean-Vlasov SDE:*

1. $dS_t^* = dt$,
2. ξ_t^* admits a density in s such that

$$\begin{cases} \frac{d\xi_t^{*0}}{dt}(s, \{w\}) = -\partial_s \xi_t^{*0}(s, \{w\}) + \beta \xi_t^{*1}(s, \{w\}) - \int_{\mathcal{P}(E_m)} \alpha(I(\xi')) \frac{\xi_t^{*0}(s, \{w\})}{\xi_t^{*0}(s, \mathbb{Z})} \mu_t^*(\{0\}, s, d\xi') \\ \xi_t^{*0}(0, \{w\}) = 0, \end{cases}$$

$$\begin{cases} \frac{d\xi_t^{*1}}{dt}(s, \{w\}) = -\partial_s \xi_t^{*1}(s, \{w\}) - \beta \xi_t^{*1}(s, \{w\}) \\ \xi_t^{*1}(0, \{w\}) = \int_{\mathbb{R}^+ \times \mathcal{P}(E_m)} \alpha(I(\xi')) \frac{p^-(S_t^*, w) \xi_t^{*0}(s', \{w+1\}) + (1-p^-(S_t^*, w)) \xi_t^{*0}(s', \{w\})}{\xi_t^{*0}(s', \mathbb{Z})} \mu_t^*(\{0\}, s', d\xi') ds', \end{cases}$$

3. At rate $\beta \mathbb{1}_{\{V_{t^-}^* = 1\}}$, $(V_{t^-}^*, S_{t^-}^*, \xi_{t^-}^*)$ jumps to $(0, S_{t^-}^*, \xi_{t^-}^*)$,

4. At rate $\alpha(I(\xi_{t^-}^*)) \mathbb{1}_{\{V_{t^-}^* = 0\}}$, $(V_{t^-}^*, S_{t^-}^*, \xi_{t^-}^*)$ jumps to $(1, 0, \nu^+(\xi_{t^-}^*))$ where

$$\begin{aligned} \nu^+(\xi_{t^-}^*)(\{v\}, A, \{w\}) &= \int_A p^+(s, w-1) \xi_{t^-}^*(\{v\}, s, \{w-1\}) ds \\ &\quad + \int_A (1-p^+(s, w)) \xi_{t^-}^*(\{v\}, s, \{w\}) ds \end{aligned}$$

for all $(\{v\}, A, \{w\}) \in \mathcal{B}(E_m)$.

We give some details on this limit system. The first and second points are the drift terms and both last points are the jump terms. The drift on ξ_t^* is composed of three parts. First, the linear increase of s which is the term in $-\partial_s$. Second, the mass transport from ξ_t^{*1} to ξ_t^{*0} at rate β . Finally, the mass transport from ξ_t^{*0} to ξ_t^{*1} at a rate depending on μ_t^* with the reset of s in 0 giving the boundary condition in $s = 0$. Note that this system of equations involves the law of its solution: $(X_t^*)_{t \geq 0}$ is thus a McKean–Vlasov process. The process $(X_t^*)_{t \geq 0}$ solves the SDE

$$X_t^* = X_0^* + \int_0^t b(X_u^*, \mu_u^*) du + \int_0^t \int_{\mathbb{R}^+} h(X_{u^-}^*, z) \zeta^*(dz, du)$$

where ζ^* is a Poisson measure on $\mathbb{R}^+ \times \mathbb{R}^+$ with intensity $dz du$, the probability measure μ_u^* is the law of the process X_u^* , and the functions b and h are defined according to the previous Conjecture:

$$\begin{aligned} h(X_{u^-}^*, z) &= \left(1, -S_{u^-}^*, 0\right) \mathbb{1}_{\{z \leq \alpha(I(\xi_{u^-}^*)) \mathbb{1}_{\{V_{u^-}^* = 0\}}\}} + (-1, 0, 0) \mathbb{1}_{\{z \leq \beta \mathbb{1}_{\{V_{u^-}^* = 1\}}\}} \\ b(X_u^*, \mu_u^*) &= \left(0, 1, -\partial_s \xi_u^* + \beta \delta_0 \otimes \xi_u^{*1} - \beta \delta_1 \otimes \xi_u^{*1} + \delta_1 \otimes \nu^{-,1}(S_u^*, \xi_u^*, \mu_u^*) - \delta_0 \otimes \nu^0(\xi_u^*, \mu_u^*)\right) \end{aligned}$$

where ν^0 and $\nu^{-,1}$ are defined respectively in (4.22) and (4.30). Note that the process $(X_t^*)_{t \geq 0}$ lives on the peculiar space $E = \{0, 1\} \times \mathbb{R}^+ \times \mathcal{P}(E_m)$ which contains the space of measures on E_m . We first study the case without plasticity in Section 4.3.3 and then with plasticity in Section 4.3.4.

4.3.3 The case without plasticity

The purpose of this section is to conjectured the evolution equations of the typical neuron $(X_t^*)_{t \geq 0}$ from the ones of finite size neural network $(X_t^N)_{t \geq 0}$, when $p^+ \equiv p^- \equiv 0$ (without plasticity). To do so, we first give the notations and definitions we need to: first, derive the infinitesimal generator of $(\mu_t^N)_{t \geq 0}$ and second, derive the dynamics of $\mathbb{E}\langle \mu_t^N, \Psi \rangle$ where Ψ is a test function with properties described below. Then, we study in detail the most complex terms of this dynamics. Finally, we conjectured the dynamics of any limit point $(X_t^*)_{t \geq 0}$.

Definitions and notations

First, we define a space suitable for our following computations.

Notation 4.19. We denote by $C_b^{1,1}(E)$ the space of functions from E to \mathbb{R} that are bounded, continuously differentiable with respect to their second variable, Fréchet differentiable (see Definition A.6) with respect to their third variable and finally, both these derivatives are bounded.

The Fréchet derivative with respect to the third variable is defined on the larger space of the signed measure on E_m , $\mathcal{M}(E_m)$, equipped with the total variation norm, $\|\cdot\|_{TV}$, whose definition is given in the appendix, Definition A.2. In particular, $(\mathcal{M}(E_m), \|\cdot\|_{TV})$ is a Banach space, see [4, Rk 1.7], enabling us to define the Fréchet derivative of $\Psi \in C_b^{1,1}(E)$. For all $\Psi \in C_b^{1,1}(E)$, for all $(v, s, \xi) \in E$ and $h \in \mathcal{M}(E_m)$, we denote by $\partial_\xi \Psi(v, s, \xi) \cdot h$ the Fréchet derivative of Ψ at (v, s, ξ) in the direction h (see Definition A.6).

Second, we define two applications, η^N and ν^N , in order to ease the writing and reading processes. The application η^N is used to describe the jumps of the empirical measures $\xi_t^{1,N}, \dots, \xi_t^{N,N}$ when

a neuron jumps from 1 to 0. The application ν^N is used to describe the jumps of $\xi_t^{1,N}, \dots, \xi_t^{N,N}$ when a neuron spikes (jump of a neuron from 0 to 1).

Notation 4.20. The symbol $\stackrel{\mathbb{E}}{=}$ refers to the equality between the expectations of two random variables:

$$\mathbb{E}[X] = \mathbb{E}[Y] \quad \Leftrightarrow \quad X \stackrel{\mathbb{E}}{=} Y.$$

We denote by η^N and ν^N two operators associating to a pair $(s, \xi) \in \mathbb{R}^+ \times \mathcal{P}_N(E_m)$ the following signed measures on E_m :

$$\eta^N(s, \xi) \stackrel{\text{def.}}{=} \frac{1}{N} \sum_{(1,s,\tilde{w}) \in \text{supp}(\xi)} \left[\delta_{(0,s,\tilde{w})} - \delta_{(1,s,\tilde{w})} \right]$$

and

$$\nu^N(s, \xi) \stackrel{\text{def.}}{=} \frac{1}{N} \sum_{(0,s,\tilde{w}) \in \text{supp}(\xi)} \left[\delta_{(1,0,\tilde{w})} - \delta_{(0,s,\tilde{w})} \right].$$

Remark 4.21. Let t be any spiking time of the neuron i , i.e. $V_{t^-}^{i,N} = 0$ and $V_t^{i,N} = 1$. Then, according to Lemma 4.2, for every j , there is a unique $(\tilde{v}^i, \tilde{s}^i, \tilde{w}^i)$ in the support of $\xi_{t^-}^{j,N}$ such that $\tilde{s}^i = S_t^{i,N}$. So, $\nu^N(S_{t^-}^{i,N}, \xi_{t^-}^{j,N})$ is a signed measure with two atoms.

$$\nu^N(S_{t^-}^{i,N}, \xi_{t^-}^{j,N}) = \frac{1}{N} \delta_{(1,0,\tilde{w}^i)} - \frac{1}{N} \delta_{(0,S_{t^-}^{i,N},\tilde{w}^i)}.$$

Similarly, at a time t when $V_{t^-}^{i,N} = 1$ and $V_t^{i,N} = 0$, there is a unique $(\tilde{v}^i, \tilde{s}^i, \tilde{w}^i)$ in the support of $\xi_{t^-}^{j,N}$ such that $\tilde{s}^i = S_t^{i,N}$ and $\eta^N(S_{t^-}^{i,N}, \xi_{t^-}^{j,N})$ is a signed measure with two atoms

$$\eta^N(S_{t^-}^{i,N}, \xi_{t^-}^{j,N}) = \frac{1}{N} \delta_{(0,S_{t^-}^{i,N},\tilde{w}^i)} - \frac{1}{N} \delta_{(1,S_{t^-}^{i,N},\tilde{w}^i)}.$$

We define here η^N and ν^N only on $\mathcal{P}_N(E_m)$, the set of empirical measures of order N over E_m . Indeed, with the changes occurring when passing to the large N limit, these measures will no longer be required. New ones are defined on the all spaces $\mathcal{P}(E_m)$ later on, see Propositions 4.22 and 4.23.

Third, we define the derivative of any probability measure ε on \mathbb{R}^+ by

$$\forall \phi \in C_b^1(\mathbb{R}^+) \text{ with bounded derivative, } \quad \langle -\varepsilon', \phi \rangle \stackrel{\text{def.}}{=} \langle \varepsilon, \phi' \rangle. \quad (4.10)$$

Finally, let us define for all $\xi \in \mathcal{P}(E_m)$ and $t \geq 0$, the probability measure $\xi \oplus t$ such that

$$\forall (v, A, w) \in \mathcal{B}(\{0, 1\}, [t, +\infty[, \mathbb{Z}), \quad (\xi \oplus t)(\{v\}, A, \{w\}) = \xi(\{v\}, A - t, \{w\}).$$

Generator of $(\mu_t^N)_{t \geq 0}$

We now look for the generator of the Markov process $(\mu_t^N)_{t \geq 0}$. We describe the generator only on the set of functions $\Phi \in C_b(\mathcal{P}(E))$ for which there exists $\psi \in C_b^{1,1}(E_m)$ such that

$$\forall \mu \in \mathcal{P}(E), \quad \Phi(\mu) = \langle \mu, \psi \rangle.$$

We split the generator into the drift term and the jump term using the equality $1 = \mathbb{1}_{t < T_1} + \mathbb{1}_{t \geq T_1}$. Let $\mu_0 \in \mathcal{P}_N(E)$ such that $\mu_0 = \frac{1}{N} \sum_i \delta_{(v^i, s^i, \xi^i)}$ with $(v^1, s^1, \xi^1), \dots, (v^N, s^N, \xi^N)$ are in E . Let $\Psi \in C_b^{1,1}(E)$, then

$$\begin{aligned} \langle \mu_t^N - \mu_0, \Psi \rangle &= \langle \mathbb{1}_{t < T_1} \mu_t^N - \mu_0, \Psi \rangle + \langle \mathbb{1}_{t \geq T_1} \mu_t^N, \Psi \rangle \\ &= \frac{1}{N} \sum_i (\Psi(v^i, s^i + t, \xi^i \oplus t) \mathbb{1}_{t < T_1} - \Psi(v^i, s^i, \xi^i)) + \langle \mathbb{1}_{t \geq T_1} \mu_t^N, \Psi \rangle \\ &= \frac{1}{N} \sum_i \underbrace{(\Psi(v^i, s^i + t, \xi^i \oplus t) - \Psi(v^i, s^i, \xi^i))}_{\textcircled{D}} \\ &\quad + \underbrace{\langle \mathbb{1}_{t \geq T_1} \mu_t^N, \Psi \rangle - \frac{1}{N} \sum_i \Psi(v^i, s^i + t, \xi^i \oplus t) \mathbb{1}_{t \geq T_1}}_{\textcircled{J}}. \end{aligned}$$

Thereby, we obtain that the drift term satisfies

$$\begin{aligned} \lim_{t \rightarrow 0} \frac{\mathbb{E} \textcircled{D}}{t} &= \lim_{t \rightarrow 0} \frac{1}{N} \sum_i \frac{\mathbb{E} [\Psi(v^i, s^i + t, \xi^i \oplus t)] - \Psi(v^i, s^i, \xi^i)}{t} \\ &= \frac{1}{N} \sum_i \partial_s \Psi(v^i, s^i, \xi^i) + \partial_\xi \Psi(v^i, s^i, \xi^i) \cdot (-\partial_s \xi^i) \\ &= \int_{\mathcal{P}(E_m)} \partial_s \Psi(v, s, \xi) + \partial_\xi \Psi(v, s, \xi) \cdot (-\partial_s \xi) \mu_0(dv, ds, d\xi), \end{aligned}$$

where $-\partial_s \xi$ is the derivative defined as in (4.10).

For the jump term \textcircled{J} , when we take the limit as t tends to 0 in $\frac{\mathbb{E} \textcircled{J}}{t}$, the only term left is the one due to the first jump at time T_1 . Indeed, all the other terms are of order t or more. We thus obtain that

$$\begin{aligned} \lim_{t \rightarrow 0} \frac{\mathbb{E} \textcircled{J}}{t} &= \frac{1}{N} \sum_i \mathbb{1}_{\{v^i=1\}} \beta \left\{ \left[\Psi(0, s^i, \xi^i + \eta^N(s^i, \xi^i)) - \Psi(1, s^i, \xi^i) \right] \right. \\ &\quad \left. + \sum_{j \neq i} \left[\Psi(v^j, s^j, \xi^j + \eta^N(s^i, \xi^j)) - \Psi(v^j, s^j, \xi^j) \right] \right\} \\ &\quad + \frac{1}{N} \sum_i \mathbb{1}_{\{v^i=0\}} \alpha(I(\xi^i)) \left\{ \left[\Psi(1, 0, \xi^i + \nu^N(s^i, \xi^i)) - \Psi(0, s^i, \xi^i) \right] \right. \\ &\quad \left. + \sum_{j \neq i} \left[\Psi(v^j, s^j, \xi^j + \nu^N(s^i, \xi^j)) - \Psi(v^j, s^j, \xi^j) \right] \right\}. \end{aligned}$$

In the previous equation, by adding and removing the missing terms in the sums inside the braces,

we get

$$\begin{aligned} \lim_{t \rightarrow 0} \textcircled{\mathbf{J}} &= \frac{1}{N} \sum_i \mathbb{1}_{\{v^i=1\}} \beta \left\{ \left[\Psi\left(0, s^i, \xi^i + \eta^N(s^i, \xi^i)\right) - \Psi\left(1, s^i, \xi^i + \eta^N(s^i, \xi^i)\right) \right] \right. \\ &\quad \left. + \sum_j \left[\Psi\left(v^j, s^j, \xi^j + \eta^N(s^i, \xi^j)\right) - \Psi\left(v^j, s^j, \xi^j\right) \right] \right\} \\ &\quad + \frac{1}{N} \sum_i \mathbb{1}_{\{v^i=0\}} \alpha(I(\xi^i)) \left\{ \left[\Psi\left(1, 0, \xi^i + \nu^N(s^i, \xi^i)\right) - \Psi\left(0, s^i, \xi^i + \nu^N(s^i, \xi^i)\right) \right] \right. \\ &\quad \left. + \sum_j \left[\Psi\left(v^j, s^j, \xi^j + \nu^N(s^i, \xi^j)\right) - \Psi\left(v^j, s^j, \xi^j\right) \right] \right\}. \end{aligned}$$

Dynamics of $\mathbb{E}\langle \mu_t^N, \Psi \rangle$

From the Markov property of the process $(\mu_t^N)_{t \geq 0}$, we deduce that for all $\Psi \in C_b^{1,1}(E)$,

$$\begin{aligned} \langle \mu_t^N, \Psi \rangle &\stackrel{\mathbb{E}}{=} \langle \mu_0^N, \Psi \rangle + \int_0^t \int_E \left(\partial_s \Psi(v, s, \xi) + \partial_\xi \Psi(v, s, \xi) \cdot (-\partial_s \xi) \right) \mu_u^N(dv, ds, d\xi) du \\ &\quad + \int_0^t \sum_i \mathbb{1}_{\{V_{u^-}^{i,N}=1\}} \frac{\beta}{N} \left\{ \left[\Psi\left(0, S_{u^-}^{i,N}, \xi_{u^-}^{i,N} + \eta^N(S_{u^-}^{i,N}, \xi_{u^-}^{i,N})\right) - \Psi\left(1, S_{u^-}^{i,N}, \xi_{u^-}^{i,N} + \eta^N(S_{u^-}^{i,N}, \xi_{u^-}^{i,N})\right) \right] \right. \\ &\quad \left. + \sum_j \left[\Psi\left(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u^-}^{j,N} + \eta^N(S_{u^-}^{i,N}, \xi_{u^-}^{j,N})\right) - \Psi\left(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u^-}^{j,N}\right) \right] \right\} du \\ &\quad + \int_0^t \sum_i \mathbb{1}_{\{V_{u^-}^{i,N}=0\}} \frac{\alpha(I_{u^-}^{i,N})}{N} \left\{ \left[\Psi\left(1, 0, \xi_{u^-}^{i,N} + \nu^N(S_{u^-}^{i,N}, \xi_{u^-}^{i,N})\right) - \Psi\left(0, S_{u^-}^{i,N}, \xi_{u^-}^{i,N} + \nu^N(S_{u^-}^{i,N}, \xi_{u^-}^{i,N})\right) \right] \right. \\ &\quad \left. + \sum_j \left[\Psi\left(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u^-}^{j,N} + \nu^N(S_{u^-}^{i,N}, \xi_{u^-}^{j,N})\right) - \Psi\left(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u^-}^{j,N}\right) \right] \right\} du. \end{aligned}$$

We finally get that

$$\begin{aligned} \langle \mu_t^N, \Psi \rangle &\stackrel{\mathbb{E}}{=} \langle \mu_0^N, \Psi \rangle + \int_0^t \int_E \left(\partial_s \Psi(v, s, \xi) + \partial_\xi \Psi(v, s, \xi) \cdot (-\partial_s \xi) \right) \mu_u^N(dv, ds, d\xi) du \\ &\quad + \int_0^t \int_{\mathbb{R}^+ \times \mathcal{P}(E_m)} \beta \left[\Psi\left(0, s, \xi + \eta^N(s, \xi)\right) - \Psi\left(1, s, \xi + \eta^N(s, \xi)\right) \right] \mu_u^N(\{1\}, ds, d\xi) du + \textcircled{1} \\ &\quad + \int_0^t \int_{\mathbb{R}^+ \times \mathcal{P}(E_m)} \alpha(I(\xi)) \left[\Psi\left(1, 0, \xi + \nu^N(s, \xi)\right) - \Psi\left(0, s, \xi + \nu^N(s, \xi)\right) \right] \mu_u^N(\{0\}, ds, d\xi) du + \textcircled{2} \end{aligned} \tag{4.11}$$

where ① and ② are the most complex elements:

$$\begin{aligned} \textcircled{1} \stackrel{\text{def.}}{=} & \int_0^t \sum_i \mathbb{1}_{\{V_{u^-}^{i,N}=1\}} \frac{\beta}{N} \sum_j \left[\Psi \left(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u^-}^{j,N} + \eta^N(S_{u^-}^{i,N}, \xi_{u^-}^{j,N}) \right) \right. \\ & \left. - \Psi \left(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u^-}^{j,N} \right) \right] du, \end{aligned} \quad (4.12)$$

$$\begin{aligned} \textcircled{2} \stackrel{\text{def.}}{=} & \int_0^t \sum_i \mathbb{1}_{\{V_{u^-}^{i,N}=0\}} \frac{\alpha(I_{u^-}^{i,N})}{N} \sum_j \left[\Psi \left(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u^-}^{j,N} + \nu^N(S_{u^-}^{i,N}, \xi_{u^-}^{j,N}) \right) \right. \\ & \left. - \Psi \left(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u^-}^{j,N} \right) \right] du. \end{aligned} \quad (4.13)$$

Drift term due to the returns to resting potential

When a neuron returns to its resting potential state 0, the total variation of the jumps of the empirical measures $(\xi_t^{1,N})_{1 \leq i \leq N}, \dots, (\xi_t^{N,N})_{1 \leq i \leq N}$ are of order $\frac{1}{N}$. Hence, as N tends to infinity, the limit of the term ① depends on the Fréchet derivative of Ψ . The direction of this derivative describes the way some of the mass of the measure ξ^* is transported. This transport can be described informally as follows: between time $t = 0$ and time $t = \epsilon$ small, for all s, w , a proportion $\beta\epsilon$ of the mass of ξ_0^* that was in $(1, s, w)$ is transported to $(0, s, w)$.

