



**HAL**  
open science

# Apprentissage et Contrôle dans les Architectures Neurales

Emmanuel Daucé

► **To cite this version:**

Emmanuel Daucé. Apprentissage et Contrôle dans les Architectures Neuronales. Réseau de neurones [cs.NE]. Aix-Marseille Université, 2016. tel-01264905

**HAL Id: tel-01264905**

**<https://theses.hal.science/tel-01264905>**

Submitted on 29 Jan 2016

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

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



Distributed under a Creative Commons Attribution 4.0 International License

Mémoire présenté en vue de l'obtention  
du diplôme d'Habilitation à Diriger les Recherches

Spécialité Informatique

# APPRENTISSAGE ET CONTRÔLE DANS LES ARCHITECTURES NEURONALES

par

Emmanuel DAUCÉ

Theoretical Neurosciences Group, Institut de Neurosciences des Systèmes  
INSERM UMR\_S 1106

et

Ecole Centrale de Marseille

Travaux présentés le 18 Janvier 2016 devant un Jury composé de :

Dr. Frédéric ALEXANDRE	INRIA Bordeaux Sud Ouest (rapporteur)
Dr. Benoît GIRARD	ISIR - Univ. Paris VI (rapporteur)
Dr. Jean-Pierre NADAL	LPS - Ecole Normale Supérieure, Paris (rapporteur)
Pr. Manuel SAMUELIDES	Institut Supérieur de l'Aéronautique et de l'Espace, Toulouse
Dr. Rémi MUNOS	Google DeepMind, Londres
Pr. Liva RALAIVOLA	LIF, Aix-Marseille Université
Pr. Laurent PÉZARD	LNIA, Aix-Marseille Université (président)

Ecole Doctorale en Mathématiques et Informatique de Marseille (ED 184)

Faculté de Sciences  
Aix-Marseille Université, France

## Remerciements

Je remercie Liva RALAIVOLA, Professeur d'Informatique à la Faculté de Sciences d'Aix-Marseille Université pour avoir accepté d'endosser le rôle de tuteur pour ce travail, pour son soutien et sa relecture minutieuse. Merci également à Bruno TORRESANI, Professeur de Mathématiques à la Faculté de Sciences d'Aix-Marseille Université, et rapporteur interne de ce document, pour son exigence critique et ses conseils rédactionnels.

J'adresse toute ma reconnaissance à mes trois rapporteurs, Frédéric ALEXANDRE, Directeur de Recherche à l'INRIA Bordeaux-Sud Ouest, Benoît GIRARD, Directeur de Recherche à l'Institut des Systèmes Intelligents et de Robotique, et Jean-Pierre NADAL, Directeur de Recherche au Laboratoire de Physique Statistique de l'École Normale Supérieure, et Directeur d'Études à l'École des Hautes Études en Sciences Sociales, pour leur disponibilité et leurs encouragements. Je remercie également Manuel SAMUELIDÈS, Professeur de mathématiques à l'Institut Supérieur de l'Aéronautique et de l'Espace, Rémi MUNOS, Directeur de Recherche à l'INRIA Lille Nord-Europe, actuellement en détachement à Google DeepMind à Londres, et Laurent PÉZARD, Professeur en Neurosciences à la Faculté de Sciences d'Aix-Marseille Université, pour leur participation à l'animation de ce jury d'HDR.

Les travaux présentés ici sont le fruit de nombreuses collaborations et interactions pluridisciplinaires au sein de différents groupes de travail, ACI, ANR, conférences, rencontres, ateliers et réunions informelles dans les différents instituts où j'ai effectué mes recherches. Les collègues, maîtres et amis qui ont participé à mes travaux sont cités dans le document et il est impossible d'établir la liste de tous ceux dont les discussions, recherches et travaux ont permis d'en tisser les fils. Je tiens donc à remercier ceux qui, en mettant à disposition les conditions matérielles de la recherche, rendent possible son développement, ses bifurcations et ses ramifications nouvelles, et en particulier, pour leur accueil et leur ouverture pluridisciplinaire, Reinoud BOOTSMA, ancien chef de l'équipe Dynamique du Mouvement et Directeur de l'École Doctorale de Sciences du Mouvement Humain, et Jean-Louis VERCHER, ancien Directeur de l'Institut des Sciences du Mouvement, unité mixte CNRS rattachée à la Faculté des Sciences du Sport d'Aix-Marseille Université, ainsi que Viktor JIRSA, Directeur de l'Institut de Neurosciences des Systèmes, unité mixte INSERM rattachée à la Faculté de Médecine d'Aix-Marseille Université.

Je remercie bien sûr mes collègues et confrères de l'Institut de Neurosciences de Systèmes, ainsi que ceux de l'équipe pédagogique Mathématiques et Informatique de l'École Centrale de Marseille, pour leur bonne humeur au quotidien et leurs conseils dans les derniers préparatifs, et merci bien sûr à ma famille et à mes amis pour la joie et le soutien de tous les jours.

---

## Résumé

Ce mémoire d'Habilitation à Diriger les Recherches présente mes contributions à la recherche en modélisation mathématique et informatique des dynamiques neuronales et de l'apprentissage. Le but de ces travaux est de mieux comprendre la logique à l'œuvre au sein du système nerveux des êtres vivants, afin d'améliorer la connaissance du cerveau et des pathologies nerveuses. Il s'agit également de s'inspirer du cerveau pour repenser la manière de concevoir des objets communicants et agissants (robots mobiles, interfaces vocales, jeux, etc.), afin qu'ils interagissent de manière plus "naturelle" avec leur environnement.

Les modèles de réseaux de neurones étudiés sont des systèmes dynamiques de grande dimension, définis par un graphe d'interactions entre des unités de calcul élémentaires. Nous étudions les principes du calcul "collectif" dans ces architectures, tels que le "temps de calcul" corresponde au temps de relaxation vers un *attracteur*. Les comportements d'ensemble sont alors interprétés comme des implémentations d'opérateurs plus simples. Nous montrons en particulier que l'utilisation d'un graphe de connexions aléatoire permet, en régime critique, de structurer l'espace des entrées sensorielles en régions distinctes, selon un principe proche de l'"auto-encodage".

L'apprentissage est abordé dans une perspective neuro-réaliste, c'est-à-dire prenant en compte les principales contraintes biologiques, comme le caractère local des changements synaptiques, le caractère récurrent du graphe de connexions, ou encore le caractère discret des signaux émis par les neurones. Les tâches d'apprentissage étudiées impliquent la détection de causalités temporelles au sein des signaux présentés en entrée du réseau, ou encore la prise en compte des causalités circulaires entre le réseau et son environnement, sur des tâches de contrôle moteur dites "en boucle fermée". Nous montrons dans ce cadre la possibilité d'une reconnaissance de formes reposant sur un principe de *rétroaction positive* (ou résonance) entre le signal présenté en entrée et un "réservoir" de formes apprises. Dans le cadre du contrôle moteur, nous montrons également comment la plasticité de Hebb, dans sa variante guidée, correspond à la recherche du maximum de vraisemblance dans un espace de couples (perception, action), en l'absence de modèle de l'environnement.

Le travail de relecture et de compilation, occasionné par la rédaction de ce mémoire, a également permis de mettre en lumière certaines difficultés fondamentales, qui freinent les possibilités d'apprentissage de lois de contrôle dans un espace d'actions continu. En particulier, si l'on souhaite construire des dispositifs artificiels capable d'apprendre des tâches motrices de la même manière qu'ils apprennent à classifier des signaux et des images, il faut établir des règles de contrôle *ne nécessitant pas* de comparaison entre les grandeurs de l'espace environnant. Nous proposons, dans ce cadre, de s'inspirer du principe de "contrôle de la position finale", suggéré par les neurosciences, par opposition au "contrôle du déplacement" utilisé en théorie du contrôle classique. Différents projets sont proposés, visant à développer ces idées sur des modèles de l'activité du cerveau à large échelle, ou encore dans le cadre des interfaces cerveau-machine.

## Abstract

This “Habilitation à Diriger les Recherches” dissertation presents my research contributions to mathematics and computer science modelling of neuronal dynamics and learning. The aim of this work is to better understand the logic inside the nervous system of the living beings, in order to improve the knowledge of the brain and nervous pathologies. Another aim is to take inspiration from the brain in order to reconsider the design of communicating tools and active devices (mobile robots, vocal interfaces, games, etc.) for they interact in a more “natural” way with their environment.

The neural network models addressed here are high dimensional dynamical systems, defined by the graph of couplings between elementary computational units. We study the principles of “collective” computation in such networks, so that the “computing time” corresponds to the relaxation time toward an *attractor*. The collective behaviors are then interpreted as the implementation of simpler operations. We show in particular how random connection graphs allow, in the critical regime, to separate the sensory input space into distinct regions, in a way close to the “auto-encoding” principle.

Learning is addressed in a neuro-realistic perspective, i.e. taking into account the main biological constraints, like the locality of the synaptic changes, the recurrent organization of the graph, or the discreteness of the neuronal signals. The learning tasks we consider imply the detection of temporal causalities within the input signals, or also to take into account the circular causalities between the network and its environment, on “closed loop” motor control tasks. We consider the possibility of a pattern recognition principle based on the *positive feedback* (or resonance) between the input signal and a “reservoir” of learned features. In the motor control framework, we also show how the Hebb plasticity, in its guided variant, corresponds to the search for the maximum likelihood in a space of (perception, action) couples, in the absence of an environment model.

The proofreading work and papers compilation, needed by the writing of this dissertation, also allowed to shed light on some fundamental impediments in learning control laws in continuous action state spaces. In particular, if one wants to build artificial devices capable to learn motor tasks the same way they learn to classify signals and images, one needs to establish control rules that *do not necessitate* comparisons between quantities of the surrounding space. We propose, in that context, to take inspiration from the “end effector control” principle, as suggested by neuroscience studies, as opposed to the “displacement control” principle used in the classical control theory. Several projects are proposed, aiming at expanding some of those ideas into large-scale brain activity models, or also for the design of brain-computer interfaces.