Proposition 4.22. *Under Assumptions 4.1 and 4.13, for all $\Psi \in C_b^{1,1}(E)$,*

$$\lim_{N \rightarrow \infty} \textcircled{1} \stackrel{\mathbb{E}}{=} \int_0^t \left\{ \int_E \partial_\xi \Psi(v, s, \xi) \cdot (\beta\delta_0 \otimes \xi^1 - \beta\delta_1 \otimes \xi^1) \mu_u^*(dv, ds, d\xi) \right\} du. \quad (4.14)$$

Proof. By Lemma 4.2, for all j , the support of $\xi_{u^-}^{j,N}$ almost surely does not possess two points with the same second coordinate. Hence, we have the following upper bound, for all $s \in \mathbb{R}^+$, almost surely,

$$\|\eta^N(s, \xi_{u^-}^{j,N})\|_{TV} \leq \frac{2}{N}. \quad (4.15)$$

We write the term of ① which is between square brackets using the Fréchet derivative:

$$\begin{aligned} & \Psi \left(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u^-}^{j,N} + \eta^N(S_{u^-}^{i,N}, \xi_{u^-}^{j,N}) \right) - \Psi \left(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u^-}^{j,N} \right) \\ & = \partial_\xi \Psi \left(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u^-}^{j,N} \right) \cdot \left(\eta^N(S_{u^-}^{i,N}, \xi_{u^-}^{j,N}) \right) + \mathcal{O}_{N\infty} \left(\|\eta^N(S_{u^-}^{i,N}, \xi_{u^-}^{j,N})\|_{TV} \right). \end{aligned}$$

Hence, we have:

$$\begin{aligned} \textcircled{1} \stackrel{\mathbb{E}}{=} & \int_0^t \sum_i \mathbb{1}_{\{V_{u^-}^{i,N}=1\}} \frac{\beta}{N} \sum_j \left[\partial_\xi \Psi \left(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u^-}^{j,N} \right) \cdot \left(\eta^N(S_{u^-}^{i,N}, \xi_{u^-}^{j,N}) \right) \right. \\ & \left. + \mathcal{O}_{N\infty} \left(\|\eta^N(S_{u^-}^{i,N}, \xi_{u^-}^{j,N})\|_{TV} \right) \right] du. \end{aligned}$$

Thus, with the upper bound (4.15), we obtain that

$$\textcircled{1} \stackrel{\mathbb{E}}{=} \int_0^t \sum_i \mathbb{1}_{\{V_{u^-}^{i,N}=1\}} \frac{\beta}{N} \sum_j \left[\partial_\xi \Psi \left(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u^-}^{j,N} \right) \cdot \left(\eta^N(S_{u^-}^{i,N}, \xi_{u^-}^{j,N}) \right) \right] du + \mathcal{O}_{N\infty}(1).$$

Now, by linearity of $\partial_\xi \Psi$ and inverting the sums, we get

$$\textcircled{1} \stackrel{\mathbb{E}}{=} \int_0^t \frac{1}{N} \sum_j \left[\partial_\xi \Psi(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u^-}^{j,N}) \cdot \left(\sum_i \mathbb{1}_{\{V_{u^-}^{i,N}=1\}} \beta \eta^N(S_{u^-}^{i,N}, \xi_{u^-}^{i,N}) \right) \right] du + \mathcal{O}_{N\infty}(1).$$

From Remark 4.11, for all j and $u \geq 0$, $\xi_{u^-}^{j,N}$ and $\mu_{u^-}^N$ have the same support in V and S , and from Lemma 4.2, these supports are N almost surely distinct points. Hence, almost surely, for any function $f : \mathbb{R}^+ \times \mathbb{Z} \rightarrow \mathbb{R}^+$,

$$\forall j, \sum_i \sum_{(1, S_{u^-}^{i,N}, \tilde{w}) \in \text{supp}(\xi_{u^-}^{j,N})} f(S_{u^-}^{i,N}, \tilde{w}) = \sum_{(1, \tilde{s}, \tilde{w}) \in \text{supp}(\xi_{u^-}^{j,N})} f(\tilde{s}, \tilde{w}).$$

We deduce that almost surely,

$$\begin{aligned} \sum_i \mathbb{1}_{\{V_{u^-}^{i,N}=1\}} \beta \eta^N(S_{u^-}^{i,N}, \xi_{u^-}^{j,N}) &= \sum_i \frac{\beta}{N} \sum_{(1, S_{u^-}^{i,N}, \tilde{w}) \in \text{supp}(\xi_{u^-}^{j,N})} \left[\delta_{(0, S_{u^-}^{i,N}, \tilde{w})} - \delta_{(1, S_{u^-}^{i,N}, \tilde{w})} \right] \\ &= \frac{\beta}{N} \sum_{(1, \tilde{s}, \tilde{w}) \in \text{supp}(\xi_{u^-}^{j,N})} \left[\delta_{(0, \tilde{s}, \tilde{w})} - \delta_{(1, \tilde{s}, \tilde{w})} \right] \\ &= \beta \delta_0 \otimes \xi_{u^-}^{j,N}(\{1\}, \cdot, \cdot) - \beta \delta_1 \otimes \xi_{u^-}^{j,N}(\{1\}, \cdot, \cdot), \end{aligned}$$

where we used Notation 4.14. We deduce that

$$\textcircled{1} \stackrel{\mathbb{E}}{=} \int_0^t \left\{ \int_E \partial_\xi \Psi(v, s, \xi) \cdot (\beta \delta_0 \otimes \xi^1 - \beta \delta_1 \otimes \xi^1) \mu_u^N(dv, ds, d\xi) \right\} du + \mathcal{O}_{N\infty}(1).$$

The application $\xi \mapsto \xi^1$ is bounded continuous. Thus, as $\Psi \in C_b^{1,1}(E)$, the application

$$(s, v, \xi) \mapsto \partial_\xi \Psi(v, s, \xi) \cdot (\beta \delta_0 \otimes \xi^1 - \beta \delta_1 \otimes \xi^1)$$

is bounded continuous. We conclude that under Assumption 4.13,

$$\lim_{N \rightarrow \infty} \textcircled{1} \stackrel{\mathbb{E}}{=} \int_0^t \left\{ \int_E \partial_\xi \Psi(v, s, \xi) \cdot (\beta \delta_0 \otimes \xi^1 - \beta \delta_1 \otimes \xi^1) \mu_u^*(dv, ds, d\xi) \right\} du.$$

□

Drift term due to action potentials

As N tends to infinity, the limit of the term $\textcircled{2}$ also gives a drift term for ξ^* . This term is more difficult to deal with because the spiking rates, $(\mathbb{1}_{\{V_t^{i,N}=0\}} \alpha(I_t^{i,N}))_{1 \leq i \leq N}$, depend on the actual state of the neural network X_t^N . Informally, the limit of the term $\textcircled{2}$ represents the following mass transport on the probability measures: between time $t = 0$ and time $t = \epsilon$ small, for all s, w , a proportion $\epsilon \int_{\mathcal{P}(E_m)} \frac{\alpha(I(\tilde{\xi}))}{\mu_0^*(\{0\}, s, \mathcal{P}(E_m))} \mu_0^*(\{0\}, s, d\tilde{\xi})$ of the mass of ξ_0^* that was at $(0, s, w)$ is transported to $(1, [0, \epsilon], w)$.

Proposition 4.23. *We assume that Assumptions 4.1 and 4.16 hold. For all Ψ Fréchet differentiable with respect to its third variable:*

$$\textcircled{2} \stackrel{\mathbb{E}}{=} \int_0^t \int_E \partial_\xi \Psi(v, s, \xi) \cdot \left(\delta_1 \otimes \nu^1(\xi, \mu_{u^-}^N) - \delta_0 \otimes \nu^0(\xi, \mu_{u^-}^N) \right) \mu_{u^-}^N(dv, ds, d\xi) du + \mathcal{O}_{N\infty}(1).$$

where ν^0 and ν^1 are defined in the proof: equations (4.18) and (4.19).

Proof. By Lemma 4.2, for any j , the support of $\xi_{u^-}^{j,N}$ almost surely does not contain two atoms with the same S . Furthermore, we have assumed that α is bounded. Thus, we have the following upper bounds, for all $i, j \in \llbracket 1, N \rrbracket$, almost surely,

$$\|\alpha(I_{u^-}^{i,N})\nu^N(S_{u^-}^{i,N}, \xi_{u^-}^{j,N})\|_{TV} \leq \frac{2\alpha_M}{N}. \quad (4.16)$$

Using the Fréchet derivative as previously, from (4.13), we obtain:

$$\begin{aligned} \textcircled{2} &\stackrel{\mathbb{E}}{=} \int_0^t \frac{1}{N} \sum_j \left[\partial_\xi \Psi(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u^-}^{j,N}) \cdot \left(\sum_i \mathbb{1}_{\{V_{u^-}^{i,N}=0\}} \alpha(I_{u^-}^{i,N})\nu^N(S_{u^-}^{i,N}, \xi_{u^-}^{j,N}) \right) \right] du \\ &\quad + \mathcal{O}_{N\infty}(1). \end{aligned}$$

We now look at the term between round brackets, it gives the direction of the Fréchet derivative. It is a measure on E_m . For all j , we denote

$$\begin{aligned} \textcircled{2.a} &\stackrel{\text{def.}}{=} \sum_i \mathbb{1}_{\{V_{u^-}^{i,N}=0\}} \alpha(I_{u^-}^{i,N})\nu^N(S_{u^-}^{i,N}, \xi_{u^-}^{j,N}) \\ &= \sum_i \mathbb{1}_{\{V_{u^-}^{i,N}=0\}} \alpha(I_{u^-}^{i,N}) \frac{1}{N} \sum_{(0, S_{u^-}^{i,N}, \tilde{w}) \in \text{supp}(\xi_{u^-}^{j,N})} \left[\delta_{(1,0,\tilde{w})} - \delta_{(0, S_{u^-}^{i,N}, \tilde{w})} \right] \\ &= \frac{1}{N} \sum_i \mathbb{1}_{\{V_{u^-}^{i,N}=0\}} \alpha(I_{u^-}^{i,N}) \left[\delta_{(1,0)} \otimes \left(\sum_{\tilde{w} \in \mathbb{Z}} \mathbb{1}_{\{(0, S_{u^-}^{i,N}, \tilde{w}) \in \text{supp}(\xi_{u^-}^{j,N})\}} \delta_{\tilde{w}} \right) \right. \\ &\quad \left. - \delta_{(0, S_{u^-}^{i,N})} \otimes \left(\sum_{\tilde{w} \in \mathbb{Z}} \mathbb{1}_{\{(0, S_{u^-}^{i,N}, \tilde{w}) \in \text{supp}(\xi_{u^-}^{j,N})\}} \delta_{\tilde{w}} \right) \right]. \end{aligned}$$

But $\sum_{\tilde{w} \in \mathbb{Z}} \mathbb{1}_{\{(0, S_{u^-}^{i,N}, \tilde{w}) \in \text{supp}(\xi_{u^-}^{j,N})\}} \delta_{\tilde{w}}$ is the conditional law $\xi_{u^-}^{j,N}(\{0\}, S_{u^-}^{i,N}, \cdot)$. This incites us to define for any $w \in \mathbb{Z}$, the Radon-Nikodym derivative $s \mapsto \frac{d\xi^0(\cdot, \{w\})}{d\xi^0(\cdot, \mathbb{Z})}(s)$ which is well defined as $\xi^0(\cdot, \{w\})$ is absolutely continuous with respect to $\xi^0(\cdot, \mathbb{Z})$. Thereby, we define the function $\gamma_\xi : \mathbb{R}^+ \times \mathbb{Z} \rightarrow [0, 1]$ such that

$$\forall (s, w) \in \mathbb{R}^+ \times \mathbb{Z}, \quad \gamma_\xi(s, w) = \frac{d\xi^0(\cdot, \{w\})}{d\xi^0(\cdot, \mathbb{Z})}(s). \quad (4.17)$$

Note that the for all $A \in \mathcal{B}(\mathbb{R}^+)$, we have

$$\xi^0(A, \{w\}) = \int_A \gamma_\xi(s, w) \xi^0(ds, \mathbb{Z}).$$

As $\xi^0(A, \{w\}) \leq \xi^0(A, \mathbb{Z})$, we deduce that taking $\gamma_\xi \leq 1$ is legitimate (γ_ξ is unique up to a $\xi^0(\cdot, \mathbb{Z})$ null set). In particular, we obtain that

$$\sum_{\tilde{w} \in \mathbb{Z}} \mathbb{1}_{\{(0, S_{u^-}^{i,N}, \tilde{w}) \in \text{supp}(\xi_{u^-}^{j,N})\}} \delta_{\tilde{w}} = \sum_{\tilde{w} \in \mathbb{Z}} \gamma_{\xi_{u^-}^{j,N}}(S_{u^-}^{i,N}, \tilde{w}) \delta_{\tilde{w}}.$$

Thereby, using the measure $\mu_{u^-}^N(\{0\}, \cdot, \cdot) = \frac{1}{N} \sum_i \mathbb{1}_{\{V_{u^-}^{i,N}=0\}} \delta_{(S_{u^-}^{i,N}, \xi_{u^-}^{i,N})}$, we obtain that

$$\begin{aligned} \textcircled{2.a} &= \sum_{\tilde{w} \in \mathbb{Z}} \frac{1}{N} \sum_i \mathbb{1}_{\{V_{u^-}^{i,N}=0\}} \alpha(I_{u^-}^{i,N}) \gamma_{\xi_{u^-}^{j,N}}(S_{u^-}^{i,N}, \tilde{w}) \left(\delta_{(1,0,\tilde{w})} - \delta_{(0, S_{u^-}^{i,N}, \tilde{w})} \right) \\ &= \sum_{\tilde{w} \in \mathbb{Z}} \int_{\mathbb{R}^+ \times \mathcal{P}(E_m)} \alpha(I(\tilde{\xi})) \gamma_{\xi_{u^-}^{j,N}}(s, \tilde{w}) \left(\delta_{(1,0,\tilde{w})} - \delta_{(0,s,\tilde{w})} \right) d\mu_{u^-}^N(\{0\}, ds, d\tilde{\xi}). \end{aligned}$$

This last equation incites us to define the two following measures. For all probability measures $\xi \in \mathcal{P}(E_m)$ and $\mu \in \mathcal{P}(E)$, we define $\nu^0(\xi, \mu)$ and $\nu^1(\xi, \mu)$ the measures on $\mathbb{R}^+ \times \mathbb{Z}$ such that for all $(A, w) \in \mathcal{B}(\mathbb{R}^+) \times \mathbb{Z}$,

$$\nu^0(\xi, \mu)(A, \{w\}) \stackrel{def.}{=} \int_{\mathcal{P}(E_m)} \int_A \alpha(I(\tilde{\xi})) \gamma_\xi(s, w) \mu(\{0\}, ds, d\tilde{\xi}), \quad (4.18)$$

$$\nu^1(\xi, \mu)(A, \{w\}) \stackrel{def.}{=} \mathbb{1}_{\{0 \in A\}} \nu^0(\xi, \mu)(\mathbb{R}^+, \{w\}). \quad (4.19)$$

Hence, we deduce that

$$\textcircled{2.a} = \delta_1 \otimes \nu^1(\xi_{u^-}^{j,N}, \mu_{u^-}^N) - \delta_0 \otimes \nu^0(\xi_{u^-}^{j,N}, \mu_{u^-}^N)$$

and then

$$\begin{aligned} \textcircled{2} &\stackrel{\mathbb{E}}{=} \int_0^t \frac{1}{N} \sum_j \partial_\xi \Psi(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u^-}^{j,N}) \cdot \textcircled{2.a} du + \mathcal{O}_{N\infty}(1) \\ &\stackrel{\mathbb{E}}{=} \int_0^t \int_E \partial_\xi \Psi(v, s, \xi) \cdot \left(\delta_1 \otimes \nu^1(\xi, \mu_{u^-}^N) - \delta_0 \otimes \nu^0(\xi, \mu_{u^-}^N) \right) \mu_{u^-}^N(dv, ds, d\xi) du + \mathcal{O}_{N\infty}(1). \end{aligned}$$

□

From this result, under Assumption 4.13, as N tends to infinity, we expect $\mathbb{E}(\textcircled{2})$ to converge to

$$\int_0^t \left\{ \int_E \partial_\xi \Psi(v, s, \xi) \cdot \left(\delta_1 \otimes \nu^1(\xi, \mu_u^*) - \delta_0 \otimes \nu^0(\xi, \mu_u^*) \right) \mu_u^*(dv, ds, d\xi) \right\} du. \quad (4.20)$$

This convergence does not hold a priori because the functions

$$(v, s, \xi) \mapsto \partial_\xi \Psi(v, s, \xi) \cdot [\delta_1 \otimes \nu^1(\xi, \mu_u^N) - \delta_0 \otimes \nu^0(\xi, \mu_u^N)] \quad (4.21)$$

are not a priori continuous. We expect to show this convergence using a sequence (indexed by N) of continuous functions both getting closer and closer to (4.21) and converging to

$$(v, s, \xi) \mapsto \partial_\xi \Psi(v, s, \xi) \cdot [\delta_1 \otimes \nu^1(\xi, \mu_u^*) - \delta_0 \otimes \nu^0(\xi, \mu_u^*)].$$

Remark 4.24. *If for all $u \geq 0$, μ_u^* admits a density in s , then for all $(A, \{w\}) \in \mathcal{B}(\mathbb{R}^+) \times \mathbb{Z}$ and $\xi \in \text{supp}(\mu_u^*)$ (μ_u^* has mass only on ξ with density in s , see Consequence 4.26), we have*

$$\begin{aligned} \nu^0(\xi, \mu_u^*)(A, \{w\}) &= \int_A \int_{\mathcal{P}(E_m)} \alpha(I(\tilde{\xi})) \frac{\xi^0(s, \{w\})}{\xi^0(s, \mathbb{Z})} \mu_u^*(\{0\}, ds, d\tilde{\xi}), \\ \nu^1(\xi, \mu_u^*)(A, \{w\}) &= \mathbb{1}_{\{0 \in A\}} \nu^0(\xi, \mu_u^*)(\mathbb{R}^+, \{w\}). \end{aligned} \quad (4.22)$$

Equation on μ_t^*

We denote by $\mu_t^* = \mathcal{L}(X_t^*) = \mathcal{L}(V_t^*, S_t^*, \xi_t^*)$. We have seen previously that under Assumption 4.1, for all i, j and $u \geq 0$, we have almost surely

$$\|\nu^N(S_{u^-}^{i,N}, \xi_{u^-}^{j,N})\|_{TV} \leq \frac{2}{N} \quad \text{and} \quad \|\eta^N(S_{u^-}^{i,N}, \xi_{u^-}^{j,N})\|_{TV} \leq \frac{2}{N}.$$

Thus, under Assumption 4.13, taking the limit as N tends to infinity in equation (4.11), we obtain that

$$\begin{aligned} \lim_{N \rightarrow \infty} \langle \mu_t^N, \Psi \rangle &\stackrel{\mathbb{E}}{=} \langle \mu_0^*, \Psi \rangle + \int_0^t \int_E \partial_s \Psi(v, s, \xi) \mu_u^*(dv, ds, d\xi) du \\ &+ \lim_{N \rightarrow \infty} \int_0^t \int_E \partial_\xi \Psi(v, s, \xi) \cdot (-\partial_s \xi) \mu_u^N(dv, ds, d\xi) du \\ &+ \int_0^t \int_{\mathbb{R}^+ \times \mathcal{P}(E_m)} \beta \left[\Psi(0, s, \xi(s, \xi)) - \Psi(1, s, \xi(s, \xi)) \right] \mu_u^*(\{1\}, ds, d\xi) du + \lim_{N \rightarrow \infty} \textcircled{1} \\ &+ \int_0^t \int_{\mathbb{R}^+ \times \mathcal{P}(E_m)} \alpha(I(\xi)) \left[\Psi(1, 0, \xi) - \Psi(0, s, \xi) \right] \mu_u^*(\{0\}, ds, d\xi) du + \lim_{N \rightarrow \infty} \textcircled{2}. \end{aligned}$$

It is reasonable to expect that the expectation of the term with the weak derivative $-\partial_s \xi$ converges, when N tends to infinity, to

$$\int_0^t \int_E \partial_\xi \Psi(v, s, \xi) \cdot (-\partial_s \xi) \mu_u^*(dv, ds, d\xi) du.$$

Then, from Propositions 4.22 and 4.23, we can formulate

Conjecture 4.25. *Assume that Assumptions 4.1, 4.13 and 4.16 hold. Then, the dynamics of μ_t^* is given by:*

- $\mathcal{L}(X_0^*) = \lim_{N \rightarrow \infty} \mathcal{L}(X_0^{1,N}) = \mu_0^*$ and in particular S_0^* is distributed by ρ_0 which is absolutely continuous with respect to the Lebesgue measure,
- for all $t \geq 0$, for all $\Psi \in C_b^{1,1}(E)$ (continuously differentiable with respect to its second variable and Fréchet differentiable with respect to its third variable), and using Notation 4.14,

$$\begin{aligned} \langle \mu_t^*, \Psi \rangle &= \langle \mu_0^*, \Psi \rangle + \int_0^t \int_E \left(\partial_s \Psi(v, s, \xi) + \partial_\xi \Psi(v, s, \xi) \cdot (-\partial_s \xi) \right) \mu_u^*(dv, ds, d\xi) du \\ &+ \int_0^t \left\{ \int_{\mathbb{R}^+ \times \mathcal{P}(E_m)} \beta \left[\Psi(0, s, \xi) - \Psi(1, s, \xi) \right] \mu_u^*(\{1\}, ds, d\xi) \right\} du \\ &+ \int_0^t \left\{ \int_E \partial_\xi \Psi(v, s, \xi) \cdot (\beta \delta_0 \otimes \xi^1 - \beta \delta_1 \otimes \xi^1) \mu_u^*(dv, ds, d\xi) \right\} du \\ &+ \int_0^t \left\{ \int_{\mathbb{R}^+ \times \mathcal{P}(E_m)} \alpha(I(\xi)) \left[\Psi(1, 0, \xi) - \Psi(0, s, \xi) \right] \mu_u^*(\{0\}, ds, d\xi) \right\} du \\ &+ \int_0^t \left\{ \int_E \partial_\xi \Psi(v, s, \xi) \cdot (\delta_1 \otimes \nu^1(\xi, \mu_u^*) - \delta_0 \otimes \nu^0(\xi, \mu_u^*)) \mu_u^*(dv, ds, d\xi) \right\} du. \end{aligned} \tag{4.23}$$

We deduce from this conjecture that the joint law of the two first components of ξ_t^* and μ_t^* would then be equal for all t .