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Neurosciences computationnelles . . . . .	1
1.1.1	Une approche pluridisciplinaire . . . . .	2
1.1.2	Développement . . . . .	3
1.1.3	Principales avancées . . . . .	5
1.2	Parcours personnel . . . . .	7
1.3	Organisation du document . . . . .	14
<b>I</b>	<b>Vue d'ensemble</b>	<b>15</b>
<b>2</b>	<b>Préambule</b>	<b>17</b>
2.1	Neurons et réseaux . . . . .	17
2.2	Traitement séquentiel et traitement récurrent . . . . .	21
2.2.1	Traitement séquentiel . . . . .	21
2.2.2	Traitement récurrent . . . . .	22
2.3	Notre approche . . . . .	23
2.3.1	Séparation des signaux . . . . .	24
2.3.2	Plasticité . . . . .	25
2.3.3	Assemblées neuronales . . . . .	26
2.4	Exemples . . . . .	27
2.4.1	Réseau de Hopfield . . . . .	27
2.4.2	Réseaux récurrents aléatoires . . . . .	28
2.4.3	Champ neuronal . . . . .	29
<b>3</b>	<b>Apprentissage et contrôle</b>	<b>31</b>
3.1	Apprentissage et plasticité . . . . .	31
3.1.1	Apprentissage automatique . . . . .	32
3.1.2	Lien entre le computationnel et le biologique . . . . .	35
3.2	Systèmes dynamiques et contrôle . . . . .	36
3.2.1	Couplages sensori-moteurs . . . . .	36
3.2.2	Approche probabiliste . . . . .	38
3.3	Approche mixte : modèles et problèmes . . . . .	40
3.3.1	Approche probabiliste . . . . .	40
3.3.2	Modèles bio-inspirés . . . . .	42

<b>4</b>	<b>Panorama des principaux résultats</b>	<b>45</b>
<b>5</b>	<b>Bilan et projets</b>	<b>57</b>
5.1	Bilan . . . . .	57
5.1.1	Résumé . . . . .	57
5.1.2	Discussion . . . . .	58
5.2	Perspectives et projets . . . . .	62
5.2.1	Construction d'un espace métrique péripersonnel . . . . .	63
5.2.2	Vers une approche computationnelle de l'activité à large échelle . . . . .	66
5.2.3	Interfaces informatiques inspirées par la biologie . . . . .	68
<b>II</b>	<b>Résultats détaillés</b>	<b>71</b>
<b>6</b>	<b>Dynamique des grands réseaux de neurones aléatoires</b>	<b>73</b>
6.1	Réseaux de neurones récurrents et modélisation biologique . . . . .	73
6.2	Contributions personnelles . . . . .	76
6.2.1	Réseaux aléatoires multi-populations . . . . .	76
6.2.2	Criticalité dans les réseaux aléatoires . . . . .	78
6.2.3	Modèles de champ neuronal . . . . .	81
6.2.4	Extension aux populations de neurones impulsionnels . . . . .	83
6.2.5	Modèles de la connectivité fonctionnelle large-échelle . . . . .	84
<b>7</b>	<b>Plasticité synaptique et apprentissage</b>	<b>87</b>
7.1	Modèles de la perception . . . . .	87
7.2	Contributions personnelles . . . . .	88
7.2.1	Route inverse . . . . .	88
7.2.2	Architectures multi-couches . . . . .	91
7.2.3	Plasticité dans les populations de neurones impulsionnels . . . . .	94
7.2.4	Discrimination de séquences spatio-temporelles . . . . .	99
<b>8</b>	<b>Architectures de contrôle</b>	<b>101</b>
8.1	Contrôleur et environnement . . . . .	101
8.2	Contributions personnelles . . . . .	106
8.2.1	Alternance comportementale en robotique autonome . . . . .	106
8.2.2	Contrôle adaptatif en boucle fermée . . . . .	108
8.2.3	Codage topographique et renforcement . . . . .	110
8.2.4	Modèles biologiques de l'orientation . . . . .	115
8.2.5	Interfaces cerveau-machine . . . . .	119
	<b>Bibliographie</b>	<b>120</b>

---

<b>III Annexes</b>	<b>141</b>
A Self-Organization and dynamics reduction in recurrent networks : stimulus presentation and learning (1998)	143
B Mean-Field theory and Synchronization in Random Recurrent Neural Networks (2001)	159
C Resonant spatiotemporal learning in large random recurrent networks (2002)	167
D Short Term Memory in Recurrent Networks of Spiking Neurons (2004)	183
E Hebbian reinforcement learning in a modular dynamic network (2004)	197
F Temporal Pattern identification using Spike-Timing Dependent Plasticity (2007)	209
G Learning an Control with large Dynamic Neural Networks (2007)	219
H Spike-Timing dependent plasticity and regime transitions in random recurrent neural networks (2008)	261
I A Model of Neuronal Specialization Using Hebbian Policy-Gradient with "Slow" Noise (2009)	269
J Online Learning with Noise : A Kernel-Based Policy-Gradient Approach (2010)	277
K How and where does the brain predict the when : a Bayesian approach to modeling temporal expectation (2011)	285
L Simple spatio-temporal transformation with sub-threshold integration in the saccadic system (2011)	289
M Toward STDP-based population action in large networks of spiking neurons (2014)	293
N Evidence build-up facilitates on-line adaptivity in dynamic environments : example of the BCI P300-speller (2014)	299
O Reward-based online learning in non-stationary environments : adapting a P300-speller with a "backspace" key (2015)	305
P Multistability in large-scale models of brain activity	315



# 1

## Introduction

Le but de ce document est de fournir un aperçu détaillé des questions et des approches que j'ai eu l'occasion d'aborder au cours des vingt dernières années, depuis mes premiers pas en recherche. Ce document est organisé autour des différentes contributions et productions scientifiques auxquelles j'ai participé, dans un ordre thématique. Il fournit également un aperçu d'un certain nombre de questions restées ouvertes, et pouvant faire l'objet de développements futurs.