Consequence 4.26. *Grant Conjecture 4.25. Then, for all $t \geq 0$,*

$$\xi_t^*(\cdot, \cdot, \mathbb{Z}) = \mu_t^*(\cdot, \cdot, \mathcal{P}(E_m)).$$

Proof. For all $\xi_0^* \in \text{supp}(\mu_0^*)$, we have $\mu_0^*(\cdot, \cdot, \mathcal{P}(E_m)) = \xi_0^*(\cdot, \cdot, \mathbb{Z})$. Then, we show that $(\mu_t^*(\cdot, \cdot, \mathcal{P}(E_m)))_{t \geq 0}$ and $(\xi_t^*(\cdot, \cdot, \mathbb{Z}))_{t \geq 0}$ have the same dynamics, which enables us to conclude the proof. First, in equation (4.23), taking a function Ψ which depends only on v and s , we obtain the dynamics of $(\mu_t^*(\cdot, \cdot, \mathcal{P}(E_m)))_{t \geq 0}$:

$$\begin{aligned} \langle \mu_t^*, \Psi \rangle &= \langle \mu_0^*, \Psi \rangle + \int_0^t \int_E \partial_s \Psi(v, s) \mu_u^*(dv, ds, d\xi) du \\ &\quad + \int_0^t \left\{ \int \beta [\Psi(0, s) - \Psi(1, s)] \mu_u^*(\{1\}, ds, d\xi) \right\} du \\ &\quad + \int_0^t \left\{ \int \alpha(I(\xi)) [\Psi(1, 0) - \Psi(0, s)] \mu_u^*(\{0\}, ds, d\xi) \right\} du. \end{aligned} \quad (4.24)$$

Second, taking Ψ depending only on ξ in equation (4.23), we obtain that

$$\begin{aligned} \langle \mu_t^*, \Psi \rangle - \langle \mu_0^*, \Psi \rangle &= \int_0^t \left\{ \int_E \partial_\xi \Psi(\xi) \cdot \left[\beta \delta_0 \otimes \xi^1 - \beta \delta_1 \otimes \xi^1 \right. \right. \\ &\quad \left. \left. + \delta_1 \otimes \nu^1(\xi, \mu_u^*) - \delta_0 \otimes \nu^0(\xi, \mu_u^*) - \partial_s \xi \right] \mu_u^*(dv, ds, d\xi) \right\} du. \end{aligned}$$

Working on the right hand side term, we obtain

$$\langle \mu_t^*, \Psi \rangle - \langle \mu_0^*, \Psi \rangle = \mathbb{E} [\Psi(\xi_t^*)] - \mathbb{E} [\Psi(\xi_0^*)] = \int_0^t \mathbb{E} \left[\partial_\xi \Psi(\xi_u^*) \cdot \left(\frac{d\xi_u^*}{du} \right) \right] du$$

and then deduce that

$$\begin{aligned} \forall \phi \in C_b^1(E_m) \text{ with bounded derivative,} \\ \langle \xi_t^*, \phi \rangle &= \langle \xi_0^*, \phi \rangle + \int_0^t \langle -\partial_s \xi_u^*, \phi \rangle du + \int_0^t \beta \langle \delta_0 \otimes \xi_u^{*1} - \delta_1 \otimes \xi_u^{*1}, \phi \rangle du \\ &\quad + \int_0^t \langle \delta_1 \otimes \nu^1(\xi_u^*, \mu_u^*) - \delta_0 \otimes \nu^0(\xi_u^*, \mu_u^*), \phi \rangle du. \end{aligned} \quad (4.25)$$

Evaluating this last equation for a function ϕ depending only on v and s gives us the dynamics of $\xi_t^*(\cdot, \cdot, \mathbb{Z})$. We obtain exactly the same equation as (4.24):

- using (4.10), we have for all $u \in [0, t]$, $\langle -\partial_s \xi_u^*, \phi \rangle = \langle \xi_u^*, \partial_s \phi \rangle$,
- as for all s in the support of $\xi \in \mathcal{P}(E_m)$, $\sum_{w \in \mathbb{Z}} \gamma_\xi(s, w) = 1$, we obtain that (see (4.18) and (4.19) for the definitions of ν^0 and ν^1) for all $A \in \mathcal{B}(\mathbb{R}^+)$:

$$\begin{aligned} \nu^0(\xi, \mu)(A, \mathbb{Z}) &= \int_{\mathcal{P}(E_m)} \int_A \alpha(I(\tilde{\xi})) \mu(\{0\}, ds, d\tilde{\xi}), \\ \nu^1(\xi, \mu)(A, \mathbb{Z}) &= \mathbb{1}_{\{0 \in A\}} \int_{\mathcal{P}(E_m)} \int_{\mathbb{R}^+} \alpha(I(\tilde{\xi})) \mu(\{0\}, ds, d\tilde{\xi}). \end{aligned}$$

□

The typical neuron dynamics

When μ_t^* admits a density in s , we use the abuse of notation $\mu_t^*(\{v\}, ds, d\xi) = \mu_t^*(\{v\}, s, d\xi)ds$.

Consequence 4.27. *Grant Conjecture 4.25 and Assumption 4.17. Then, we have*

- $dS_t^* = dt$.
- ξ_t^* admits a density in s that satisfies the following equations,

$$\begin{cases} \frac{d\xi_t^{*0}}{dt}(s, \{w\}) &= -\partial_s \xi_t^{*0}(s, \{w\}) + \beta \xi_t^{*1}(s, \{w\}) - \int_{\mathcal{P}(E_m)} \alpha(I(\xi')) \frac{\xi_t^{*0}(s, \{w\})}{\xi_t^{*0}(s, \mathbb{Z})} \mu_t^*(\{0\}, s, d\xi') \\ \xi_t^{*0}(0, \{w\}) &= 0, \end{cases}$$
- At rate $\beta \mathbb{1}_{\{V_{t^-}^* = 1\}}$, $(V_{t^-}^*, S_{t^-}^*, \xi_{t^-}^*)$ jumps to $(0, S_{t^-}^*, \xi_{t^-}^*)$.
- At rate $\alpha(I(\xi_{t^-}^*)) \mathbb{1}_{\{V_{t^-}^* = 0\}}$, $(V_{t^-}^*, S_{t^-}^*, \xi_{t^-}^*)$ jumps to $(1, 0, \xi_{t^-}^*)$.

Proof. Both the first two and last two points are clear with Conjecture 4.25. We detail the third point. First, from the regularity assumption on the density in s of μ_t^* , Consequence 4.26 tells us that ξ_t^* admits a density C^1 in s and satisfies (4.25). Using Remark 4.24, we thus obtain by integration by parts the following equations on the density functions $s \mapsto \xi_t^{*v}(s, \{w\})$:

$$\begin{aligned} \frac{d\xi_t^{*0}}{dt}(s, \{w\}) &= -\partial_s \xi_t^{*0}(s, \{w\}) - \delta_0(s) \xi_t^{*0}(0, \{w\}) + \beta \xi_t^{*1}(s, \{w\}) \\ &\quad - \int_{\mathcal{P}(E_m)} \alpha(I(\xi')) \frac{\xi_t^{*0}(s, \{w\})}{\xi_t^{*0}(s, \mathbb{Z})} \mu_t^*(\{0\}, s, d\xi'), \\ \frac{d\xi_t^{*1}}{dt}(s, \{w\}) &= -\partial_s \xi_t^{*1}(s, \{w\}) - \delta_0(s) \xi_t^{*1}(0, \{w\}) - \beta \xi_t^{*1}(s, \{w\}) \\ &\quad + \delta_0(s) \int_{\mathbb{R}^+ \times \mathcal{P}(E_m)} \alpha(I(\xi')) \frac{\xi_t^{*0}(s', \{w\})}{\xi_t^{*0}(s', \mathbb{Z})} \mu_t^*(\{0\}, ds', d\xi'). \end{aligned}$$

By integrating the previous equations on $s \in [0, \epsilon]$ and making ϵ tends to 0 we obtain the equations with conditions at boundaries $s = 0$ of the corollary. \square

4.3.4 The case with plasticity

We now deal with the synaptic weight jumps. As soon as a neuron spikes, several different synaptic weight jumps are possible. We denote by J_i^N the possible increments of the weights when the neuron i spikes

$$J_i^N = \{A = (a^{kl})_{1 \leq k, l \leq N} : a^{ii} \in \{-1, 0, 1\} \text{ and } \forall k, l \neq i, a^{il} \in \{0, 1\}, a^{ki} \in \{0, -1\}, a^{kl} = 0\}.$$

For all $\Delta \in J_i^N$, we denote by $P_{t,i}^\Delta$ the probability that the weight matrix be incremented by Δ conditionally to X_t^N . In order to ease its understanding, we express it with the weight matrix W_t^N rather than using the empirical measures $(\xi_t^{i,N})_{1 \leq i \leq N}$:

$$\begin{aligned} P_{t,i}^\Delta = & \left[\frac{\Delta^{ii}(1 + \Delta^{ii})}{2} p^+(S_{t^-}^{i,N}, W_{t^-}^{ii,N}) (1 - p^-(S_{t^-}^{i,N}, W_{t^-}^{ii,N})) \right. \\ & + (1 - |\Delta^{ii}|) \left(p^+(S_{t^-}^{i,N}, W_{t^-}^{ii,N}) p^-(S_{t^-}^{i,N}, W_{t^-}^{ii,N}) \right. \\ & \quad \left. \left. + (1 - p^+(S_{t^-}^{i,N}, W_{t^-}^{ii,N})) (1 - p^-(S_{t^-}^{i,N}, W_{t^-}^{ii,N})) \right) \right. \\ & \left. + \frac{\Delta^{ii}(\Delta^{ii} - 1)}{2} p^-(S_{t^-}^{i,N}, W_{t^-}^{ii,N}) (1 - p^+(S_{t^-}^{i,N}, W_{t^-}^{ii,N})) \right] \\ & \prod_{k,l \neq i} \left(\Delta^{il} p^+(S_{t^-}^{l,N}, W_{t^-}^{il,N}) + (1 - \Delta^{il}) (1 - p^+(S_{t^-}^{l,N}, W_{t^-}^{il,N})) \right) \\ & \left(-\Delta^{ki} p^-(S_{t^-}^{k,N}, W_{t^-}^{ki,N}) + (1 + \Delta^{ki}) (1 - p^-(S_{t^-}^{k,N}, W_{t^-}^{ki,N})) \right). \end{aligned}$$

After the spike of the neuron i at time t and assuming the weights are incremented of $\Delta \in J_i^N$ at this time, $\xi_{t^-}^{i,N}$ jumps to $\xi_{t,\Delta}^{i,N}$ and for all j , $\xi_{t^-}^{j,N}$ jumps to $\xi_{t,\Delta,i}^{j,N}$ such that

$$\begin{aligned} \xi_{t,\Delta}^{i,N} = & \xi_{t^-}^{i,N} + \frac{1}{N} \sum_{(0, S_{t^-}^{i,N}, \tilde{W}) \in \text{supp}(\xi_{t^-}^{i,N})} (\delta_{(1,0, \tilde{W} + \Delta^{ii})} - \delta_{(0, S_{t^-}^{i,N}, \tilde{W})}) \\ & + \sum_{l \neq i} \frac{1}{N} \sum_{(\tilde{V}, S_{t^-}^{l,N}, \tilde{W}) \in \text{supp}(\xi_{t^-}^{i,N})} (\delta_{(\tilde{V}, S_{t^-}^{l,N}, \tilde{W} + \Delta^{il})} - \delta_{(\tilde{V}, S_{t^-}^{l,N}, \tilde{W})}) \end{aligned} \quad (4.26)$$

$$\xi_{t,\Delta,i}^{j,N} = \xi_{t^-}^{j,N} + \frac{1}{N} \sum_{(0, S_{t^-}^{i,N}, \tilde{W}) \in \text{supp}(\xi_{t^-}^{i,N})} (\delta_{(1,0, \tilde{W} + \Delta^{ji})} - \delta_{(0, S_{t^-}^{i,N}, \tilde{W})}). \quad (4.27)$$

For all $\Psi \in C_b^{1,1}(E)$, we now have

$$\begin{aligned} \langle \mu_t^N, \Psi \rangle & \stackrel{\mathbb{E}}{=} \langle \mu_0^N, \Psi \rangle + \int_0^t \int_E \left(\partial_s \Psi(v, s, \xi) + \partial_\xi \Psi(v, s, \xi) \cdot (-\partial_s \xi) \right) \mu_u^N(dv, ds, d\xi) du \\ & + \int_0^t \sum_i \mathbb{1}_{\{V_{u^-}^{i,N} = 1\}} \frac{\beta}{N} \left\{ \left[\Psi\left(0, S_{u^-}^{i,N}, \xi_{u^-}^{i,N} + \eta^N(S_{u^-}^{i,N}, \xi_{u^-}^{i,N})\right) - \Psi\left(1, S_{u^-}^{i,N}, \xi_{u^-}^{i,N}\right) \right] \right. \\ & \quad \left. + \sum_{j \neq i} \left[\Psi\left(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u^-}^{j,N} + \eta^N(S_{u^-}^{i,N}, \xi_{u^-}^{j,N})\right) - \Psi\left(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u^-}^{j,N}\right) \right] \right\} du \\ & + \int_0^t \sum_i \mathbb{1}_{\{V_{u^-}^{i,N} = 0\}} \frac{\alpha(I_{u^-}^{i,N})}{N} \sum_{\Delta \in J_i^N} P_{u,i}^\Delta \left\{ \left[\Psi\left(1, 0, \xi_{u,\Delta}^{i,N}\right) - \Psi\left(0, S_{u^-}^{i,N}, \xi_{u^-}^{i,N}\right) \right] \right. \\ & \quad \left. + \sum_{j \neq i} \left[\Psi\left(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u,\Delta,i}^{j,N}\right) - \Psi\left(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u^-}^{j,N}\right) \right] \right\} du. \end{aligned}$$

We complete the sum (last term) by adding and removing the term

$$\mathbb{E} \left[\int_0^t \sum_i \mathbb{1}_{\{V_{u^-}^{i,N} = 0\}} \frac{\alpha(I_{u^-}^{i,N})}{N} \sum_{\Delta \in J_i^N} P_{u,i}^\Delta \left[\Psi\left(0, S_{u^-}^{i,N}, \xi_{u,\Delta,i}^{i,N}\right) - \Psi\left(0, S_{u^-}^{i,N}, \xi_{u^-}^{i,N}\right) \right] du \right].$$

We obtain

$$\begin{aligned}
\langle \mu_t^N, \Psi \rangle &\stackrel{\mathbb{E}}{=} \langle \mu_0^N, \Psi \rangle + \int_0^t \int_E \left(\partial_s \Psi(v, s, \xi) + \partial_\xi \Psi(v, s, \xi) \cdot (-\partial_s \xi) \right) \mu_u^N(dv, ds, d\xi) du \\
&+ \int_0^t \sum_i \mathbb{1}_{\{V_{u^-}^{i,N}=1\}} \frac{\beta}{N} \left\{ \left[\Psi\left(0, S_{u^-}^{i,N}, \xi_{u^-}^{i,N} + \eta^N(S_{u^-}^{i,N}, \xi_{u^-}^{i,N})\right) - \Psi\left(1, S_{u^-}^{i,N}, \xi_{u^-}^{i,N}\right) \right] \right. \\
&\quad \left. + \sum_{j \neq i} \left[\Psi\left(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u^-}^{j,N} + \eta^N(S_{u^-}^{i,N}, \xi_{u^-}^{j,N})\right) - \Psi\left(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u^-}^{j,N}\right) \right] \right\} du \\
&+ \int_0^t \sum_i \mathbb{1}_{\{V_{u^-}^{i,N}=0\}} \frac{\alpha(I_{u^-}^{i,N})}{N} \sum_{\Delta \in J_i^N} P_{u,i}^\Delta \left\{ \left[\Psi\left(1, 0, \xi_{u,\Delta}^{i,N}\right) - \Psi\left(0, S_{u^-}^{i,N}, \xi_{u,\Delta}^{i,N}\right) \right] \right. \\
&\quad \left. + \sum_j \left[\Psi\left(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u,\Delta}^{j,N}\right) - \Psi\left(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u^-}^{j,N}\right) \right] \right\} du
\end{aligned} \tag{4.28}$$

We then denote by

$$\textcircled{3} = \sum_i \mathbb{1}_{\{V_{u^-}^{i,N}=0\}} \frac{\alpha(I_{u^-}^{i,N})}{N} \sum_{\Delta \in J_i^N} P_{u,i}^\Delta \left[\Psi\left(1, 0, \xi_{u,\Delta}^{i,N}\right) - \Psi\left(0, S_{u^-}^{i,N}, \xi_{u,\Delta}^{i,N}\right) \right] \tag{4.29}$$

and

$$\textcircled{4} = \sum_i \mathbb{1}_{\{V_{u^-}^{i,N}=0\}} \frac{\alpha(I_{u^-}^{i,N})}{N} \sum_{\Delta \in J_i^N} P_{u,i}^\Delta \sum_j \left[\Psi\left(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u,\Delta}^{j,N}\right) - \Psi\left(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u^-}^{j,N}\right) \right].$$

The depression term

We first deal with the term $\textcircled{4}$ which describes the depression of the weights.

Proposition 4.28. *Then, for all $u \geq 0$,*

$$\textcircled{4} \stackrel{\mathbb{E}}{=} \int \partial_\xi \Psi(v, s, \xi) \cdot [\delta_1 \otimes \nu^{-,1}(s, \xi, \mu_u^N) - \delta_0 \otimes \nu^0(\xi, \mu_u^N)] \mu_u^N(dv, ds, d\xi) + \mathcal{O}(1).$$

where $\nu^{-,1}$ is defined in the proof, see (4.30).

Proof. First, by Lemma 4.2, we have almost surely for all i, j ,

$$\left\| \frac{1}{N} \sum_{(0, S_{u^-}^{i,N}, \tilde{W}) \in \text{supp}(\xi_{u^-}^{j,N})} (\delta_{(1,0, \tilde{W} + \Delta^{ji})} - \delta_{(0, S_{u^-}^{i,N}, \tilde{W})}) \right\|_{TV} \leq \frac{2}{N}.$$

Then, we use Fréchet derivative to obtain that

$$\begin{aligned}
&\Psi\left(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u,\Delta}^{j,N}\right) - \Psi\left(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u^-}^{j,N}\right) \\
&= \partial_\xi \Psi\left(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u^-}^{j,N}\right) \cdot \left(\frac{1}{N} \sum_{(0, S_{u^-}^{i,N}, \tilde{W}) \in \text{supp}(\xi_{u^-}^{j,N})} (\delta_{(1,0, \tilde{W} + \Delta^{ji})} - \delta_{(0, S_{u^-}^{i,N}, \tilde{W})}) \right) + \mathcal{O}\left(\frac{1}{N}\right).
\end{aligned}$$

We deduce by linearity that

$$\begin{aligned} \textcircled{4} &= \frac{1}{N} \sum_j \partial_\xi \Psi \left(V_{u^-}^{j,N}, S_{u^-}^{j,N}, \xi_{u^-}^{j,N} \right) \\ &\quad \underbrace{\left(\sum_i \mathbb{1}_{\{V_{u^-}^{i,N}=0\}} \frac{\alpha(I_{u^-}^{i,N})}{N} \sum_{(0, S_{u^-}^{i,N}, \tilde{W}) \in \text{supp}(\xi_{u^-}^{j,N})} \sum_{\Delta \in J_i^N} P_{u,i}^\Delta (\delta_{(1,0, \tilde{W} + \Delta^{jj})} - \delta_{(0, S_{u^-}^{i,N}, \tilde{W})}) \right)}_{\textcircled{4a}} + \mathcal{O}(1). \end{aligned}$$

But for any function f on $\{1, 0, -1\}$ we have for all $i \neq j$,

$$\begin{aligned} \sum_{\Delta \in J_i^N} P_{u,i}^\Delta f(\Delta^{jj}) &= \sum_{\Delta \in J_i^N, \Delta^{jj}=-1} P_{u,i}^\Delta f(-1) + \sum_{\Delta \in J_i^N, \Delta^{jj}=0} P_{u,i}^\Delta f(0) \\ &= p^-(S_{u^-}^{j,N}, W_{u^-}^{jj})f(-1) + (1 - p^-(S_{u^-}^{j,N}, W_{u^-}^{jj}))f(0), \end{aligned}$$

and for $i = j$,

$$\begin{aligned} &\sum_{\Delta \in J_j^N} P_{u,j}^\Delta f(\Delta^{jj}) \\ &= p^-(S_{u^-}^{j,N}, W_{u^-}^{jj})(1 - p^+(S_{u^-}^{j,N}, W_{u^-}^{jj}))f(-1) + p^+(S_{u^-}^{j,N}, W_{u^-}^{jj})(1 - p^-(S_{u^-}^{j,N}, W_{u^-}^{jj}))f(1) \\ &\quad + \left((1 - p^-(S_{u^-}^{j,N}, W_{u^-}^{jj}))(1 - p^+(S_{u^-}^{j,N}, W_{u^-}^{jj})) + p^+(S_{u^-}^{j,N}, W_{u^-}^{jj})p^-(S_{u^-}^{j,N}, W_{u^-}^{jj}) \right) f(0). \end{aligned}$$

In $\textcircled{4a}$, the term for $i = j$ is of order $\frac{1}{N}$. Hence, we obtain that

$$\begin{aligned} &\textcircled{4a} + \mathcal{O}\left(\frac{1}{N}\right) \\ &= \sum_i \mathbb{1}_{\{V_{u^-}^{i,N}=0\}} \frac{\alpha(I_{u^-}^{i,N})}{N} \sum_{(0, S_{u^-}^{i,N}, \tilde{W}) \in \text{supp}(\xi_{u^-}^{j,N})} \left[p^-(S_{u^-}^{j,N}, \tilde{W}) (\delta_{(1,0, \tilde{W}-1)} - \delta_{(0, S_{u^-}^{i,N}, \tilde{W})}) \right. \\ &\quad \left. + (1 - p^-(S_{u^-}^{j,N}, \tilde{W})) (\delta_{(1,0, \tilde{W})} - \delta_{(0, S_{u^-}^{i,N}, \tilde{W})}) \right] \\ &= \sum_i \mathbb{1}_{\{V_{u^-}^{i,N}=0\}} \frac{\alpha(I_{u^-}^{i,N})}{N} \sum_{\tilde{W} \in \mathbb{Z}} \mathbb{1}_{\{(0, S_{u^-}^{i,N}, \tilde{W}) \in \text{supp}(\xi_{u^-}^{j,N})\}} \left[p^-(S_{u^-}^{j,N}, \tilde{W}) (\delta_{(1,0, \tilde{W}-1)} - \delta_{(0, S_{u^-}^{i,N}, \tilde{W})}) \right. \\ &\quad \left. + (1 - p^-(S_{u^-}^{j,N}, \tilde{W})) (\delta_{(1,0, \tilde{W})} - \delta_{(0, S_{u^-}^{i,N}, \tilde{W})}) \right] \\ &= \sum_i \mathbb{1}_{\{V_{u^-}^{i,N}=0\}} \frac{\alpha(I_{u^-}^{i,N})}{N} \sum_{\tilde{W} \in \mathbb{Z}} \gamma_{\xi_{u^-}^{j,N}}(S_{u^-}^{i,N}, \tilde{W}) \left[p^-(S_{u^-}^{j,N}, \tilde{W}) (\delta_{(1,0, \tilde{W}-1)} - \delta_{(0, S_{u^-}^{i,N}, \tilde{W})}) \right. \\ &\quad \left. + (1 - p^-(S_{u^-}^{j,N}, \tilde{W})) (\delta_{(1,0, \tilde{W})} - \delta_{(0, S_{u^-}^{i,N}, \tilde{W})}) \right]. \end{aligned}$$

Hence, we deduce that

$$\begin{aligned} \textcircled{4a} &= \sum_{\tilde{W} \in \mathbb{Z}} \int_{\mathcal{P}(E_m)} \int_{\mathbb{R}^+} \alpha(I(\tilde{\xi})) \gamma_{\xi_{u^-}^{j,N}}(\tilde{s}, \tilde{W}) \left[p^-(S_{u^-}^{j,N}, \tilde{W}) (\delta_{(1,0, \tilde{W}-1)} - \delta_{(0, \tilde{s}, \tilde{W})}) \right. \\ &\quad \left. + (1 - p^-(S_{u^-}^{j,N}, \tilde{W})) (\delta_{(1,0, \tilde{W})} - \delta_{(0, \tilde{s}, \tilde{W})}) \right] \mu_{u^-}^N(\{0\}, d\tilde{s}, d\tilde{\xi}) + \mathcal{O}\left(\frac{1}{N}\right). \end{aligned}$$

This incites us to define the following measure. For all triple $(s, \xi, \mu) \in \mathbb{R}^+ \times \mathcal{P}(E_m) \times \mathcal{P}(E)$, we denote by $\nu^{-,1}(s, \xi, \mu)$ the measure on $\mathbb{R}^+ \times \mathbb{Z}$ such that for all $(A, \{w\}) \in \mathcal{B}(\mathbb{R}^+) \times \mathbb{Z}$,

$$\begin{aligned} \nu^{-,1}(s, \xi, \mu)(A, \{w\}) &\stackrel{def.}{=} \mathbb{1}_{\{0 \in A\}} \int_{\mathcal{P}(E_m)} \alpha(I(\tilde{\xi})) p^-(s, w+1) \gamma_\xi(\tilde{s}, w+1) \mu(\{0\}, d\tilde{s}, d\tilde{\xi}) \\ &\quad + \mathbb{1}_{\{0 \in A\}} \int_{\mathcal{P}(E_m)} \alpha(I(\tilde{\xi})) (1 - p^-(s, w)) \gamma_\xi(\tilde{s}, w) \mu(\{0\}, d\tilde{s}, d\tilde{\xi}), \end{aligned} \quad (4.30)$$

where we used the definition of γ_ξ given by (4.17). Thereby, using the function ν^0 defined in (4.18), we get

$$\textcircled{4a} = \delta_1 \otimes \nu^{-,1}(S_{u^-}^{j,N}, \xi_{u^-}^{j,N}, \mu_{u^-}^N) \delta_0 \otimes \nu^0(\xi_{u^-}^{j,N}, \mu_{u^-}^N) + \mathcal{O}(1).$$

We finally obtain that

$$\textcircled{4} \stackrel{\mathbb{E}}{=} \int \partial_\xi \Psi(v, s, \xi) \cdot [\delta_1 \otimes \nu^{-,1}(s, \xi, \mu_u^N) - \delta_0 \otimes \nu^0(\xi, \mu_u^N)] \mu_u^N(dv, ds, d\xi) + \mathcal{O}(1).$$

It ends the proof. \square

Under Assumption 4.13, it is reasonable to expect that, as N tends to infinity, $\mathbb{E}\textcircled{4}$ converges to

$$\int \partial_\xi \Psi(v, s, \xi) \cdot [\delta_1 \otimes \nu^{-,1}(s, \xi, \mu_u^*) - \delta_0 \otimes \nu^0(\xi, \mu_u^*)] \mu_u^*(dv, ds, d\xi).$$

As noted in the remark given just after Proposition 4.23, this convergence does not hold a priori because the functions

$$(v, s, \xi) \mapsto \partial_\xi \Psi(v, s, \xi) \cdot [\delta_1 \otimes \nu^{-,1}(s, \xi, \mu_u^N) - \delta_0 \otimes \nu^0(\xi, \mu_u^N)] \quad (4.31)$$

are not a priori continuous. We expect to show this convergence using a sequence (indexed by N) of continuous functions both getting closer and closer to (4.31) and converging to

$$(v, s, \xi) \mapsto \partial_\xi \Psi(v, s, \xi) \cdot [\delta_1 \otimes \nu^{-,1}(s, \xi, \mu_u^*) - \delta_0 \otimes \nu^0(\xi, \mu_u^*)].$$

The potentiation term

Now, for the term $\textcircled{3}$ describing potentiation, see (4.29), we need to assume that Ψ is linear in its third variable ξ .

Proposition 4.29. *Under Assumptions 4.1, 4.13 and 4.16 we find that for all $u \geq 0$, for all $\Psi \in C_b^{1,1}(E)$ linear in its third variable ξ ,*

$$\lim_{N \rightarrow \infty} \textcircled{3} \stackrel{\mathbb{E}}{=} \int_E \alpha(I(\xi)) \left[\Psi(1, 0, \nu^+(\xi)) - \Psi(0, s, \xi) \right] \mu_u^*(\{0\}, ds, d\xi),$$

where ν^+ is defined in the proof, see (4.32).

Proof. By linearity of Ψ in its third variable, we obtain that

$$\textcircled{3} = \sum_i \mathbb{1}_{\{V_{u^-}^{i,N}=0\}} \frac{\alpha(I_{u^-}^{i,N})}{N} \sum_{\Delta \in J_i^N} P_{u,i}^\Delta \left[\Psi \left(1, 0, \sum_{\Delta \in J_i^N} P_{u,i}^\Delta \xi_{u,\Delta}^{i,N} \right) - \Psi \left(0, S_{u^-}^{i,N}, \sum_{\Delta \in J_i^N} P_{u,i}^\Delta \xi_{u,\Delta}^{i,N} \right) \right].$$

and then by definition of $\xi_{u,\Delta}^{i,N}$, see (4.26), we have

$$\begin{aligned} \sum_{\Delta \in J_i^N} P_{u,i}^\Delta \xi_{u,\Delta}^{i,N} &= \xi_{u^-}^{i,N} + \frac{1}{N} \sum_{(0, S_{u^-}^{i,N}, \tilde{W}) \in \text{supp}(\xi_{u^-}^{i,N})} \sum_{\Delta \in J_i^N} P_{u,i}^\Delta (\delta_{(1,0, \tilde{W} + \Delta^{ii})} - \delta_{(0, S_{u^-}^{i,N}, \tilde{W})}) \\ &\quad + \sum_{l \neq i} \frac{1}{N} \sum_{(\tilde{V}, S_{u^-}^{l,N}, \tilde{W}) \in \text{supp}(\xi_{u^-}^{i,N})} \sum_{\Delta \in J_i^N} P_{u,i}^\Delta (\delta_{(\tilde{V}, S_{u^-}^{l,N}, \tilde{W} + \Delta^{il})} - \delta_{(\tilde{V}, S_{u^-}^{l,N}, \tilde{W})}). \end{aligned}$$

Noting the second term of the right hand side of the last equation is of order $\frac{1}{N}$, we deduce that

$$\begin{aligned} &\sum_{\Delta \in J_i^N} P_{u,i}^\Delta \xi_{u,\Delta}^{i,N} + \mathcal{O}\left(\frac{1}{N}\right) \\ &= \xi_{u^-}^{i,N} + \sum_{l \neq i} \frac{1}{N} \sum_{(\tilde{V}, S_{u^-}^{l,N}, \tilde{W}) \in \text{supp}(\xi_{u^-}^{i,N})} p^+(S_{u^-}^{l,N}, \tilde{W}) (\delta_{(\tilde{V}, S_{u^-}^{l,N}, \tilde{W} + 1)} - \delta_{(\tilde{V}, S_{u^-}^{l,N}, \tilde{W})}) \\ &= \frac{1}{N} \sum_{(\tilde{V}, \tilde{S}, \tilde{W}) \in \text{supp}(\xi_{u^-}^{i,N})} p^+(\tilde{S}, \tilde{W}) \delta_{(\tilde{V}, \tilde{S}, \tilde{W} + 1)} + (1 - p^+(\tilde{S}, \tilde{W})) \delta_{(\tilde{V}, \tilde{S}, \tilde{W})} \\ &= \nu^+(\xi_{u^-}^{i,N}) \end{aligned}$$

where

$$\nu^+(\xi)(\{v\}, A, \{w\}) = \int_A p^+(s, w-1) \xi(\{v\}, ds, \{w-1\}) + \int_A (1 - p^+(s, w)) \xi(\{v\}, ds, \{w\}). \quad (4.32)$$

But p^+ takes values in $[0, 1]$ so using the triangular inequality, $\xi \mapsto \nu^+(\xi)$ is continuous. We can then conclude using Assumption 4.13. \square

Justifying Conjecture 4.18

We are now in position to detail why Conjecture 4.18 is reasonable. Under Assumptions 4.1, 4.13 and 4.16, by making N tends to infinity in equation (4.28) and using the result of Proposition 4.28, we should get the following dynamics of μ_t^* :

- $\mathcal{L}(X_0^*) = \lim_{N \rightarrow \infty} \mathcal{L}(X_0^{1,N}) = \mu_0^*$ and in particular S_0^* is distributed by the ρ_0 law which is absolutely continuous with respect to the Lebesgue measure,

- using Notation 4.14 we have, for all $t \geq 0$, for all $\Psi \in C_b^{1,1}(E)$,

$$\begin{aligned}
\langle \mu_t^*, \Psi \rangle &= \langle \mu_0^*, \Psi \rangle + \int_0^t \int_E \left(\partial_s \Psi(v, s, \xi) + \partial_\xi \Psi(v, s, \xi) \cdot (-\partial_s \xi) \right) \mu_u^*(dv, ds, d\xi) du \\
&\quad + \int_0^t \left(\int_{\mathbb{R}^+ \times \mathcal{P}(E_m)} \beta \left[\Psi(0, s, \xi) - \Psi(1, s, \xi) \right] \mu_u^*({1}, ds, d\xi) \right) du \\
&\quad + \int_0^t \left\{ \int_E \partial_\xi \Psi(v, s, \xi) \cdot (\beta \delta_0 \otimes \xi^1 - \beta \delta_1 \otimes \xi^1) \mu_u^*(dv, ds, d\xi) \right\} du \\
&\quad + \int_0^t \left(\int_{\mathbb{R}^+ \times \mathcal{P}(E_m)} \alpha(I(\xi)) \left[\Psi(1, 0, \nu^+(\xi)) - \Psi(0, s, \xi) \right] \mu_u^*({0}, ds, d\xi) \right) du \\
&\quad + \int_0^t \left(\int_E \partial_\xi \Psi(v, s, \xi) \cdot [\delta_1 \otimes \nu^{-,1}(s, \xi, \mu_u^*) - \delta_0 \otimes \nu^0(\xi, \mu_u^*)] \mu_u^*(dv, ds, d\xi) \right) du,
\end{aligned} \tag{4.33}$$

where $\nu^{-,1}$ is defined in (4.30) and ν^0 is defined in (4.18).

With the same arguments as the one used in Consequence 4.26, one has that for all $t \geq 0$,

$$\xi_t^*(\cdot, \cdot, \mathbb{Z}) = \mu_t^*(\cdot, \cdot, \mathcal{P}(E_m)).$$

Then, from Conjecture 4.18, μ_t^* is assumed to admit a density C^1 in s . Therefore, the last equation tells us that ξ_t^* admits a density C^1 in s . Hence, with the same arguments as in Consequence 4.26 but with equation (4.33), we have:

$$\frac{d}{dt} \xi_t^* = -\partial_s \xi_t^* + \beta \left(\delta_0 \otimes \xi_t^{*1} - \delta_1 \otimes \xi_t^{*1} \right) + \delta_1 \otimes \nu^{-,1}(S_t^*, \xi_t^*, \mu_t^*) - \delta_0 \otimes \nu^0(\xi_t^*, \mu_t^*).$$

From this equation and Remark 4.24, we thus obtain the following equations on the density functions $s \mapsto \xi_t^{*v}(s, \{w\})$:

$$\begin{aligned}
\frac{d\xi_t^{*0}}{dt}(s, \{w\}) &= -\partial_s \xi_t^{*0}(s, \{w\}) - \delta_0(s) \xi_t^{*0}(0, \{w\}) + \beta \xi_t^{*1}(s, \{w\}) \\
&\quad - \int_{\mathcal{P}(E_m)} \alpha(I(\xi')) \frac{\xi_t^{*0}(s, \{w\})}{\xi_t^{*0}(s, \mathbb{Z})} \mu_t^*({0}, s, d\xi'),
\end{aligned}$$

$$\begin{aligned}
\frac{d\xi_t^{*1}}{dt}(s, \{w\}) &= -\partial_s \xi_t^{*1}(s, \{w\}) - \delta_0(s) \xi_t^{*1}(0, \{w\}) - \beta \xi_t^{*1}(s, \{w\}) \\
&\quad + \delta_0(s) \int_{\mathbb{R}^+ \times \mathcal{P}(E_m)} \frac{\alpha(I(\xi'))}{\xi_t^{*0}(s', \mathbb{Z})} \left[p^-(S_t^*, w) \xi_t^{*0}(s', \{w+1\}) \right. \\
&\quad \left. + (1 - p^-(S_t^*, w)) \xi_t^{*0}(s', \{w\}) \right] \mu_t^*({0}, ds', d\xi').
\end{aligned}$$

By integrating the previous equations on $s \in [0, \epsilon]$ and making ϵ tends to 0, we obtain the equations with conditions at boundaries $s = 0$ of the conjecture.

4.4 Numerical comparison with the neural network

We simulate the stochastic STDP model. In particular, we illustrate our results by comparing the simulation of the finite size neural network versus the simulation of the limit system.

We simulate N neurons $X_t^{1,*}, \dots, X_t^{N,*}$ having the same dynamics as X_t^* , except that instead of μ_t^* , we use $\mu_t^{*,N} = \frac{1}{N} \sum_i \delta_{X_t^{i,*}}$. Hence, we do not need to compute μ_t^* . For instance, we use the $\xi_t^{i,*}$ to compute the $I_t^{i,*} = I(\xi_t^{i,*})$. For more details on the code, see Appendix A.4. The results obtained with these simulations are compared with those obtained with the simulation of the finite size neural network: $(V_t^{i,N}, S_t^{i,N}, W_t^{ij,N})_{1 \leq i, j \leq N, t \geq 0}$. We compare the expectations over time of the synaptic weights, of the time since the last spikes and of the potentials, see Figures 4.2 and 4.3. Moreover, we compare, at the end of the simulation, the distributions of the intensities of the incoming currents onto the neurons (Figure 4.3), the distributions of the time since the last spikes and the one of the potentials (Figure 4.1).

We use a sigmoid for the function α and the classical STDP curves for p^+ and p^- ,

$$\alpha(x) = \frac{\alpha_M - \alpha_m}{1 + e^{\sigma(\theta - x)}} + \alpha_m \quad \text{and} \quad p^{+/-}(s, w) = A_{+/-} e^{-\frac{s}{\tau_{+/-}}} \mathbb{1}_{[w_{min}, w_{max}]}(w),$$

with the following parameters:

$$N = 5000, \alpha_m = 0.05, \alpha_M = \beta = 1, \tau_+ = 1.5 \text{ ms}, \tau_- = 2 \text{ ms}, A_+ = 0.8, A_- = 0.6, \\ dt = 0.05 \text{ ms}, \sigma = 1.5, \theta = 0, w_{max} = 10, w_{min} = -10.$$

The simulations show a very good match between the initial network and the limit one.