Ces différentes contributions appartiennent à un champ de recherche assez large à l'intersection des neurosciences et de l'informatique, généralement appelé “*neurosciences computationnelles*”. Le terme computationnel fait référence à la notion de calcul. On peut donc également parler de neuro-calcul. *Les neurosciences computationnelles visent à identifier ce qui, dans l'activité des neurones, s'apparente à un calcul; à identifier le type et la nature des opérations réalisées par le système nerveux.*

### 1.1 Neurosciences computationnelles

Si les principes de fonctionnement élémentaires du système nerveux central (neurones, synapses) commencent à être bien connus, le fonctionnement intégré du système dans son ensemble reste l'objet de nombreuses conjectures. Les modèles mathématiques et les simulations informatiques permettent de proposer des hypothèses de fonctionnement et/ou de valider certaines hypothèses issues de l'observation.

Ainsi, l'étude du système nerveux, sous l'angle computationnel, a deux objectifs :

1. Explorer des mécanismes de calcul alternatifs
2. Comprendre le fonctionnement du cerveau

**Mécanismes de calcul alternatifs** Le premier objectif est de s'inspirer du fonctionnement du système nerveux pour découvrir de nouveaux principes pour le calcul artificiel, ou de nouvelles architectures, pour résoudre des problèmes qui résistent encore aux calculateurs actuels. Mieux comprendre le cerveau, c'est donc potentiellement étendre le domaine de la science informatique, étendre le type de données qu'elles sont capables de traiter (son, langage, images, ...), augmenter leur capacité à contrôler leur environ-

nement, les rendre plus proches, en terme de fonctionnement, des humains avec lesquels elles interagissent.

**Fonctionnement du cerveau** Le deuxième objectif consiste à utiliser les connaissances informatiques et mathématiques actuelles pour mieux comprendre le système nerveux. Il s’agit d’une certaine façon d’identifier ce qui dans l’activité nerveuse s’apparente aux algorithmes utilisés dans des tâches computationnelles variées, comme le contrôle, l’apprentissage, le traitement des images et des données ; voir si les solutions mises au point dans des contextes d’ingénierie proches de celui face auquel se trouve le système nerveux ont des équivalents au sein de l’activité nerveuse (ou si, au contraire, la nature utilise des principes différents). L’objectif est, à partir des observations, de mieux identifier les rôles et les interactions à l’intérieur du système nerveux, et potentiellement de mieux soigner les désordres et maladies neurologiques.

### 1.1.1 Une approche pluridisciplinaire

Les neurosciences computationnelles sont par définition pluridisciplinaires :

- Elles reposent en premier lieu sur le champ des *neurosciences*, qui ont connu un développement considérable au cours des quarante dernières années, avec l’apparition de nouvelles techniques de mesure permettant d’observer le cerveau en fonctionnement à différentes échelles.
- Elles reposent en second lieu sur la *modélisation mathématique et informatique*, qui a également connu un développement important lié à l’augmentation des capacités de calcul des ordinateurs, ainsi que de la capacité à traiter des données massives grâce à l’apprentissage automatique (*Machine learning*).
- Un dernier champ de recherche, celui de la *robotique autonome* et du contrôle moteur, est également impliqué, dans la mesure où le développement d’interfaces bioniques nécessite une meilleure connaissance des mécanismes de contrôle à l’œuvre chez l’homme et l’animal.

**Trois niveaux d’analyse** Selon la typologie classique proposée par David Marr (Marr, 1982), un modèle peut appartenir à différents niveaux de généralité, à savoir le niveau computationnel (la logique), le niveau algorithmique (encodage et opérations réalisées), et la réalisation matérielle (le réseau de neurones ou “substrat”).

Cette distinction en trois niveaux permet principalement de distinguer ce qui relève de la définition du problème : définition de la tâche, en fonction des contraintes physiques et de l’appareillage sensoriel et moteur, d’une part, et d’autre part ce qui relève de la réalisation (algorithmique) de la tâche, c’est-à-dire principalement le circuit de traitement. C’est, en terminologie informatique, la distinction entre l’équation à résoudre, la brique logicielle qui la résout et le circuit logique qui l’implémente. Marr postule une relative indépendance entre les différents niveaux de description, en établissant une claire séparation entre le substrat neuronal, le circuit de traitement et l’espace de la tâche.

**Approche ascendante (ou “bottom-up”)** La méthodologie “bottom-up” consiste à étudier les *conditions d’émergence* de fonctions de haut niveau à partir des caractéris-

tiques d'un réseau de neurones plongé dans un environnement contraignant. Les réseaux de neurones sont en effet capables de manifester des formes d'organisation spontanée : comportement critique, transitions entre différentes configurations d'activité, activité persistante et mémoire à court terme (rémanence) etc. L'ajout de mécanismes de plasticité synaptique vise à exploiter cette richesse intrinsèque pour que le réseau puisse acquérir des connaissances et des compétences nouvelles.

**Approche descendance (ou “top-down”)** La méthodologie “top-down” consiste à étudier les *conditions de réalisation* d'une certaine tâche par étapes descendantes, de la contrainte générale du corps engagé dans la tâche à la réalisation matérielle du circuit qui la résout. L'approche descendante consiste souvent à identifier les aspects non-triviaux d'une tâche. L'étude des temps de réponse, des courbes d'apprentissage et des taux d'erreur permet d'explorer les caractéristiques limites d'un comportement et de mieux appréhender les contraintes posées par le système qui l'implémente physiquement. Il peut s'agir des contraintes physiques (limites d'extension, de course, etc.) mais également des limites provenant des caractéristiques physiques du réseau de neurones sous-jacent. L'élaboration d'un modèle descriptif prenant en compte plusieurs échelles permet alors d'émettre des hypothèses sur les sources (physiques ou logiques) des limitations comportementales observées.