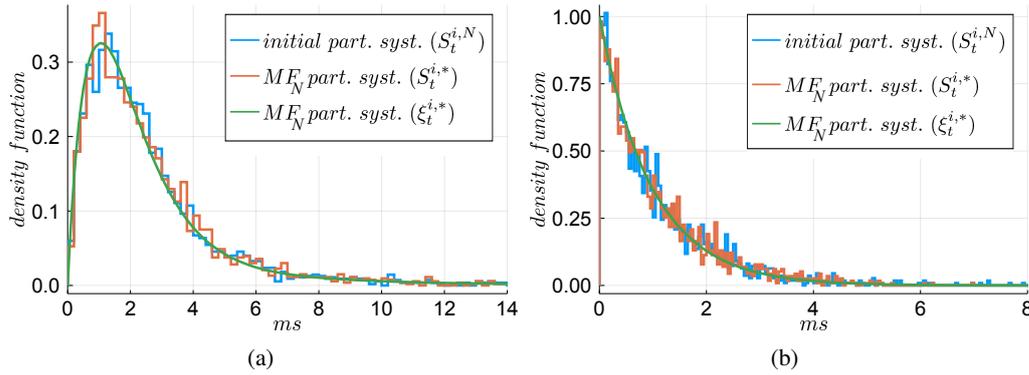


Figure 4.1: Distributions of the time from last spikes for neurons in state 0 in (4.1a) and for neurons in state 1 in (4.1b), at final time $t = 500 \text{ ms}$.

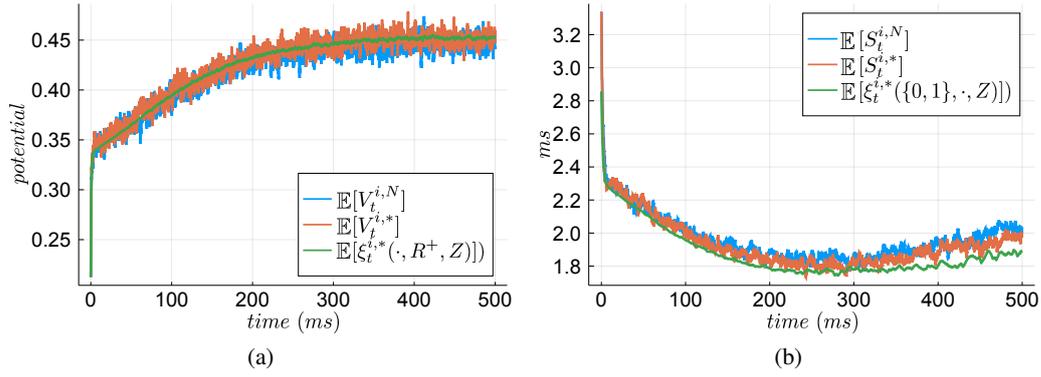


Figure 4.2: Expectation over time of the potentials (4.2a) and of the time from last spikes (4.2b)

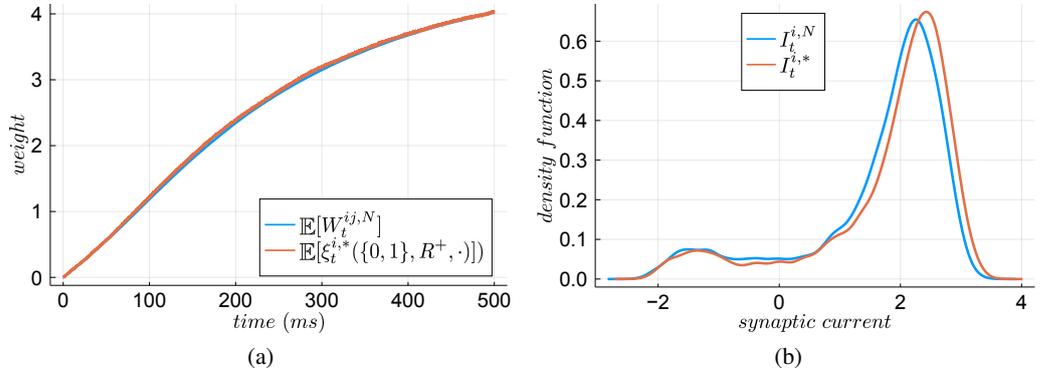


Figure 4.3: (4.3a): Expectation over time of the weights. (4.3b): Distribution of the synaptic currents onto the neurons.

4.5 Discussion

In this chapter, we introduced the averaging parameter $\frac{1}{N}$ in the synaptic currents. Because of the dependence between the neurons, in the large N asymptotic, the expectation of these currents is not sufficient: we need their law. It can be obtained from the law of the presynaptic weights. Considering the system composed of *new* variables (potentials, times from last spike, vector of the incoming weights) and its empirical measure, we can find a closed system of equations. Under the assumption that this empirical law converges to a measure as N tends to infinity, we conjecture the dynamics of any limit point. The typical neuron dynamics is solution to a McKean-Vlasov SDE on the space $\{0, 1\} \times \mathbb{R}^+ \times \mathcal{P}(E_m)$ where $\mathcal{P}(E_m)$ is the space of probability measures on $\{0, 1\} \times \mathbb{R}^+ \times \mathbb{Z}$. We illustrated this limit dynamics with simulations by comparing the finite size neural network to the mean field limit system. There is a good match between the two.

To the best of our knowledge, this mean field analysis is the first one done on a network of interacting neurons with plasticity. It opens the door to new mathematical questions. After showing our conjecture holds true, the next steps are: show that the limit system has a unique solution and then show the tightness of the empirical measures to conclude on their convergence in law to a unique deterministic limit measure. Finally, studying the limit system would give insight in the model and a particularly interesting study would be its long time behaviour. Improving the actual

code for simulation is an important step in order to help the theoretical study.

Chapter 5

Conclusion and Perspectives

In this thesis, we propose the mathematical analysis of memory lifetime in a model of neural network with plasticity (Chapter 2) as well as the long time behaviour (Chapter 3) and the mean field approximation (Chapter 4) of a new model of STDP. In both models, the synaptic weight changes are stochastic. The mathematical analysis of plastic neural network models using this kind of randomness (in plasticity rule) is quite new in this field and this is the main objective of this thesis. We summarise our results on both models and give some perspectives in continuity with our work.

Chapter 2

Chapter 2 is devoted to the mathematical analysis of the model proposed by Amit and Fusi in [5]. In particular, we studied the synaptic current distribution dynamics of the neurons selective or not to a given signal. The progressive overlapping of these two distributions is responsible for the forgetting of the signal. Using a Chernoff type inequality, we can control this overlap and then the probability that the neuron misevaluates the signal.

In order to go beyond our results, it would be interesting to understand the structure created by the learning of a signal on the entire network. Indeed, our analysis focuses only on one neuron and its incoming synaptic weights, the natural next step is to study the N neurons case and thus the N^2 synaptic weights. This is difficult in part because of the correlations. We think that using the Kronecker product, such as in Appendix A.2, is one of the possible key of this work. Moreover, we did not answer an interesting question which is: what is the scaling of f_N (the coding level of the signals) in N that maximises the time to forget the signal? Also, does it exist a better estimator than the simple threshold one? Finally, the main advantage of this model is its simplicity allowing its mathematical analysis. This simplicity is obtained at the expense of its reality. We give some ideas to remedy to this lack of biological realism. For example, the neural dynamics is imposed by the external signals whereas a more interesting model would take into account a more realistic interacting neuron dynamics. This would for example enables the memories to be *replayed* (without the need of any external signal) in order to maintain them. For instance, a time continuous neural dynamics would open many possibilities of study: changing the synaptic plasticity rules, proposing a new readout procedure based on richer information extracted from the past activity of the neurons. Poisson and Hawkes processes are good candidates in order to model the neural dynamics, see [30] for a presentation of Hawkes processes used in neuroscience. Two other changes of the model can be interesting to study: add a dependence between the external signals and add some neural layers modelling for example how the external signal leads to a neural representation of it in order to store it.

Chapters 3 and 4

In Chapters 3 and 4, we introduced a new stochastic STDP model that we implemented in a well known neural network model. This model is more complex than the one of Chapter 2. The neural network is composed of N binary probabilistic Wilson-Cowan neurons and discrete synaptic weights following a stochastic STDP rule. We analysed this model under two different asymptotic approaches: we first assumed that the plasticity is infinitely slow compared to the neural dynamics and then, we assumed that the number of neurons tends to infinity (changing a bit the model by dividing by N the weights in the synaptic current formula). The natural next step is to combined the two in order to obtain an even simpler approximation of the model. We did not do it in this thesis but it is not obvious that the two limits permute.

In Chapter 3, we showed the convergence of the slow process (the synaptic weights) to a limit process of which we have studied the long time behaviour. We proposed two techniques leading to conditions for positive recurrence or transience of the limit chain. The first one consists in coupling the limit process with two other ones. These two processes have independent components with similar dynamics that are determined by the extremal values of the limit process jumping rates. We think that the processes used for the bounding in this method can be improved to get much better bounds. The second uses of the Laplace transform of this invariant measure. In particular, we show how to compute this transform for any N using the Laplace transforms of the marginals of the invariant measure. Inverting the Laplace transform in two dimension is already hard in our case because the formulas are not easy to handle with. We used the Fokker-Planck equation to obtain it in the Appendix A.3.

In the continuity of this work, it would be very interesting to test the learning capacity of our model and also to compare it with real biological data. Also, we could simplify a bit this model by deleting the potentials and modelling the synaptic current using the weights W and the time from the last spikes S instead of the potentials V .

In Chapter 4, after defining *new* variables describing the neural network (from the initial description) adapted to the mean field analysis and assuming that its empirical law converges to a certain deterministic law in the large N asymptotic, we conjecture the McKean-Vlasov limit dynamics of this law. This conjecture is based on computations done on the dynamics of the empirical measure. We did not show yet rigorously the convergence of some term of this dynamics. The conjectured dynamics is particularly hard to handle with as it deals with laws on a space of measures. We nevertheless managed to compare the finite size neural network to the limit system using simulations and obtained a good match.

The natural questions arising from this chapter are multiple. First, the existence and uniqueness of the limit dynamics has to be studied. Then, the convergence as well as the propagation of chaos have to be shown. Finally, analysing the limit system is a challenging step maybe enabling us to understand a part of the weight structure formation.

Chapter A

Appendix

A.1 Definitions

A.1.1 Convergence in law (or weakly)

Definition A.1. (See for instance equation (10) in [19]) Let X be a metric space and \mathcal{X} its Borel sets and consider probability measures μ^N and μ on \mathcal{X} . Then, μ^N converges in law (or weakly) to μ if and only if

$$\forall f \in C_b(X), \quad \int_X f(x) \mu^N(dx) \rightarrow_{N \rightarrow \infty} \int_X f(x) \mu(dx).$$

A.1.2 Variation of signed measures

We first give some definitions on the space of signed measures on a measurable space, see [4, Thm 1.6 and Rk 1.7] for more details.

Definition A.2. Let (X, \mathcal{X}) be a measurable space equipped with a signed measure η . One associates to η the (unique) pair of positive measures η^+ and η^- defined for any $A \in \mathcal{X}$ by

$$\begin{aligned} \eta^+(A) &= \sup\{\eta(B), B \in \mathcal{X}, B \subset A\} \\ \eta^-(A) &= \sup\{-\eta(B), B \in \mathcal{X}, B \subset A\}. \end{aligned}$$

The Jordan decomposition of the signed measure η writes

$$\eta(A) = \eta^+(A) - \eta^-(A).$$

The variation $|\eta|$ of the signed measure η is defined, for all $A \in \mathcal{X}$, by

$$|\eta|(A) \stackrel{\text{def.}}{=} \eta^+(A) + \eta^-(A).$$

We deduce the total variation of η as

$$\|\eta\|_{TV} \stackrel{\text{def.}}{=} |\eta|(X) = \eta^+(X) + \eta^-(X).$$

The space of signed measures embedded with the norm $\|\cdot\|_{TV}$ is a Banach space.

A.1.3 Around Poisson

The first part of this section is taken from the book [62, Sec 4.2.1].

We first introduce non-degenerate point processes.

Definition A.3. (*Point process*) A point process on \mathbb{R}^+ with an infinite number of (strictly) positive jump instants without accumulation point, is a process $(N_t)_{t \geq 0}$ with values in \mathbb{N} , vanishing at 0, non-decreasing, right continuous, with unit jumps, and with infinite limit: for $0 \leq s \leq t < \infty$,

$$0 = N_0 \leq N_s \leq N_t = N_{t+}, \quad N_t - N_{t-} \in \{0, 1\}, \quad \lim_{t \rightarrow \infty} N_t = \infty.$$

This kind of random object has two other equivalent descriptions, to which it makes implicit reference. We give the jump instants description, the other one being the inter-arrivals description. Let $(T_n)_{n \geq 1}$ be the sequence of positive random variables increasing to infinity,

$$0 < T_1 < T_2 < \dots < \infty, \quad \lim_{n \rightarrow \infty} T_n = \infty,$$

such that the T_n are the jump instants of $(N_t)_{t \geq 0}$. Then $(N_t)_{t \geq 0}$ is given by their counting process:

$$N_t = \sum_{n \geq 1} \mathbb{1}_{\{T_n \leq t\}}.$$

What follows is taken from the book [32, Sec 2.1]. An (inhomogeneous) Poisson process of time-varying rate $\lambda_t > 0$ (λ is as regular as needed) is a point process satisfying the two following conditions. The number of points within disjoint intervals are independent and in each finite interval $(a_i, b_i]$, this number has a Poisson distribution of intensity $\int_{a_i}^{b_i} \lambda_x dx$. Formally,

Definition A.4. $(N_t)_{t \geq 0}$ is said to be a Poisson process of time-varying rate $\lambda_t > 0$ if it is a point process such that

1. for all $k \geq 1$, $t_0 = 0 < t_1 < \dots < t_k$, the random variables $(N_{t_k} - N_{t_{k-1}}), \dots, (N_{t_1} - N_0)$ are independent,
2. in each finite interval $(a_i, b_i]$, the number of points has a Poisson distribution of intensity $\int_{a_i}^{b_i} \lambda_x dx$: for all $k \in \mathbb{N}$,

$$\mathbb{P}(N_{b_i} - N_{a_i} = k) = \frac{(\int_{a_i}^{b_i} \lambda_x dx)^k}{k!} e^{-\int_{a_i}^{b_i} \lambda_x dx}.$$

The Poisson neuron is a spiking neuron model where the spiking times of a neuron is obtained from the jump instants of a Poisson process.

We now generalise this definition by giving the definition of a Poisson random measure. This definition is taken from the book [12, Def A.1].

Definition A.5. Let (E, \mathcal{E}) be a measurable space and μ a σ -finite measure on this space. A (homogeneous) Poisson point measure N with intensity $\mu(dh)dt$ on $\mathbb{R}^+ \times E$ is a $(\mathbb{R}^+ \times E, \mathcal{B}(\mathbb{R}^+) \otimes \mathcal{E})$ -random measure defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ which satisfies the following properties:

- N is a counting measure: $\forall A \in \mathcal{B}(\mathbb{R}^+) \otimes \mathcal{E}, \forall \omega \in \Omega, N(\omega, A) \in \mathbb{N} \cup \{+\infty\}$.

- $\forall \omega \in \Omega, N(\omega, \{0\} \times E) = 0$: no jump at time 0.
- $\forall A \in \mathcal{B}(\mathbb{R}^+) \otimes \mathcal{E}, \mathbb{E}N(A) = \nu(A)$, where $\nu(dt, dh) = \mu(dh)dt$.
- If A and B are disjoint in $\mathcal{B}(\mathbb{R}^+) \otimes \mathcal{E}$ and if $\nu(A) < +\infty, \nu(B) < +\infty$, then the random variables $N(A)$ and $N(B)$ are independent.

A.1.4 Fréchet differentiation

We recall the notion of Fréchet differentiation. We use the definition given in [66, p.6].

Definition A.6. Let F and G be two Banach spaces embedded with the norms $\|\cdot\|_F$ and $\|\cdot\|_G$. Let D be an open set of F . We say that $\varphi : D \rightarrow G$ is Fréchet differentiable in $y \in D$ if there exists a linear bounded operator $L(y) : h \in F \mapsto L(y)h \in G$ such that for all $h \in F$ satisfying $y + h \in D$, we have

$$\|\varphi(y + h) - \varphi(y) - L(y)h\|_G \leq \rho(\|h\|_F, y),$$

where $\rho(\|h\|_F, y)$ satisfies $\frac{\rho(\|h\|_F, y)}{\|h\|_F} \rightarrow 0$ when $\|h\|_F \rightarrow 0$. We call $L(y)$ the Fréchet derivative of φ in y and $L(y)h$ the Fréchet differential of φ in y in the direction h .

We say that φ is Fréchet differentiable on D if for all y in the interior of F , φ is Fréchet differentiable in y .

A.1.5 Semigroup Theory and Probability Theory

This section is taken from the book [62, Sec 6.1].

A transition kernel P is a measurable mapping

$$x \in E \mapsto P(x, \cdot) = P(x, dy) \in \mathcal{P}(E).$$

A transition kernel $P = (P(x, dy))_{x \in E}$ can be identified with the operator $f \mapsto Pf$ on L^∞ defined by

$$\forall x \in E, \quad Pf(x) = \int_E f(y)P(x, dy).$$

One can also be interested in the adjoint of P , denoted P^* defined on the dual $\mathcal{M}(E)$ of L^∞ by

$$\forall \mu \in \mathcal{M}(E), \forall f \in L^\infty, \quad P^* \mu f = \mu Pf = \int_E f(y)P(s, dy)\mu(dx).$$

From this definition, one can define a Markov process as follows.

Definition A.7. A process $(X_t)_{t \geq 0}$ with values in E is a Markov process if there exists a family $(P_t)_{t \geq 0}$ of Markovian kernels $P_t = (P_t(x, dy))_{x \in E}$ such that, for all $s, t \geq 0$ and $A \in \mathcal{B}(E)$,

$$\mathbb{P}(X_{s+t} \in A | (X_\theta)_{0 \leq \theta \leq s}) = P_t(X_s, A).$$

The P_t are called the transition kernels of the process. In particular, for f bounded we have

$$\mathbb{E}[f(X_t) | X_0 = x] = P_t f(x)$$

and $P_t(x, dy)$ is the conditional law of X_t given $X_0 = x$. The Chapman-Kolmogorov equation satisfied by the transition kernels P_t leads to the semi-group property of $(P_t)_{t \geq 0}$

$$\forall t, s \geq 0, \quad P_{t+s} = P_t P_s$$

This semi-group is characterized by its left derivative at 0.

Definition A.8. *The infinitesimal generator of the Markov process, or of its semigroup, is the operator defined by $\mathcal{A} : f \in D(\mathcal{A}) \subset L^\infty \mapsto \mathcal{A}f \in L^\infty$ such that*

$$\forall x \in E, \quad \mathcal{A}f(x) = \lim_{t \rightarrow 0} \frac{P_t f(x) - f(x)}{t} = \lim_{t \rightarrow 0} \frac{\mathbb{E}[f(X_t)] - f(x)}{t}.$$

where $D(\mathcal{A})$ is composed of the functions for which this limit exists. It characterizes the law of the evolution of the Markov process.

A.1.6 Markov chains

This definition is taken from the book [105, Sec 3.4].

Definition A.9. [96, Def] *Let $(X_t)_{t \geq 0}$ be a Markov chain minimal (chain that dies after explosion) on a countable state space Σ . We say that a state i is transient if*

$$\mathbb{P}_i \left(\sum_{t=1}^{\infty} \mathbb{1}_{\{X_t=i\}} = \infty \right) = 0.$$

We say that a state i is recurrent if

$$\mathbb{P}_i \left(\sum_{t=1}^{\infty} \mathbb{1}_{\{X_t=i\}} = \infty \right) = 1.$$

We denote by T_i the random variable giving the first passage time to state i ,

$$T_i \stackrel{\text{def.}}{=} \inf \{t \geq 0 : X_t = i\},$$

where $\inf \emptyset \stackrel{\text{def.}}{=} \infty$. A recurrent state i is said to be positive recurrent if $\mathbb{E}_i[T_i]$ is finite and it is said to be null recurrent otherwise.

We say that $(X_t)_{t \geq 0}$ is irreducible if every state of Σ communicates with each other:

$$\forall i, j \in \Sigma, \exists t_1, t_2 > 0, \quad \mathbb{P}_i(X_{t_1} = j) > 0 \quad \text{and} \quad \mathbb{P}_j(X_{t_2} = i) > 0.$$

This theorem is taken from the book [96, Thm 2.6.4].