**Approche “Multi-échelles”** Il importe dans tous les cas de prendre en compte la différence de nature (ou l'hétérogénéité) entre d'une part le milieu extérieur, caractérisé par des grandeurs extensives (des masses et des articulations en déplacement, des changements continus) et le milieu intérieur (milieu logiciel) caractérisé par des activités bornées, discrètes, changeant rapidement et de manière abrupte. Entre les deux, un niveau architectural (dit “mésoscopique”) doit également être considéré, constitué de circuits imbriqués intégrant l'activité de larges portions du cerveau.

Chacun de ces niveaux de description est traditionnellement considéré de manière séparée, et l'effort consistant à traduire un des niveaux de description dans un autre, effort ayant pris une place importante dans les neurosciences computationnelles, s'appelle l'approche “multi-échelles” (Wilson & Bower, 1989; Varela et al., 2001). Cette approche utilise le formalisme de la physique statistique pour décrire un “milieu” neuronal dont les variables d'état évoluent (ou peuvent être décrites) selon différentes échelles spatiales et temporelles, chaque échelle de description étant explicative du comportement observé à l'échelle supérieure (par des variables dites “macroscopiques”).

### 1.1.2 Développement

Les neurosciences computationnelles se sont fortement développées, et ont commencé à s'organiser en tant que champ de recherche, à partir de la fin des années 90, avec le soutien de plusieurs programmes nationaux et internationaux. On citera, parmi les plus actifs, le réseau Bernstein pour les neurosciences computationnelles en Allemagne, développé depuis 2004, le collectif INCF (International Neuroinformatics Coordination Facility) qui tente de coordonner les nombreux projets de recherche de la discipline au niveau international (basé en Suède). Elles ont bénéficié de nombreux soutiens financiers

européens (NEST, Brain scales, Blue Brain project, ...) qui se sont récemment concrétisés par l'attribution d'un financement 'Flagship' pour le projet "Human Brain Project" qui regroupe les principaux acteurs des projets précédents. On notera également l'initiative BRAIN (Brain Research through Advancing Innovative Neurotechnologies) pilotée par la NIH et la NSF aux Etats-Unis. Ce champ de recherche commence également à s'insérer dans les programmes de deuxième cycle, comme le montrent les nombreux cours en ligne disponibles sur le Web (voir par exemple ceux de Sébastien Seung au MIT, Andrew Ng à Stanford, Rajesh Rao à l'Université de Washington, etc.)

Au niveau des revues scientifiques, les neurosciences computationnelles ne bénéficient pas de la même couverture académique que les disciplines plus anciennement implantées. Les revues qui ont historiquement participé à son développement sont les revues traitant des réseaux de neurones artificiels apparues à la fin des années 80 (Neural Networks, Neural Computation, Neural Computing Letters, Biological Cybernetics, etc.). Plusieurs revues de neurosciences, comme Neuroimage ou J. Neurosciences, acceptent également des contributions du domaine. Enfin, les revues de physique (Physica-D) hébergent également des articles traitant des neurosciences computationnelles. Suite au développement du champ de recherche et de sa couverture relativement lacunaire par les revues scientifiques, sont apparues des revues en "Open Acces" spécialement dédiées au domaine. On notera en particulier PLOS-Computational Biology et Frontiers in Computational Neurosciences qui ont aidé à compenser le relatif manque de visibilité du domaine. Il existe enfin quelques conférences spécifiquement dédiées à ce champ de recherche, comme la conférence CNS (Computational Neurosciences), la conférence CoSyNe (Computational and Systems Neuroscience) ainsi que la conférence SAB (Simulation of Adaptive Behavior) dans le domaine de la robotique bio-inspirée.

**Les neurosciences computationnelles en France** Les neurosciences computationnelles ont émergé en France dans le cadre des programmes de recherche interdisciplinaires de sciences cognitives montés dans les années 90 à l'intersection des sections 27 (Neurosciences) et 34 (Linguistique) du CNRS, ainsi que différentes actions concertées incitatives (ACI "cognitive" en 1999) du ministère de la Recherche. Ces soutiens institutionnels ont permis la constitution d'une communauté des Sciences Cognitives, avec l'apparition de plusieurs instituts et de masters de sciences cognitives. Les neurosciences computationnelles ont ensuite reçu un soutien spécifique avec l'ACI "Neurosciences intégratives et computationnelles" en 2001, achevée en 2007, qui a vu l'émergence de nouvelles questions dans le champ des neurosciences : notion de connectivité fonctionnelle, approche "multi-échelles", dynamiques de populations et codage par populations (Faugeras et al., 2007)... La mise en place de l'agence nationale de la recherche en 2005 a rebattu les cartes, avec un système de financement centré sur des projets non-thématiques. En parallèle avec l'apparition des grands programmes internationaux mentionnés, la communauté s'est auto-organisée, avec des initiatives nationales comme le groupe "Neurocomp" qui a tenté de fédérer la communauté française de neurosciences computationnelles en organisant des événements scientifiques, et les conférences "Ladislav Tauc", sur des thématiques à l'interface des neurosciences et des systèmes dynamiques.

Si on constate à l’heure actuelle une baisse relative de soutien institutionnel national<sup>1</sup>, la communauté des neurosciences computationnelles reste relativement stable avec une représentation assez équilibrée entre des laboratoires de sciences de la vie, de physique théorique, d’informatique et de robotique, mais continuent de bénéficier globalement des budgets importants alloués aux neurosciences en général (cf. le projet “Neurospin”) et des appels d’offres internationaux.