Theorem A.10. (Foster's criterion) *An irreducible Markov chain $(X_n)_{n \geq 0}$ on a countable state space Σ is positive recurrent if and only if there exist a positive function $f : \Sigma \rightarrow \mathbb{R}^+$, a finite non-empty set $A \subset \Sigma$ and $\epsilon > 0$ such that*

$$\begin{aligned} \forall x \in \Sigma \setminus A, \quad & \mathbb{E}[f(X_{n+1}) - f(X_n) | X_n = x] \leq -\epsilon, \\ \forall x \in A, \quad & \mathbb{E}[f(X_{n+1}) | X_n = x] < \infty. \end{aligned}$$

A.2 Kronecker Product for Memory Lifetime

We can generalise Proposition 2.12. Instead of computing the spectrum of P_K^N , we study the spectrum of the transition matrix $M_{V,W}^N$ of $\left(V_t^N, \left(W_t^{1j,N} \right)_{2 \leq j \leq N} \right)_{t \geq 1}$. In the following we use the notation $W_t^{1,N} = \left(W_t^{1j,N} \right)_{2 \leq j \leq N}$.

We denote by

- v_1, \dots, v_{2^N} an enumeration of $\{0, 1\}^N$,
- $p_{v_k}^N = \mathbb{P}(V_t^N = v_k)$,
- $M_{v_k}^N$ is the $2^{N-1} \times 2^{N-1}$ transition matrix of $\left(W_t^{1,N} \right)_{t \geq 1}$ knowing that $V_t^N = v_k$,
- M_W^N is the $2^{N-1} \times 2^{N-1}$ transition matrix of $\left(W_t^{1,N} \right)_{t \geq 1}$. Thus, $M_W^N = \sum_{k=1}^{2^N} p_{v_k}^N M_{v_k}^N$.

The form of $M_{V,W}^N$ is:

$$M_{V,W}^N = \underbrace{\begin{bmatrix} p_{v_1}^N M_{v_1}^N & p_{v_2}^N M_{v_1}^N & \dots & p_{v_{2^K}}^N M_{v_1}^N \\ p_{v_1}^N M_{v_2}^N & p_{v_2}^N M_{v_2}^N & \dots & p_{v_{2^K}}^N M_{v_2}^N \\ \vdots & \vdots & \dots & \vdots \\ p_{v_1}^N M_{v_{2^N}}^N & p_{v_2}^N M_{v_{2^N}}^N & \dots & p_{v_{2^N}}^N M_{v_{2^N}}^N \end{bmatrix}}_{2^{2^N-1} \times 2^{2^N-1} \text{ matrix}}$$

Proposition A.11. *The spectrum of $M_{V,W}^N$ is*

$$\Sigma(M_{V,W}^N) = \Sigma(M_W^N) \cup \{0\} = \{\lambda_i^N, 0 \leq i \leq N\} \cup \{0\},$$

where λ_i^N has multiplicity $\binom{N-1}{i}$ and 0 has multiplicity $2^{N-1}(2^N - 1)$.

First, let prove a lemma useful for the proof of the proposition. It is based on the Kronecker product properties and the following 2×2 matrices:

$$M_{00} = I_2, \quad M_{01} = \begin{bmatrix} 1 & 0 \\ q_{01}^{-,N} & 1 - q_{01}^{-,N} \end{bmatrix}, \quad M_{10} = \begin{bmatrix} 1 & 0 \\ q_{10}^{-,N} & 1 - q_{10}^{-,N} \end{bmatrix}, \quad M_{11} = \begin{bmatrix} 1 - q^+ & q^+ \\ 0 & 1 \end{bmatrix}.$$

Definition A.12. *Let $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{p \times q}$. Then, the Kronecker product (or tensor product) of A and B is defined as the matrix*

$$A \otimes B = \begin{bmatrix} a_{11}B & \dots & a_{1n}B \\ a_{21}B & \dots & a_{2n}B \\ \vdots & \dots & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{bmatrix} \in \mathbb{R}^{mp \times nq}. \quad (\text{A.1})$$

Lemma A.13. *With the notation $\otimes_n M = \underbrace{M \otimes M \otimes \dots \otimes M}_{n \text{ times}}$, we have*

$$\begin{aligned} M_W^N &= (1 - f_N) \otimes_{N-1} \underbrace{\left((1 - f_N)M_{00} + f_N M_{01} \right)}_{M_0} \\ &\quad + f_N \otimes_{N-1} \underbrace{\left((1 - f_N)M_{10} + f_N M_{11} \right)}_{M_1} \end{aligned}$$

Proof. By recurrence on the number of synapses we look at. For one synapse, the transition matrix is $(1 - f_N)M_0 + f_N M_1$. In order to find the matrix for one more synapse given the matrix for n synapses, we add the value of the new synapse at the end of the previous enumeration: $e^n = (e_k^n)_{1 \leq k \leq 2^n}$ where $e_k^n \in \{0, 1\}^n$ becomes e^{n+1} such that for all $1 \leq k \leq 2^n$, $e_k^{n+1} = (0, e_k^n)$ and $e_{2^n+k}^{n+1} = (1, e_k^n)$.

The transition probabilities of the synapses of $W_t^{1,N}$ are independent knowing $V_t^{1,N}$. Hence, we condition by V_t^N and thus, we can split in two the transition matrix $M_W^{n,N}$ of n synapses of $W_t^{1,N}$. For all $1 \leq n \leq N - 1$,

$$M_W^{n,N} \stackrel{\text{def.}}{=} f_N M_{1,W}^{n,N} + (1 - f_N) M_{0,W}^{n,N},$$

where $M_{y,W}^{n,N}$ is the transition matrix knowing that $V_t^N = y$. The recurrence assumption is the following. Assume that for a $1 \leq n \leq N - 2$, $y \in \{0, 1\}$, $M_{y,W}^{n,N} = \otimes_n M_y$. Then, by independence of the transition probabilities of the synapses knowing $V_t^{1,N}$ and the fact that the transition probability matrix of the $(n + 1)^{\text{th}}$ knowing that $V_t^N = y$ is M_y , we get

$$M_{y,W}^{n+1,N} = M_y \otimes M_{y,W}^{n,N},$$

which ends the proof. \square

We now prove the previous Proposition.

Proof of Proposition A.11. First, as the number of collinear vectors of the matrix is $2^{N-1}(2^N - 1)$ thus

$$\dim(\text{Ker}(M_{V,W}^N)) \geq 2^{N-1}(2^N - 1).$$

So, $0 \in \Sigma(M_{V,W}^N)$ and has multiplicity at least $2^{N-1}(2^N - 1)$.

Then, we show that $\Sigma(M_{V,W}^N) \setminus \{0\} \subset \Sigma(M_W^N)$. Let $u = \begin{bmatrix} u^1 \\ \vdots \\ u^{2^N} \end{bmatrix} \in \mathbb{R}^{2^{2N-1}}$, with $u^i \in \mathbb{R}^{2^{N-1}}$,

and $\lambda \in]0, 1]$ such that:

$$(M_{V,W}^N)^T u = \lambda u \quad \Leftrightarrow \quad \forall i, \underbrace{p_{v_i}^N \sum_{k=1}^{2^N} (M_{v_k}^N)^T u^k}_{\tilde{u}} = \lambda u^i \quad \Leftrightarrow \quad \forall i, j, u^i p_{v_j}^N = u^j p_{v_i}^N.$$

Thus, combining the two last equalities gives

$$\underbrace{\left(\sum_{k=1}^{2^N} (M_{v_k}^N)^T p_{v_k}^N \right)}_{M_W^N} u^i = \lambda u^i \quad \Rightarrow \quad \Sigma(M_{V,W}^N) \setminus \{0\} \subset \Sigma(M_W^N).$$

Conversely, assume there exists $x \in \mathbb{R}^{2^{N-1}}$ such that $\left(\sum_{k=1}^{2^N} (M_{v_k}^N)^T p_{v_k}^N \right) x = \lambda x$. Then, defining u such that for all i , $u^i = p_{v_i}^N x$, this vector u checks

$$(M_{V,W}^N)^T u = \lambda u \quad \Rightarrow \quad \Sigma(M_W^N) \subset \Sigma(M_{V,W}^N) \setminus \{0\}.$$

We conclude that,

$$\Sigma(M_{V,W}^N) = \Sigma((M_{V,W}^N)^T) = \Sigma(M_W^N) \cup \{0\} = \{\lambda_i^N, 0 \leq i \leq N\} \cup \{0\}.$$

Finally, we find the multiplicity of these eigenvalues. We denote by $e_0 = [1 \ 1]^T$ and $e_2 = [0 \ 1]^T$, so:

$$M_0^T e_0 = e_0 = M_1^T e_0, \quad M_0^T e_2 = \Lambda_0^N e_2 \text{ and } M_1^T e_2 = \Lambda_1^N e_2 + f_N q^+ e_0$$

Now, we denote by $u_{i,N} = \left(u_{i,N}^1, \dots, u_{i,N}^{\binom{N}{i}}\right)$ the sequence of vectors that can be written as the Kronecker product of i vectors e_2 and $(N-i-1)$ vectors e_0 . For example, $u_{i,N}^1 = \otimes_i e_2 \otimes_{N-1-i} e_0$. We compute the matrix M_W^N in this basis using the linearity of the Kronecker product:

$$\begin{aligned} (M_W^N)^T u_{i,N}^j &= (1-f) \otimes_{N-1} M_0^T u_{i,N}^j + f \otimes_{N-1} M_1^T u_{i,N}^j \\ &= (1-f)(\Lambda_0^N)^i u_{i,N}^j + f \left[(\Lambda_1^N)^i u_{i,N}^j + \sum_{k=0}^{i-1} \sum_l \alpha_{k,l} u_{k,N}^l \right] \\ &= \left(\underbrace{(1-f)(\Lambda_0^N)^i + f(\Lambda_1^N)^i}_{\lambda_i^N} \right) u_{i,N}^j + f \sum_{k=0}^{i-1} \sum_l \alpha_{k,l} u_{k,N}^l \end{aligned}$$

Where $\alpha_{k,l} \in \{0, (fq^+)^{i-k}\}$. Hence, in this basis, $(M_W^N)^T$ is upper triangular with λ_i^N on the diagonal with multiplicity $\binom{K-1}{i}$, it ends the proof on $\Sigma(M_{V,W}^N)$. \square

In hindsight, denoting by $M_W^{i,N}$ the transition matrix of $(W_t^{i,N})_{t \geq 1}$, we note that it can be written using the Bayes' formula as

$$\begin{aligned} M_W^{i,N} &= \sum_{v \in \{0,1\}^N} \mathbb{P}(V_t^N = v) \left[(1-v^i) \otimes_{k \neq i} \left((1-v^k) M_{00} + v^k M_{01} \right) \right. \\ &\quad \left. + v^i \otimes_{k \neq i} \left((1-v^k) M_{10} + v^k M_{11} \right) \right]. \end{aligned}$$

A.3 Invariant measure of the fast process in dimension 2

Using the same argument as in Lemma 4.3, we can show that π_w admits a density in s and we write for all $(v, s) \in \{0, 1\}^2 \times \mathbb{R}_+^2$ and a fixed w ,

$$\pi_w(v, ds) = \sum_{k=1}^4 \delta_{v,k}(v) p_k(s^1, s^2) ds^1 ds^2 \quad \text{with} \quad \int_{\mathbb{R}_+^2} p_k(s^1, s^2) ds^1 ds^2 = \nu_w^{v_k}.$$

Moreover, we assume that for all i , $\alpha_i(v, w) = \alpha(\sum_j w^{ij} v^j)$. We denote by $\alpha_0 = \alpha(0)$, $\alpha_1 = \alpha(w^{12})$ and $\alpha_2 = \alpha(w^{21})$. We denote by $D_{(1,1)} f(u) = \partial_{s^1} f(u) + \partial_{s^2} f(u)$.

The Fokker-Planck equation on the invariant measure obtained from the generator \mathcal{B}_w , see (3.6),

gives

$$\begin{aligned}
D_{(1,1)}p_1(s) &= \beta(p_2(s) + p_3(s)) - 2\alpha_0p_1(s) - \delta_0(s^1)p_1(s) - \delta_0(s^2)p_1(s) \\
D_{(1,1)}p_2(s) &= -(\beta + \alpha_1)p_2(s) + \beta p_4(s) + \delta_0(s^2)\left(-p_2(s) + \alpha_0 \int_{\mathbb{R}_+} p_1(s)ds^2\right) - \delta_0(s^1)p_2(s) \\
D_{(1,1)}p_3(s) &= -(\beta + \alpha_2)p_3(s) + \beta p_4(s) + \delta_0(s^1)\left(-p_3(s) + \alpha_0 \int_{\mathbb{R}_+} p_1(s)ds^1\right) - \delta_0(s^2)p_3(s) \\
D_{(1,1)}p_4(s) &= -2\beta p_4(s) + \delta_0(s^2)\left(-p_4(s) + \alpha_2 \int_{\mathbb{R}_+} p_3(s)ds^2\right) \\
&\quad + \delta_0(s^1)\left(-p_4(s) + \alpha_1 \int_{\mathbb{R}_+} p_2(s)ds^1\right).
\end{aligned}$$

Hence, denoting by

$$A = \begin{bmatrix} -2\alpha_0 & \beta & \beta & 0 \\ 0 & -\alpha_1 - \beta & 0 & \beta \\ 0 & 0 & -\alpha_2 - \beta & \beta \\ 0 & 0 & 0 & -2\beta \end{bmatrix} \quad \text{and} \quad p = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{bmatrix},$$

we get, for all s such that both $s^1 \neq 0$ and $s^2 \neq 0$,

$$D_{(1,1)}p = Ap \tag{A.2}$$

with the following boundary conditions

$$\begin{aligned}
p_1(0, s^2) = 0 & \quad , \quad \alpha_0 \int_{\mathbb{R}_+} p_1(s^1, \theta)d\theta = p_2(s^1, 0), \\
p_1(s^1, 0) = 0 & \quad , \quad \alpha_0 \int_{\mathbb{R}_+} p_1(\theta, s^2)d\theta = p_3(0, s^2), \\
p_2(0, s^2) = 0 & \quad , \quad \alpha_1 \int_{\mathbb{R}_+} p_2(\theta, s^2)d\theta = p_4(0, s^2), \\
p_3(s^1, 0) = 0 & \quad , \quad \alpha_2 \int_{\mathbb{R}_+} p_3(s^1, \theta)d\theta = p_4(s^1, 0).
\end{aligned} \tag{A.3}$$

We make a change of variable in order to solve equation (A.2). Let $\phi : \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow D$ such that $\phi(s^1, s^2) = (u, v) = (s^1, s^2 - s^1)$ and $D = \{(x, y), x \in \mathbb{R}^+, y \in \mathbb{R} : y \geq x\}$. Therefore,

we search p_k such that $p_k(s) = f_k(\phi(s^1, s^2)) = f_k(u, v)$ and we get with $f = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{bmatrix}$ and $\forall u \neq$

$0, u \neq v$:

$$\partial_u f = Af \Leftrightarrow \exists g(v) = \begin{bmatrix} g_1(v) \\ g_2(v) \\ g_3(v) \\ g_4(v) \end{bmatrix}, \quad f(u, v) = e^{Au}g(v).$$

After computing the exponential of the matrix A , we modify the functions g_2 and g_3 as follows (we keep the same name for simplicity):

$$g_2 \leftarrow g_2 - \frac{\beta}{\alpha_1 - \beta}g_4 \quad \text{and} \quad g_3 \leftarrow g_3 - \frac{\beta}{\alpha_2 - \beta}g_4.$$

We finally obtain that the p_k have the following form:

$$\begin{aligned}
p_1(s^1, s^2) &= e^{-2\alpha_0 s^1} g_1(s^2 - s^1) + \frac{\beta}{2\alpha_0 - \alpha_1 - \beta} e^{-(\alpha_1 + \beta)s^1} g_2(s^2 - s^1) \\
&\quad + \frac{\beta}{2\alpha_0 - \alpha_2 - \beta} e^{-(\alpha_2 + \beta)s^1} g_3(s^2 - s^1) \\
&\quad + \frac{\beta^2(\alpha_1 + \alpha_2 - 2\beta)}{(\alpha_1 - \beta)(\alpha_2 - \beta)(2\alpha_0 - 2\beta)} e^{-2\beta s^1} g_4(s^2 - s^1). \\
p_2(s^1, s^2) &= e^{-(\alpha_1 + \beta)s^1} g_2(s^2 - s^1) + \frac{\beta}{\alpha_1 - \beta} e^{-2\beta s^1} g_4(s^2 - s^1), \\
p_3(s^1, s^2) &= e^{-(\alpha_2 + \beta)s^1} g_3(s^2 - s^1) + \frac{\beta}{\alpha_2 - \beta} e^{-2\beta s^1} g_4(s^2 - s^1), \\
p_4(s^1, s^2) &= e^{-2\beta s^1} g_4(s^2 - s^1).
\end{aligned}$$

We now use the Laplace transforms and the boundary conditions. The first column of (A.3) gives

$$\begin{aligned}
0 &= g_1(s^2) + \frac{\beta}{2\alpha_0 - \alpha_1 - \beta} g_2(s^2) + \frac{\beta}{2\alpha_0 - \alpha_2 - \beta} g_3(s^2) + \frac{\beta^2(\alpha_1 + \alpha_2 - 2\beta)}{(\alpha_1 - \beta)(\alpha_2 - \beta)(2\alpha_0 - 2\beta)} g_4(s^2), \\
0 &= g_1(-s^1) e^{-2\alpha_0 s^1} + \frac{\beta}{2\alpha_0 - \alpha_1 - \beta} g_2(-s^1) e^{-(\alpha_1 + \beta)s^1} + \frac{\beta}{2\alpha_0 - \alpha_2 - \beta} g_3(-s^1) e^{-(\alpha_2 + \beta)s^1} \\
&\quad + \frac{\beta^2(\alpha_1 + \alpha_2 - 2\beta)}{(\alpha_1 - \beta)(\alpha_2 - \beta)(2\alpha_0 - 2\beta)} g_4(-s^1) e^{-2\beta s^1}, \\
0 &= g_2(s^2) + \frac{\beta}{\alpha_1 - \beta} g_4(s^2) \\
0 &= g_3(-s^1) e^{-(\alpha_2 + \beta)s^1} + \frac{\beta}{\alpha_2 - \beta} g_4(-s^1) e^{-2\beta s^1}.
\end{aligned} \tag{A.4}$$

The second column of (A.3) gives

$$\begin{aligned}
\alpha_2 \mathcal{L}\{p_3\}(\lambda^1, 0) &= \int e^{-\lambda^1 s^1} p_4(s^1, 0) ds^1 = \int e^{-(\lambda^1 + 2\beta)s^1} g_4(-s^1) ds^1, \\
\alpha_1 \mathcal{L}\{p_2\}(0, \lambda^2) &= \int e^{-\lambda^2 s^2} p_4(0, s^2) ds^2 = \int e^{-\lambda^2 s^2} g_4(s^2) ds^2, \\
\alpha_0 \mathcal{L}\{p_1\}(\lambda^1, 0) &= \int e^{-\lambda^1 s^1} p_2(s^1, 0) ds^1 \\
&= \int g_2(-s^1) e^{-(\lambda^1 + \alpha_2 + \beta)s^1} + \frac{\beta}{\alpha_1 - \beta} g_4(-s^1) e^{-(\lambda^1 + 2\beta)s^1} ds^1, \\
\alpha_0 \mathcal{L}\{p_1\}(0, \lambda^2) &= \int e^{-\lambda^2 s^2} p_3(0, s^2) ds^2 = \int e^{-\lambda^2 s^2} g_3(s^2) + \frac{\beta}{\alpha_2 - \beta} g_4(s^2) ds^2.
\end{aligned}$$

We denote by g_k^- and g_k^+ the functions from \mathbb{R}_+ to \mathbb{R} such that $g_k(s) = g_k^+(s) \mathbb{1}_{\mathbb{R}_+}(s) + g_k^-(-s) \mathbb{1}_{\mathbb{R}_-}(s)$. We deduce that

$$\begin{aligned}
\mathcal{L}\{g_4^-\}(\lambda^1) &= \alpha_2 \mathcal{L}\{p_3\}(\lambda^1 - 2\beta, 0), \\
\mathcal{L}\{g_4^+\}(\lambda^2) &= \alpha_1 \mathcal{L}\{p_2\}(0, \lambda^2), \\
\mathcal{L}\{g_2^-\}(\lambda^1) &= \alpha_0 \mathcal{L}\{p_1\}(\lambda^1 - \alpha_2 - \beta, 0) - \frac{\beta \alpha_2}{\alpha_1 - \beta} \mathcal{L}\{p_3\}(\lambda^1 - \alpha_2 - \beta, 0), \\
\mathcal{L}\{g_3^+\}(\lambda^2) &= \alpha_0 \mathcal{L}\{p_1\}(0, \lambda^2) - \frac{\beta \alpha_1}{\alpha_2 - \beta} \mathcal{L}\{p_2\}(0, \lambda^2).
\end{aligned}$$

We saw how to compute the Laplace transforms in dimension 2 in Lemma 3.36. We obtain that

$$g_4(y) = \begin{cases} c_{14}^+ e^{-\beta y} + c_{24}^+ e^{-(2\beta+\alpha_1)y} & , \forall y \in \mathbb{R}^+ \\ c_{14}^- e^{-\beta y} + c_{24}^- e^{\alpha_2 y} & , \forall y \in \mathbb{R}^- \end{cases}$$

$$g_3(y) = \begin{cases} c_{13} e^{-\lambda_{13}y} + c_{23} e^{-\lambda_{23}y} + c_{33} e^{-\beta y} + c_{43} e^{-(\alpha_1+2\beta)y} & , \forall y \in \mathbb{R}^+ \\ -\frac{\beta}{\alpha_2-\beta} e^{(\beta-\alpha_2)y} g_4(y) & , \forall y \in \mathbb{R}^- \end{cases}$$

$$g_2(y) = \begin{cases} -\frac{\beta}{\alpha_1-\beta} g_4(y) & , \forall y \in \mathbb{R}^+ \\ c_{12} e^{\lambda_{12}y} + c_{22} e^{\lambda_{22}y} + c_{32} e^{-\alpha_1 y} + c_{43} e^{(\alpha_2+\beta-\alpha_1)y} & , \forall y \in \mathbb{R}^- \end{cases}$$

where

$$\lambda_{12} = \frac{2\alpha_0 - \alpha_1 - \beta + \sqrt{(2\alpha_0 - \alpha_1 - \beta)^2 + 4\alpha_1\beta}}{2}$$

$$\lambda_{22} = \frac{2\alpha_0 - \alpha_1 - \beta - \sqrt{(2\alpha_0 - \alpha_1 - \beta)^2 + 4\alpha_1\beta}}{2}$$

$$\lambda_{13} = \frac{2\alpha_0 + \alpha_2 + \beta + \sqrt{(2\alpha_0 + \alpha_2 + \beta)^2 + 4\alpha_2\beta}}{2}$$

$$\lambda_{23} = \frac{2\alpha_0 + \alpha_2 + \beta - \sqrt{(2\alpha_0 + \alpha_2 + \beta)^2 + 4\alpha_2\beta}}{2}$$

and the constants depend on the parameters. These constants are unfortunately not clear enough to be written here. The function g_1 is obtained from g_2 , g_3 and g_4 using equation (A.4).