### 1.1.3 Principales avancées

Le champ de recherche s’organise autour de plusieurs pôles.

1. Le premier pôle concerne les simulations neuro-réalistes. Le but est de mieux comprendre le fonctionnement du neurone, et des interactions entre neurones. Ces modèles étudient en détail (i) la structure de la cellule neuronale, (ii) les échanges entre la cellule et son milieu, via les synapses et l’émission de potentiels d’action, et enfin (iii) la plasticité et les changements de ces échanges au cours du temps.

Les principales avancées dans ce cadre sont :

- Le modèle Hodgkin et Huxley (Hodgkin & Huxley, 1952) et les modèles à compartiments (Rall, 1959) qui servent encore à l’heure actuelle de fondement pour les modèles neuro-réalistes, aboutissant aux modèles détaillés de colonnes corticales constituées de dizaines de milliers de cellules (Binzegger et al., 2004; Markram, 2006; Markram, 2012).
  - L’étude de la plasticité synaptique (Bliss & Lomo, 1973; Bi & Poo, 2001) et des mécanismes cellulaires associés (Harris et al., 1984; Shouval et al., 2002; O’Donnell et al., 2011; Graupner & Brunel, 2012; Cui et al., 2015) conduit à mieux comprendre la sélectivité neuronale et la formation de champs récepteurs (Bienenstock et al., 1982; Oja, 1982; Van Rossum et al., 2000; Sen Song, 2000; Billings & van Rossum, 2009) ;
2. Le second pôle concerne l’étude des dynamiques de populations de neurones. Les neurones sont dans ce cadre des unités de traitement simplifiées (neurones à spikes, neurones stochastiques, neurones fréquentiels,...). Le but est de comprendre les interactions entre de grands groupes de neurones, et d’étudier les conditions d’émergence de certaines signatures d’activité (corrélations large échelle, couplages de phase, champs récepteurs, ...) observés en électrophysiologie chez le sujet sain et dans les pathologies nerveuses.
    - Les neurones travaillent dans un régime dit “sous-critique” (Beggs & Plenz, 2004; Petermann et al., 2009; Rubinov et al., 2011; Haimovici et al., 2013), encore appelé régime de “haute conductance” (Shadlen & Newsome, 1998; Destexhe et al., 2003), qui explique à la fois le caractère très irrégulier de leur activité (Dean, 1981; Softky & Koch, 1993; Van Vreeswijk & Sompolinsky, 1998), et leur tendance à se synchroniser à courte (Gray et al., 1989; Wang & Buzsáki, 1996; Brunel, 2000a; Wang, 2010) et longue distance (Rodriguez et al., 1999; Palva et al., 2005; Busch et al., 2009).

---

1. Disparition en 2011 de la commission interdisciplinaire du CNRS “Cognition, langage, traitement de l’information, systèmes naturels et artificiels” historiquement liée aux sciences cognitives.

- Les populations de neurones peuvent conserver une activité soutenue en l’absence de stimulation par des mécanismes de réverbération internes (Hebb, 1949; Amari, 1977; Hopfield, 1982; Funahashi et al., 1989; Camperi & Wang, 1998; Brunel & Wang, 2001; Lansner, 2009).
3. Enfin, le troisième pôle concerne les modèles conceptuels, principes de traitement de l’information, qui expliquent la réception et la transmission de l’information nerveuse, les codes neuronaux, les mécanismes de la mémoire et de l’apprentissage ;

Les principales avancées dans ce cadre sont :

- Les modèles du système visuel, inspirés des observations de Hubel et Wiesel (Hubel & Wiesel, 1962), mettent en évidence une organisation hiérarchique des couches de traitement visuelles, dont la formation est modélisée par des algorithmes d’auto-encodage non-supervisés (Olshausen & Field, 1997; Lee & Seung, 1999; Rao & Ballard, 1999; Hinton et al., 2006; LeCun et al., 2015).
- Les principaux sous-systèmes du système nerveux central ont fait l’objet de modélisations spécifiques, comme le complexe hippocampe-cortex entorhinal pour la mémoire et la navigation (O’Keefe & Nadel, 1978; Treves & Rolls, 1994; Skaggs et al., 1996; Buzsáki & Moser, 2013), le cervelet pour le contrôle adaptatif (Wolpert et al., 1998; Haruno et al., 2001; Dean et al., 2010) et le complexe cortex-ganglions de la base-thalamus pour la sélection de l’action (Schultz et al., 1997; Gurney et al., 2001; Joel et al., 2002).
- Le principe du codage par population, inspiré par l’activité des régions prémotrices et motrices du cortex (Georgopoulos et al., 1986) et les champs récepteurs à l’orientation de V1 (Bonhoeffer et al., 1991), permet d’interpréter l’activité irrégulière des populations de neurones comme un échantillonnage de grandeurs continues “latentes” (Pouget & Sejnowski, 1997; Pouget et al., 2002; Fiser et al., 2010; Stewart et al., 2011; Boerlin et al., 2013).

De nombreuses bibliothèques et environnements de simulation ont par ailleurs été développés, de la cellule jusqu’aux populations de neurones à large échelle (Eliasmith & Anderson, 2004; Brette et al., 2007; Furber et al., 2013; Leon et al., 2013).