A.4 Code of Chapter 4

We simulate N neurons $X_t^{1,*}, \dots, X_t^{N,*}$ having the same dynamics as X_t^* , except that instead of μ_t^* we use $\mu_t^{*,N} = \frac{1}{N} \sum_i \delta_{X_t^{i,*}}$. Hence, we do not need to compute μ_t^* . For instance, we use the $\xi_t^{i,*}$ to compute the $I_t^{i,*} = I(\xi_t^{i,*})$. The neuron i satisfies the following equations

- $\mathcal{L}(X_0^{i,*}) = \mu_0^*$, has weight only on ξ such that $\xi(\cdot, \cdot, \mathbb{Z}) = \mu_0^*(\cdot, \cdot, \mathcal{P}(E_m))$, and $S_0^{i,*}$ with law ρ_0 admitting a density.
- $dS_t^{i,*} = dt$.
- $\xi_t^{i,*}$ admits a density in s and satisfies the following equation where we used $\mu_t^{*,N}$ instead

of μ_t^* with $\mu_t^{*,N} = \frac{1}{N} \sum_i \delta_{(V_t^{i,*}, S_t^{i,*}, \xi_t^{i,*})}$:

$$\begin{aligned} \partial_t \xi_t^{i,*0}(s, w) &= -\partial_s \xi_t^{i,*0}(s, w) + \beta \xi_t^{i,*1}(s, w) - \xi_t^{i,*0}(s, w) \underbrace{\frac{\frac{1}{N} \sum_k \alpha(I(\xi_t^{k,*})) \mathbb{1}_{\{V_t^{k,*}=0, S_t^{k,*}=s\}}}{\frac{1}{N} \sum_l \mathbb{1}_{\{V_t^{l,*}=0, S_t^{l,*}=s\}}}}_{a_t^0(s)} \\ \xi_t^{i,*0}(0, w) &= 0, \\ \partial_t \xi_t^{i,*1}(s, w) &= -\partial_s \xi_t^{i,*1}(s, w) - \beta \xi_t^{i,*1}(s, w) \\ \xi_t^{i,*1}(0, w) &= \frac{\Delta t \sum_k \alpha(I(\xi_t^{k,*})) [p^-(S_t^{i,*}) \xi_t^{i,*0}(S_t^{k,*}, w+1) + (1-p^-(S_t^{i,*})) \xi_t^{i,*0}(S_t^{k,*}, w)]}{\frac{1}{N} \sum_l \mathbb{1}_{\{V_t^{l,*}=0, S_t^{l,*}=S_t^{k,*}\}}} \end{aligned} \quad (\text{A.5})$$

- At rate $\beta \mathbb{1}_{\{V_{t^-}^{i,*}=1\}}$, $(V_{t^-}^{i,*}, S_{t^-}^{i,*}, \xi_{t^-}^{i,*})$ jumps to $(0, S_{t^-}^{i,*}, \xi_{t^-}^{i,*})$.
- At rate $\alpha(I(\xi_{t^-}^{i,*})) \mathbb{1}_{\{V_{t^-}^{i,*}=0\}}$, $(V_{t^-}^{i,*}, S_{t^-}^{i,*}, \xi_{t^-}^{i,*})$ jumps to $(1, 0, \xi_{t^-}^{i,*})$ with

$$\xi_t^{i,*}(\{v\}, A, \{w\}) = \int_A p^+(s, w-1) \xi_{t^-}^{i,*}(\{v\}, ds, \{w-1\}) + \int_A (1-p^+(s, w)) \xi_{t^-}^{i,*}(\{v\}, ds, \{w\}).$$

We use a simulation with a fixed Δt time step instead of following exactly the jumping times one by one. In addition, two-dimensional histograms are used for the measures $\xi_t^{i,*0}$ and $\xi_t^{i,*1}$. For the continuous variable s , we use a grid of size Δt starting from 0 and ending in $m_s = M_s \Delta t$ ($m_s = 15\text{ms}$ in the simulations of Section 4.4) with $M_s \in \mathbb{N}$. It is also necessary to choose bounds for synaptic weights, we denote $w_{min}, w_{max} \in \mathbb{Z}$ such that $W_t^{ij,N} \in \llbracket w_{min}, w_{max} \rrbracket$. Finally, thereafter, we denote by $m_t^* = (m_t^{1,*}, \dots, m_t^{N,*})$ and $a_t^0 = (a_t^0(0), a_t^0(\Delta t), \dots, a_t^0(M_s \Delta t))$ the vectors such that for any $m \in \mathbb{N}^*$:

$$m_t^{k,*} = \left\lfloor \frac{S_t^{k,*}(1 - V_t^{k,*})}{\Delta t} \right\rfloor \quad \text{and} \quad a_t^0(m \Delta t) = \frac{\sum_k \alpha(I(\xi_t^{k,*})) \mathbb{1}_{\{m_t^{k,*}=m\}}}{\sum_l \mathbb{1}_{\{m_t^{l,*}=m\}}}.$$

In order to discretize the equations given in (A.5), an explicit Euler scheme is used to estimate the derivatives: for a function $u : (t, s) \mapsto u(t, s)$,

$$\partial_t u(t, s) = \frac{u(t + \Delta t, s) - u(t, s)}{\Delta t} \quad \text{and} \quad -\partial_s u(t, s) = \frac{u(t, s - \Delta t) - u(t, s)}{\Delta t}.$$

Initial conditions

The probability measures $\xi_0^{i,*}$ are initially identical and distributed with the laws described in what follows. These laws admit a densities in s that we approximate with histograms. We have to be careful with boundary conditions in $s = 0$. We ensure that these conditions are initially satisfied. The laws proposed are closed to the one obtained after a long time of simulation.

We draw the initial variables as follows: with $0 < p_v < 1$,

1. We draw the potentials from the Bernoulli distribution of parameter p_v , $V_0^{i,*} \stackrel{\mathcal{L}}{=} \text{Bern}(p_v)$,
2. We draw the $S_0^{i,*}$ such that $V_0^{i,*} = 0$ from the distribution $\text{LogNormal}(x, y)$, and then we compute $m_0^{1,*}, \dots, m_0^{N,*}$ and $a_0^0(0), \dots, a_0^0(M_s)$,

3. we define $\xi_0^{i,*} = (\delta_0 \otimes \xi_0^{i,*0} + \delta_1 \otimes \xi_0^{i,*1}) \otimes \mathcal{U}(\{w_{min}, \dots, w_{max}\})$ where

$$\xi_0^{i,*0}(\cdot, \mathbb{Z}) \stackrel{\mathcal{L}}{=} \text{LogNormal}(x, y) \quad \text{and} \quad \xi_0^{i,*0}(\cdot, \mathbb{Z}) \stackrel{\mathcal{L}}{=} \mathcal{E}(\xi_0^{i,*1}(0, \mathbb{Z}))$$

with

$$\xi_0^{i,*1}(0, w) = \Delta t \sum_m a_0^0(m) \xi_0^{i,*0}(m \Delta t, w),$$

4. Finally, we draw the $S_0^{i,*}$ such that $V_0^{i,*} = 1$ from the distribution $\mathcal{E}(\xi_0^{i,*1}(0, \mathbb{Z}))$.

From $t = 0$ to $t = \Delta t$

We describe how to pass from $t = 0$ to $t = \Delta t$, and then we perform the same computations to pass from $t = \Delta t$ to $t = 2\Delta t$ and so forth.

1. We compute the $I_0^{i,*}$ (initially all the same and deterministic).
2. We draw the jumping times $\tau_0^{i,*}$ from the exponential distributions of parameters

$$\alpha(I_0^{i,*}) \mathbb{1}_{\{V_0^{i,*}=0\}} + \beta \mathbb{1}_{\{V_0^{i,*}=1\}}.$$

3. We compute $m_0^* = (m_0^{1,*}, \dots, m_0^{N,*})$ and $a_0^0 = (a_0^0(0), \dots, a_0^0(M_s))$ where for all $m \in \mathbb{N}$:

$$m_0^{k,*} = \left\lfloor \frac{S_0^{k,*}}{\Delta t} \right\rfloor \quad \text{and} \quad a_0^0(m) = \frac{\sum_k \alpha(I(\xi_0^{k,*})) \mathbb{1}_{\{m_0^{k,*}=m\}}}{\sum_l \mathbb{1}_{\{m_0^{l,*}=m\}}}.$$

We compute the terms due to the drift of the $\xi_{\Delta t}^{i,*}$. For all $m \in \llbracket 0, M_s \rrbracket$:

$$\begin{aligned} \xi_{\Delta t}^{i,*0}(m \Delta t, w) &= \xi_0^{i,*0}((m-1)\Delta t, w) + \Delta t \beta \xi_0^{i,*1}((m-1)\Delta t, w) \\ &\quad - \Delta t \xi_0^{i,*0}((m-1)\Delta t, w) a_0^0(m-1) \\ \xi_{\Delta t}^{i,*0}(0, w) &= 0, \end{aligned}$$

$$\xi_{\Delta t}^{i,*1}(m \Delta t, w) = \xi_0^{i,*1}((m-1)\Delta t, w) (1 - \Delta t \beta)$$

$$\begin{aligned} \xi_{\Delta t}^{i,*1}(0, w) &= \Delta t \sum_k \frac{\alpha(I(\xi_0^{k,*})) \mathbb{1}_{\{m_0^{k,*} \geq 1\}}}{\sum_l \mathbb{1}_{\{m_0^{l,*} = m_0^{k,*}\}}} \left[p^-(S_0^{i,*}) \xi_0^{i,*0}(m_0^{k,*} \Delta t, w + 1) \right. \\ &\quad \left. + (1 - p^-(S_0^{i,*})) \xi_0^{i,*0}(m_0^{k,*} \Delta t, w) \right] \\ &= \Delta t \sum_m \frac{\sum_k \alpha(I(\xi_0^{k,*})) \mathbb{1}_{\{m_0^{k,*} \geq 1, m_0^{k,*} = m\}}}{\sum_l \mathbb{1}_{\{m_0^{l,*} = m\}}} \left[p^-(S_0^{i,*}) \xi_0^{i,*0}(m \Delta t, w + 1) \right. \\ &\quad \left. + (1 - p^-(S_0^{i,*})) \xi_0^{i,*0}(m \Delta t, w) \right] \\ &= \Delta t \sum_m a_0^0(m) \left[p^-(S_0^{i,*}) \xi_0^{i,*0}(m \Delta t, w + 1) + (1 - p^-(S_0^{i,*})) \xi_0^{i,*0}(m \Delta t, w) \right] \end{aligned}$$

4. If $\tau_0^{i,*} < \Delta t$ and $V_0^{i,*} = 0$, for all $m \in \llbracket 0, M_s \rrbracket$,

$$\xi_{\Delta t}^{i,*}(\{v\}, m\Delta t, \{w\}) = p^+(m\Delta t)\xi_0^{i,*}(\{v\}, m\Delta t, \{w-1\}) + (1-p^+(m\Delta t))\xi_0^{i,*}(\{v\}, m\Delta t, \{w\}).$$

5. We compute

$$S_{\Delta t}^{i,*} = (S_0^{i,*} + \Delta t)\mathbb{1}_{\{\tau_0^{i,*} \geq \Delta t\}} + (\Delta t - \tau_0^{i,*})\mathbb{1}_{\{\tau_0^{i,*} < \Delta t\}},$$

6. We make jump the potentials of the neurons i such that $\tau_0^{i,*} < \Delta t$.

7. We compute the elements that we keep in memory at each time step : expectations of the potentials, of the time from last the spike, of the presynaptic weights, distribution of the synaptic current input...

Bibliography

- [1] More about Neurotransmitters and How Neurons Communicate. <https://ohiostate.pressbooks.pub/swk3805modules34/chapter/ch-4-more-about-neurotransmitters-and-how-neurons-communicate/>. [Online; accessed 20-January-2021].
- [2] ABBOTT, L. F., AND NELSON, S. B. Synaptic plasticity: taming the beast. *Nature Neuroscience* 3, 11 (2000), 1178–1183.
- [3] AMARI, S.-I. Characteristics of sparsely encoded associative memory. *Neural Networks* 2, 6 (1989), 451–457.
- [4] AMBROSIO, L., FUSCO, N., AND PALLARA, D. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2000.
- [5] AMIT, D. J., AND FUSI, S. Learning in Neural Networks with Material Synapses. *Neural Computation* 6, 5 (1994), 957–982.
- [6] AMIT, D. J., GUTFREUND, H., AND SOMPOLINSKY, H. Spin-glass models of neural networks. *Physical Review A* 32, 2 (1985), 1007.
- [7] AMIT, D. J., AND MONGILLO, G. Spike-Driven Synaptic Dynamics Generating Working Memory States. *Neural Computation* 15, 3 (Mar. 2003), 565–596.
- [8] AMIT, Y., AND HUANG, Y. Precise Capacity Analysis in Binary Networks with Multiple Coding Level Inputs. *Neural Computation* 22, 3 (2010), 660–688. PMID: 19842984.
- [9] APPLEBY, P. A., AND ELLIOTT, T. Synaptic and temporal ensemble interpretation of Spike-Timing-Dependent Plasticity. *Neural Computation* 17, 11 (2005), 2316–2336.
- [10] APPLEBY, P. A., AND ELLIOTT, T. Stable competitive dynamics emerge from multispikes interactions in a stochastic model of Spike-Timing-Dependent Plasticity. *Neural Computation* 18, 10 (2006), 2414–2464.
- [11] BANSAYE, V., CLOEZ, B., AND GABRIEL, P. Ergodic behavior of non-conservative semi-groups via generalized Doeblin’s conditions. *Acta Appl. Math.* 166 (2020), 29–72.
- [12] BANSAYE, V., AND MÉLÉARD, S. *Stochastic models for structured populations*, vol. 1 of *Mathematical Biosciences Institute Lecture Series. Stochastics in Biological Systems*. Springer, Cham; MBI Mathematical Biosciences Institute, Ohio State University, Columbus, OH, 2015. Scaling limits and long time behavior.
- [13] BELL, C. C., HAN, V. Z., SUGAWARA, Y., AND GRANT, K. Synaptic plasticity in a cerebellum-like structure depends on temporal order. *Nature* 387, 6630 (May 1997), 278–281.

- [14] BELLINGHAM, M. C., LIM, R., AND WALMSLEY, B. Developmental changes in EPSC quantal size and quantal content at a central glutamatergic synapse in rat. *The Journal of Physiology* 511, 3 (1998), 861–869.
- [15] BENAYOUN, M., COWAN, J. D., VAN DRONGELEN, W., AND WALLACE, E. Avalanches in a Stochastic Model of Spiking Neurons. *PLOS Computational Biology* 6, 7 (2010), 1–13. Publisher: Public Library of Science.
- [16] BENNA, M. K., AND FUSI, S. Computational principles of synaptic memory consolidation. *Nature Neuroscience* 19, 12 (Oct. 2016), 1697–1706.
- [17] BI, G.-Q., AND POO, M.-M. Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *Journal of neuroscience* 18, 24 (1998), 10464–10472.
- [18] BIENENSTOCK, E. L., COOPER, L. N., AND MUNRO, P. W. Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. *Journal of Neuroscience* 2, 1 (1982), 32–48.
- [19] BILLINGSLEY, P. *Convergence of probability measures*, second ed. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, 1999. A Wiley-Interscience Publication.
- [20] BLISS, T. V., AND LØMO, T. Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *The Journal of physiology* 232, 2 (1973), 331–356.
- [21] BRESSLOFF, P. C. Metastable states and quasicycles in a stochastic wilson-cowan model of neuronal population dynamics. *Physical Review E* 82, 5 (2010), 051903.
- [22] BRETTE, R., AND GERSTNER, W. Adaptive Exponential Integrate-and-Fire Model as an Effective Description of Neuronal Activity. *Journal of neurophysiology* 94, 5 (2005), 3637–3642.
- [23] BRUNEL, N., CARUSI, F., AND FUSI, S. Slow stochastic Hebbian learning of classes of stimuli in a recurrent neural network. *Network: Computation in Neural Systems* 9, 1 (1998), 123–152.
- [24] BRZOSKO, Z., MIERAU, S. B., AND PAULSEN, O. Neuromodulation of Spike-Timing-Dependent Plasticity: Past, Present, and Future. *Neuron* 103, 4 (2019), 563–581.
- [25] BURKITT, A. N., GILSON, M., AND VAN HEMMEN, J. L. Spike-Timing-Dependent Plasticity for neurons with recurrent connections. *Biological Cybernetics* 96, 5 (2007), 533–546.
- [26] CAÑIZO, J. A., AND YOLDAŞ, H. Asymptotic behaviour of neuron population models structured by elapsed-time. *Nonlinearity* 32, 2 (2019), 464–495.
- [27] CÁCERES, M. J., CARRILLO, J. A., AND PERTHAME, B. Analysis of nonlinear noisy integrate & fire neuron models: blow-up and steady states. *J. Math. Neurosci.* 1 (2011), Art. 7, 33.
- [28] CASELLA, G., AND BERGER, R. L. *Statistical inference*. The Wadsworth & Brooks/Cole Statistics/Probability Series. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, 1990.

- [29] CHERNOFF, H. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Statistics* 23 (1952), 493–507.
- [30] CHEVALLIER, J. Mean-field limit of generalized Hawkes processes. *Stochastic Process. Appl.* 127, 12 (2017), 3870–3912.
- [31] CLOPATH, C., BÜSING, L., VASILAKI, E., AND GERSTNER, W. Connectivity reflects coding: a model of voltage-based STDP with homeostasis. *Nature Neuroscience* 13, 3 (Mar. 2010), 344–352.
- [32] DALEY, D. J., AND VERE-JONES, D. *An introduction to the theory of point processes. Vol. I*, second ed. Probability and its Applications (New York). Springer-Verlag, New York, 2003. Elementary theory and methods.
- [33] DAVIS, M. H. A. Piecewise-deterministic Markov Processes: A General Class of Non-diffusion Stochastic Models. *J. Roy. Statist. Soc. Ser. B* 46, 3 (1984), 353–388. With discussion.
- [34] DAVIS, M. H. A. *Markov models and optimization*, vol. 49 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London, 1993.
- [35] DE MASI, A., GALVES, A., LÖCHERBACH, E., AND PRESUTTI, E. Hydrodynamic limit for interacting neurons. *J. Stat. Phys.* 158, 4 (2015), 866–902.
- [36] DE PITTÀ, M., BRUNEL, N., AND VOLTERRA, A. Astrocytes: Orchestrating synaptic plasticity? *Neuroscience* 323 (2016), 43–61.
- [37] DEFELIPE, J. Brain plasticity and mental processes: Cajal again. *Nature Reviews Neuroscience* 7, 10 (2006), 811–817.
- [38] DELARUE, F., INGLIS, J., RUBENTHALER, S., AND TANRÉ, E. Global solvability of a networked integrate-and-fire model of McKean-Vlasov type. *Ann. Appl. Probab.* 25, 4 (2015), 2096–2133.
- [39] DELARUE, F., INGLIS, J., RUBENTHALER, S., AND TANRÉ, E. Particle systems with a singular mean-field self-excitation. Application to neuronal networks. *Stochastic Process. Appl.* 125, 6 (2015), 2451–2492.
- [40] DELATTRE, S., AND FOURNIER, N. Statistical inference versus mean field limit for Hawkes processes. *Electron. J. Stat.* 10, 1 (2016), 1223–1295.
- [41] DUBREUIL, A. M., AMIT, Y., AND BRUNEL, N. Memory Capacity of Networks with Stochastic Binary Synapses. *PLOS Computational Biology* 10, 8 (2014), 1–15. Publisher: Public Library of Science.
- [42] ELLIOTT, T. Memory Nearly on a Spring: A Mean First Passage Time Approach to Memory Lifetimes. *Neural Computation* 26, 9 (Sept. 2014), 1873–1923.
- [43] FITZHUGH, R. Impulses and physiological states in theoretical models of nerve membrane. *Biophysical journal* 1, 6 (1961), 445.
- [44] FOURNIER, N., AND LÖCHERBACH, E. On a toy model of interacting neurons. *Ann. Inst. Henri Poincaré Probab. Stat.* 52, 4 (2016), 1844–1876.