Enfin, le domaine connaît actuellement une forte tentative d’unification autour de principes issus de l’inférence statistique. Le cerveau construirait une interprétation de son environnement perceptif à l’aide d’un petit nombre de variables explicatives, ou “facteurs causaux”. Le calcul probabiliste sur les distributions multimodales complexes issues des capteurs serait réalisé par échantillonnage (Buesing et al., 2011), via des mécanismes d’optimisation fondés sur la parcimonie (minimiser les causes) et la vraisemblance (maximiser la justesse) (Rao & Ballard, 1999; Friston & Kiebel, 2009). Même si ces principes apparaissent séduisants, avec des connexions intéressantes vers les algorithmes d’auto-encodage développés en apprentissage automatique (Hinton et al., 2006), leur réalisation matérielle dans le substrat neuronal reste à découvrir. Il manque également, même si des tentatives sont en cours (Sutskever et al., 2009; Perrinet et al., 2014; Huang & Rao, 2014), de nombreuses “briques” pour construire un modèle général de l’apprentissage des flux sensoriels et des transformations sensori-motrices, comme nous le verrons plus loin.

## 1.2 Parcours personnel

### Les années 90

J'ai suivi un parcours classique d'élève-ingénieur via le système des classes préparatoires et des concours de Grandes Ecoles. J'ai intégré l'Ecole Nationale Supérieure d'Electronique, d'Electrotechnique, d'Informatique, d'Hydraulique de Toulouse (ENSEEIH), dans la filière informatique, en Septembre 1992.

Ma spécialisation vers les sciences cognitives s'est tout d'abord faite avec la rencontre en début de troisième année avec Bernard Doyon, médecin et chercheur à l'INSERM (Hôpital de Purpan) suite aux assises du Pôle de Recherche en Science Cognitives de Toulouse (PRESCOT), tenues à Labège, les 23 et 24 septembre 1994. Bernard Doyon y présentait l'étude de comportements chaotiques obtenus par simulation de réseaux de neurones, en faisant le parallèle avec le fonctionnement du cerveau (en particulier avec l'étude des dimensions fractales des signaux électrophysiologiques pour différents états de veille et de sommeil, développée à l'époque par l'équipe d'Agnès Babloyantz à Bruxelles (Babloyantz, 1985), ainsi que les travaux de Walter Freeman (Skarda & Freeman, 1987) sur la réponse aux stimuli olfactifs sur des grilles d'électrodes chez le lapin).

J'ai décidé de m'engager dans cette voie de recherche (simulation de réseaux de neurones chaotiques) dans le cadre d'un stage du DEA de Representation de la Connaissance et Formalisation du Raisonnement (RCFR) de l'Université Paul Sabatier, en parallèle de ma troisième année d'études d'ingénieur. Les cours étaient centrés sur la logique formelle, en particulier les grammaires génératives, les systèmes experts, les modèles linguistiques et la théorie de jeux. Les réseaux de neurones et l'apprentissage y étaient peu abordés.

Le stage était situé au sein du Département d'Etudes et de Recherche en Informatique (DERI), dans un laboratoire de l'Office Nationale de Recherche d'Etudes et de Recherche Aérospatiales (ONERA) rattaché à l'Ecole Nationale Supérieure de l'Aéronautique et de l'espace (Sup'Aéro). J'ai intégré le groupe de recherche animé par Manuel Samuelidès, professeur à Supaéro, et Bernard Doyon, chercheur à l'INSERM, sur la thématique des réseaux de neurones récurrents à connectivité aléatoire.

Cette étude avait démarré au début des années 90. Les premières études sur ce type de modèle avaient été faites par l'équipe de Sompolinsky (Sompolinsky et al., 1988) de l'université de Jérusalem, dans le cadre des systèmes dynamiques continus. L'analyse d'une variante à temps discret avait été proposée par Bruno Cessac et Manuel Samuelidès (Cessac et al., 1994). Cette étude s'est poursuivie pendant une dizaine d'années, pour s'achever par la démonstration de la propriété dite de "chaos local" dans le cas discret, par Moynot et Samuelides (Moynot & Samuelides, 2002).

Lors de mon arrivée, deux étudiants venaient de passer leur thèse sur ce thème. D'un côté, Bruno Cessac, diplômé en physique théorique de l'université Paul Sabatier, avait travaillé sur l'analyse en champ moyen des réseaux de neurones récurrents aléatoires (Cessac, 1995). De l'autre, Mathias Quoy, diplômé de l'ENSEEIH, avait travaillé sur le comportement à taille finie (route vers le chaos) (Doyon et al., 1993; Cessac et al., 1994). Mon rôle était alors principalement de poursuivre l'analyse des propriétés du réseau, et en particulier d'approfondir la question de la plasticité synaptique et de l'apprentissage.

Je connaissais alors peu de choses sur les systèmes dynamiques et le chaos, et mes

connaissances en neurosciences se limitaient à une culture grand public. Je me suis formé à ces deux domaines au cours de mon DEA, puis de ma thèse. L'ambiance de travail était très stimulante et nous avons rapidement pu produire de nouveaux résultats qui ont constitué mon rapport de fin d'études et de DEA (Daucé, 1995). J'ai obtenu mon diplôme d'ingénieur informaticien en juin 1995 et mon diplôme de DEA en septembre 1995. J'ai pu bénéficier d'une bourse ministérielle qui m'a permis de poursuivre en thèse, toujours sous la direction de Bernard Doyon, au sein de l'équipe animée par Manuel Samuelidès au DERI.