- [45] FOX, K., AND STRYKER, M. Integrating Hebbian and homeostatic plasticity: introduction. *Philosophical Transactions of the Royal Society B: Biological Sciences* 372, 1715 (2017), 20160413.
- [46] FROEMKE, R. C. Temporal modulation of Spike-Timing-Dependent Plasticity. *Frontiers in Synaptic Neuroscience* (2010).
- [47] FUSI, S. Computational models of long term plasticity and memory. *arXiv:1706.04946 [q-bio]* (June 2017). arXiv: 1706.04946.
- [48] FUSI, S., AND ABBOTT, L. Limits on the memory storage capacity of bounded synapses. *Nature Neuroscience* 10, 4 (2007), 485–493.
- [49] FUSI, S., DREW, P. J., AND ABBOTT, L. Cascade Models of Synaptically Stored Memories. *Neuron* 45, 4 (Feb. 2005), 599–611.
- [50] GABRIEL, P. Measure solutions to the conservative renewal equation. In *CIMPA School on Mathematical Models in Biology and Medicine*, vol. 62 of *ESAIM Proc. Surveys*. EDP Sci., Les Ulis, 2018, pp. 68–78.
- [51] GALTIER, M., AND WAINRIB, G. Multiscale analysis of slow-fast neuronal learning models with noise. *J. Math. Neurosci.* 2 (2012), Art. 13 64.
- [52] GALTIER, M. N., AND WAINRIB, G. A biological gradient descent for prediction through a combination of STDP and homeostatic plasticity. *Neural Computation* 25, 11 (2013), 2815–2832.
- [53] GALVES, A., LÖCHERBACH, E., POUZAT, C., AND PRESUTTI, E. A system of interacting neurons with short term synaptic facilitation. *J. Stat. Phys.* 178, 4 (2020), 869–892.
- [54] GARDNER, E., AND DERRIDA, B. Optimal storage properties of neural network models. *J. Phys. A* 21, 1 (1988), 271–284.
- [55] GERSTNER, W., KEMPTER, R., VAN HEMMEN, J. L., AND WAGNER, H. A neuronal learning rule for sub-millisecond temporal coding. *Nature* 383, 6595 (1996), 76–78.
- [56] GERSTNER, W., AND KISTLER, W. M. *Spiking neuron models*. Cambridge University Press, Cambridge, 2002. Single neurons, populations, plasticity.
- [57] GERSTNER, W., KISTLER, W. M., NAUD, R., AND PANINSKI, L. *Neuronal dynamics: from single neurons to networks and models of cognition*. Cambridge University Press, Cambridge, United Kingdom, 2014.
- [58] GILSON, M., BURKITT, A. N., GRAYDEN, D. B., THOMAS, D. A., AND VAN HEMMEN, J. L. Emergence of network structure due to Spike-Timing-Dependent Plasticity in recurrent neuronal networks. i. input selectivity–strengthening correlated input pathways. *Biological Cybernetics* 101, 2 (2009), 81–102.
- [59] GILSON, M., BURKITT, A. N., GRAYDEN, D. B., THOMAS, D. A., AND VAN HEMMEN, J. L. Emergence of network structure due to Spike-Timing-Dependent Plasticity in recurrent neuronal networks v: self-organization schemes and weight dependence. *Biological Cybernetics* 103, 5 (2010), 365–386.
- [60] GILSON, M., AND FUKAI, T. Stability versus Neuronal Specialization for STDP: Long-Tail Weight Distributions Solve the Dilemma. *PLOS ONE* 6, 10 (2011), 1–18. Publisher: Public Library of Science.

- [61] GILSON, M., FUKAI, T., AND BURKITT, A. N. Spectral Analysis of Input Spike Trains by Spike-Timing-Dependent Plasticity. *PLOS Computational Biology* 8, 7 (2012), 1–22. Publisher: Public Library of Science.
- [62] GRAHAM, C., AND TALAY, D. *Stochastic simulation and Monte Carlo methods*, vol. 68 of *Stochastic Modelling and Applied Probability*. Springer, Heidelberg, 2013. Mathematical foundations of stochastic simulation.
- [63] GRAUPNER, M., AND BRUNEL, N. Calcium-based plasticity model explains sensitivity of synaptic changes to spike pattern, rate, and dendritic location. *Proceedings of the National Academy of Sciences* 109, 10 (2012), 3991–3996.
- [64] GRIPON, V., HEUSEL, J., LÖWE, M., AND VERMET, F. A comparative study of sparse associative memories. *J. Stat. Phys.* 164, 1 (2016), 105–129.
- [65] GÜTIG, R., AHARONOV, R., ROTTER, S., AND SOMPOLINSKY, H. Learning input correlations through nonlinear temporally asymmetric hebbian plasticity. *Journal of Neuroscience* 23, 9 (2003), 3697–3714.
- [66] HALE, J. K. *Ordinary differential equations*, second ed. Robert E. Krieger Publishing Co., Inc., Huntington, N.Y., 1980.
- [67] HEBB, D. O. *The organization of behavior: a neuropsychological theory*. J. Wiley; Chapman & Hall, 1949.
- [68] HELSON, P. A Mathematical Analysis of Memory Lifetime in a Simple Network Model of Memory. *Neural Computation* 32, 7 (2020), 1322–1354. PMID: 32433900.
- [69] HODGKIN, A. L., AND HUXLEY, A. F. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology* 117, 4 (1952), 500.
- [70] HOPFIELD, J. J. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences* 79, 8 (1982), 2554–2558. Publisher: National Academy of Sciences _eprint: <https://www.pnas.org/content/79/8/2554.full.pdf>.
- [71] HUANG, Y., AND AMIT, Y. Capacity analysis in multi-state synaptic models: a retrieval probability perspective. *Journal of Computational Neuroscience* 30, 3 (June 2011), 699–720.
- [72] IZHIKEVICH, E. M. *Dynamical systems in neuroscience: the geometry of excitability and bursting*. Computational Neuroscience. MIT Press, Cambridge, MA, 2007.
- [73] KANDEL, E., MACK, S., JESSELL, T., SCHWARTZ, J., SIEGELBAUM, S., AND HUDSPETH, A. *Principles of Neural Science, Fifth Edition*. McGraw-Hill’s AccessMedicine. McGraw-Hill Education, 2013.
- [74] KANDEL, E. R., DUDAI, Y., AND MAYFORD, M. R. The Molecular and Systems Biology of Memory. *Cell* 157, 1 (2014), 163 – 186.
- [75] KANG, H.-W., AND KURTZ, T. G. Separation of time-scales and model reduction for stochastic reaction networks. *Ann. Appl. Probab.* 23, 2 (2013), 529–583.

- [76] KECHRIS, A. S. *Classical descriptive set theory*, vol. 156 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995.
- [77] KEENER, J. P., HOPPENSTEADT, F., AND RINZEL, J. Integrate-and-fire models of nerve membrane response to oscillatory input. *SIAM Journal on Applied Mathematics* 41, 3 (1981), 503–517.
- [78] KEMPTER, R., GERSTNER, W., AND HEMMEN, J. L. v. Intrinsic Stabilization of Output Rates by Spike-Based Hebbian Learning. *Neural Computation* 13, 12 (2001), 2709–2741.
- [79] KEMPTER, R., GERSTNER, W., AND VAN HEMMEN, J. L. Hebbian learning and spiking neurons. *Phys. Rev. E* 59 (Apr 1999), 4498–4514.
- [80] KORTE, M., AND SCHMITZ, D. Cellular and System Biology of Memory: Timing, Molecules, and Beyond. *Physiological Reviews* 96, 2 (2016), 647–693. PMID: 26960344.
- [81] KURTZ, T. G. Averaging for martingale problems and stochastic approximation. In *Applied stochastic analysis (New Brunswick, NJ, 1991)*, vol. 177 of *Lect. Notes Control Inf. Sci.* Springer, Berlin, 1992, pp. 186–209.
- [82] LAHIRI, S., AND GANGULI, S. A memory frontier for complex synapses. In *Advances in neural information processing systems* (2013), pp. 1034–1042.
- [83] LANGILLE, J. J., AND BROWN, R. E. The Synaptic Theory of Memory: A Historical Survey and Reconciliation of Recent Opposition. *Frontiers in Systems Neuroscience* 12 (Oct. 2018), 52.
- [84] LAPICQUE, L. Recherches quantitatives sur l’excitation électrique des nerfs traitée comme une polarisation. *Journal de Physiologie et de Pathologie Générale* 9 (1907), 620–635.
- [85] LEIBOLD, C., AND KEMPTER, R. Memory capacity for sequences in a recurrent network with biological constraints. *Neural Computation* 18, 4 (2006), 904–941.
- [86] LEWIS, T. J., AND RINZEL, J. Dynamics of spiking neurons connected by both inhibitory and electrical coupling. *Journal of Computational Neuroscience* 14, 3 (2003), 283–309.
- [87] LINDNER, B., DOIRON, B., AND LONGTIN, A. Theory of oscillatory firing induced by spatially correlated noise and delayed inhibitory feedback. *Phys. Rev. E* (3) 72, 6 (2005), 061919, 14.
- [88] LINSKER, R. From basic network principles to neural architecture: Emergence of spatial-opponent cells. *Proceedings of the National Academy of Sciences* 83, 19 (1986), 7508–7512.
- [89] LITWIN-KUMAR, A., AND DOIRON, B. Formation and maintenance of neuronal assemblies through synaptic plasticity. *Nature Communications* 5, 1 (2014), 1–12.
- [90] LOUKIANOVA, D. Capacité de mémoire dans le modèle de Hopfield. *Comptes rendus de l’Académie des sciences. Série 1, Mathématique* 318, 2 (1994), 157–160.
- [91] LYNCH, G. S., DUNWIDDIE, T., AND GRIBKOFF, V. Heterosynaptic depression: a post-synaptic correlate of long-term potentiation. *Nature* 266, 5604 (1977), 737–739.
- [92] MARKRAM, H., LÜBKE, J., FROTSCHER, M., AND SAKMANN, B. Regulation of Synaptic Efficacy by Coincidence of Postsynaptic APs and EPSPs. *Science* 275, 5297 (1997), 213–215.

- [93] MASQUELIER, T., AND DECO, G. Learning and Coding in Neural Networks. In *Principles of Neural Coding*. CRC Press, May 2013, pp. 513–526.
- [94] MASQUELIER, T., GUYONNEAU, R., AND THORPE, S. J. Spike Timing Dependent Plasticity Finds the Start of Repeating Patterns in Continuous Spike Trains. *PLOS ONE* 3, 1 (2008), 1–9. Publisher: Public Library of Science.
- [95] MASQUELIER, T., GUYONNEAU, R., AND THORPE, S. J. Competitive STDP-based spike pattern learning. *Neural Computation* 21, 5 (2009), 1259–1276.
- [96] MENSHIKOV, M., POPOV, S., AND WADE, A. *Non-homogeneous random walks*, vol. 209 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2017. Lyapunov function methods for near-critical stochastic systems.
- [97] MEYN, S., AND TWEEDIE, R. L. *Markov chains and stochastic stability*, second ed. Cambridge University Press, Cambridge, 2009. With a prologue by Peter W. Glynn.
- [98] MILLER, A. *Neural network models for Recognition Memory*. PhD thesis, Hebrew University of Jerusalem, 2012.
- [99] MONGILLO, G., BARAK, O., AND TSODYKS, M. Synaptic Theory of Working Memory. *Science* 319, 5869 (Mar. 2008), 1543–1546.
- [100] MONGILLO, G., RUMPEL, S., AND LOEWENSTEIN, Y. Intrinsic volatility of synaptic connections — a challenge to the synaptic trace theory of memory. *Current Opinion in Neurobiology* 46 (2017), 7 – 13.
- [101] MORRIS, C., AND LECAR, H. Voltage oscillations in the barnacle giant muscle fiber. *Biophysical journal* 35, 1 (1981), 193–213.
- [102] MORRISON, A., DIESMANN, M., AND GERSTNER, W. Phenomenological models of synaptic plasticity based on spike timing. *Biological Cybernetics* 98, 6 (June 2008), 459–478.
- [103] NAGUMO, J., ARIMOTO, S., AND YOSHIKAWA, S. An active pulse transmission line simulating nerve axon. *Proceedings of the IRE* 50, 10 (1962), 2061–2070.
- [104] NEWMAN, C. M. Memory capacity in neural network models: Rigorous lower bounds. *Neural Networks* 1, 3 (1988), 223–238.
- [105] NORRIS, J. R. *Markov chains*, vol. 2 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1998. Reprint of 1997 original.
- [106] OCKER, G. K., AND DOIRON, B. Training and spontaneous reinforcement of neuronal assemblies by spike timing plasticity. *Cerebral Cortex* 29, 3 (2019), 937–951.
- [107] OCKER, G. K., LITWIN-KUMAR, A., AND DOIRON, B. Self-Organization of Microcircuits in Networks of Spiking Neurons with Plastic Synapses. *PLOS Computational Biology* 11, 8 (2015), 1–40. Publisher: Public Library of Science.
- [108] O’CONNOR, D. H., WITTENBERG, G. M., AND WANG, S. S.-H. Graded bidirectional synaptic plasticity is composed of switch-like unitary events. *Proceedings of the National Academy of Sciences of the United States of America* 102, 27 (2005), 9679–9684.

- [109] OSTOJIC, S., BRUNEL, N., AND HAKIM, V. How connectivity, background activity, and synaptic properties shape the cross-correlation between spike trains. *Journal of Neuroscience* 29, 33 (2009), 10234–10253.
- [110] PERTHAME, B., SALORT, D., AND WAINRIB, G. Distributed synaptic weights in a LIF neural network and learning rules. *Phys. D* 353/354 (2017), 20–30.
- [111] POO, M.-M., PIGNATELLI, M., RYAN, T. J., TONEGAWA, S., BONHOEFFER, T., MARTIN, K. C., RUDENKO, A., TSAI, L.-H., TSIEN, R. W., FISHELL, G., MULLINS, C., GONÇALVES, J. T., SHTRAHMAN, M., JOHNSTON, S. T., GAGE, F. H., DAN, Y., LONG, J., BUZSÁKI, G., AND STEVENS, C. What is memory? The present state of the engram. *BMC Biology* 14, 1 (Dec. 2016).
- [112] REDONDO, R. L., AND MORRIS, R. G. M. Making memories last: the synaptic tagging and capture hypothesis. *Nature Reviews Neuroscience* 12, 1 (Jan. 2011), 17–30.
- [113] REVUZ, D., AND YOR, M. *Continuous martingales and Brownian motion*, third ed., vol. 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999.
- [114] ROBERT, P., AND VIGNOUD, G. Stochastic Models of Neural Plasticity : Averaging Principles. *arXiv preprint arXiv:2010.08790* (2020).
- [115] ROBERT, P., AND VIGNOUD, G. Stochastic Models of Neural Synaptic Plasticity. *arXiv preprint arXiv:2010.08195* (2020).
- [116] ROMANI, S., AMIT, D. J., AND AMIT, Y. Optimizing one-shot learning with binary synapses. *Neural Computation* 20, 8 (2008), 1928–1950.
- [117] ROXIN, A., AND FUSI, S. Efficient Partitioning of Memory Systems and Its Importance for Memory Consolidation. *PLOS Computational Biology* 9, 7 (2013), 1–13. Publisher: Public Library of Science.
- [118] RUBIN, J., LEE, D. D., AND SOMPOLINSKY, H. Equilibrium properties of temporally asymmetric hebbian plasticity. *Phys. Rev. Lett.* 86 (Jan 2001), 364–367.
- [119] SOMMER, F. T., AND DAYAN, P. Bayesian retrieval in associative memories with storage errors. *IEEE Transactions on Neural Networks* 9, 4 (1998), 705–713.
- [120] SONG, S., MILLER, K. D., AND ABBOTT, L. F. Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. *Nature Neuroscience* 3, 9 (2000), 919–926.
- [121] STANDING, L. Learning 10000 pictures. *Quarterly Journal of Experimental Psychology* 25, 2 (May 1973), 207–222.
- [122] SUSMAN, L., BRENNER, N., AND BARAK, O. Stable memory with unstable synapses. *Nature Communications* 10, 1 (2019), 1–9.
- [123] SZNITMAN, A.-S. Topics in propagation of chaos. In *École d’Été de Probabilités de Saint-Flour XIX—1989*, vol. 1464 of *Lecture Notes in Math*. Springer, Berlin, 1991, pp. 165–251.
- [124] TALAGRAND, M. Rigorous results for the Hopfield model with many patterns. *Probab. Theory Related Fields* 110, 2 (1998), 177–276.
- [125] TITLEY, H. K., BRUNEL, N., AND HANSEL, C. Toward a neurocentric view of learning. *Neuron* 95, 1 (2017), 19–32.

- [126] TROUSDALE, J., HU, Y., SHEA-BROWN, E., AND JOSIĆ, K. Impact of Network Structure and Cellular Response on Spike Time Correlations. *PLOS Computational Biology* 8, 3 (2012), 1–15. Publisher: Public Library of Science.
- [127] TSODYKS, M. V., AND FEIGEL'MAN, M. V. The enhanced storage capacity in neural networks with low activity level. *EPL (Europhysics Letters)* 6, 2 (1988), 101.
- [128] TURRIGIANO, G. G. The dialectic of Hebb and homeostasis. *Philosophical Transactions of the Royal Society B: Biological Sciences* 372, 1715 (Mar. 2017), 20160258.
- [129] VAN ROSSUM, M. C. W., SHIPPI, M., AND BARRETT, A. B. Soft-bound Synaptic Plasticity Increases Storage Capacity. *PLOS Computational Biology* 8, 12 (2012), 1–11. Publisher: Public Library of Science.
- [130] WANG, H.-X., GERKIN, R. C., NAUEN, D. W., AND BI, G.-Q. Coactivation and timing-dependent integration of synaptic potentiation and depression. *Nature Neuroscience* 8, 2 (Feb. 2005), 187–193.
- [131] WIKIPEDIA. Action potential — Wikipedia, the free encyclopedia. <http://en.wikipedia.org/w/index.php?title=Action%20potential&oldid=998242198>. [Online; accessed 20-January-2021].
- [132] WIKIPEDIA. Synapse — Wikipedia, the free encyclopedia. <http://en.wikipedia.org/w/index.php?title=Synapse&oldid=1001008447>. [Online; accessed 20-January-2021].
- [133] WILLSHAW, D. J., BUNEMAN, O. P., AND LONGUET-HIGGINS, H. C. Non-holographic associative memory. *Nature* 222, 5197 (1969), 960–962.
- [134] YGER, P., AND GILSON, M. Models of Metaplasticity: A Review of Concepts. *Frontiers in Computational Neuroscience* 9 (Nov. 2015).
- [135] ZENKE, F. *Memory formation and recall in recurrent spiking neural networks*. PhD thesis, Ecole Polytechnique Federale de Lausanne, 2014.
- [136] ZENKE, F., GERSTNER, W., AND GANGULI, S. The temporal paradox of Hebbian learning and homeostatic plasticity. *Current Opinion in Neurobiology* 43 (2017), 166 – 176.
- [137] ZENKE, F., HENNEQUIN, G., AND GERSTNER, W. Synaptic Plasticity in Neural Networks Needs Homeostasis with a Fast Rate Detector. *PLOS Computational Biology* 9, 11 (2013), 1–14. Publisher: Public Library of Science.