Mon travail de thèse portait sur l'étude d'un mécanisme de *réduction* de la dynamique par plasticité synaptique au sein de réseaux de neurones récurrents aléatoires. Il s'agissait d'un sujet assez vaste, qui s'est progressivement orienté vers l'apprentissage et la reconnaissance de séquences spatio-temporelles dans les réseaux de neurones récurrents organisés en plusieurs couches de traitement.

Les résultats de la thèse ont donné lieu à de nombreuses publications et présentations en conférences. La première série de résultats concerne l'étude du mécanisme d'apprentissage par "route inverse". J'ai travaillé en particulier sur les comportements à taille finie du modèle de réseau de neurones aléatoire. Le premier article qui présente ces idées est celui de Neural Networks (Daucé et al., 1998), étudiant la réaction (spontanée ou acquise) d'un réseau de neurones récurrent aléatoire à des patrons (stimuli) également aléatoires (voir section 7.2.1 page 88 et annexe A).

Un deuxième travail, portant sur des architectures multi-couches, a été réalisé en collaboration avec Olivier Moynot, étudiant diplômé de l'ENS Cachan, qui a entamé sa thèse sous la direction de Manuel Samuelidès un an après moi, et Olivier Pinaud, étudiant en DEA de physique théorique, qui a effectué son stage à l'ONERA au printemps 1998 (Daucé et al., 1999; Daucé et al., 2001) (voir section 6.1 page 74 et annexe B). Nous avons, dans le cadre de réseaux équilibrés (constitués d'une population excitatrice et d'une population inhibitrice), mis en évidence l'existence de deux paramètres de contrôle caractéristiques (l'amplification et l'homogénéité d'un faisceau d'axones) jouant sur la nature des régimes développés (extinction, activité chaotique non spécifique et oscillations synchronisées) (Daucé et al., 2001).

La troisième série de résultats de la thèse et des publications associées (Daucé, 2000; Daucé et al., 2002) porte sur une architecture neuronale permettant d'implémenter un modèle de la perception des signaux spatio-temporels. Ces résultats reposent sur une architecture de réseaux récurrents aléatoires à plusieurs couches (deux ou trois). Le formalisme multi-couches est le même que celui utilisé dans les réseaux équilibrés (voir section 7.2.2 page 91 et annexe C).

Cette thèse m'a également permis d'assister à l'essor des neurosciences computationnelles au cours des années 1995-2000, avec les premiers neurones à impulsions (Thorpe et al., 1996; Lindblad & Kinser, 1998), l'émergence du concept de connectivité fonctionnelle large échelle (Rodriguez et al., 1999), l'application des modèles de champ neuronal au contrôle moteur (Schöner et al., 1995), les modèles de formation des champs récepteurs (Olshausen & Field, 1997), etc. Dans le même temps, le domaine des réseaux de neurones artificiels se rapproche de plus en plus de l'apprentissage statistique (qui deviendra le "Machine Learning") avec l'essor des machines à vecteurs supports (SVM) à la fin des années 90 (Vapnik, 1998).

## Couplages sensori-moteurs (2000-2009)

Les années 90 ont également vu un certain renouveau dans le champ des sciences cognitives, avec, entre autres, l'apparition des concepts de cognition "située" (Brooks, 1991) et d'"énaction" (Varela et al., 1991). Ces nouvelles approches mettent en avant le rôle de l'expérimentation motrice dans l'acquisition de connaissances (autrement dit lient la connaissance – le "connaissable" – aux compétences sensori-motrices de l'agent).

Une expérience de robotique mobile inspirée de ces idées a été effectuée dans le cadre d'un court séjour de fin de thèse (juillet 1999) au sein de l'équipe de neurocybernetique dirigée par Philippe Gaussier au laboratoire ETIS (Cergy-Pontoise). Cette équipe, qui met en œuvre l'approche "située" dans une perspective développementale et cognitive, a développé plusieurs modèles d'apprentissage sensori-moteur reposant sur des cartes "cognitives" inspirées du fonctionnement de l'hippocampe (O'Keefe & Nadel, 1978). L'idée était ici d'étudier le comportement de l'architecture d'apprentissage développée dans le cadre de ma thèse dans un contexte de contrôle en boucle fermée. L'apprentissage dans ce contexte revient à stabiliser par plasticité synaptique certaines séquences d'interaction (certaines séquences d'associations sensori-motrices). L'expérience, qui met en évidence un mécanisme spontané de couplage/découplage entre l'activité récurrente interne et de la séquence motrice, a été conduite sur la plateforme de contrôle robotique PerAc (Gaussier & Zrehen, 1995), et publiée dans (Quoy & Daucé, 2000; Daucé & Quoy, 2000; Quoy et al., 2001; Daucé et al., 2002) (voir section 8.2.1 page 106).

J'ai été recruté comme maître de conférences en informatique en septembre 2001 à l'École Supérieure de Mécanique de Marseille (ESM2), depuis devenue "École Centrale de Marseille" suite à la fusion entre ESM2, l'ENSSPICAM (École Nationale Supérieure de Synthèses, Procédés et Ingénierie Chimiques d'Aix-Marseille), l'ENSP (École Nationale Supérieure de Physique) et l'ESIM (École Supérieure d'Ingénieurs de Marseille). Le profil de recherche était défini au sein du laboratoire "Mouvement et Perception", une unité mixte CNRS associée à la faculté des Sciences du Sport de Luminy (Université de la Méditerranée – Aix-Marseille II). Ce laboratoire, devenu depuis "Institut des Sciences du Mouvement" regroupe de nombreuses thématiques allant de la "biorobotique" (N. Franceschini) aux sciences sociales, en passant par les neurosciences, la psychologie expérimentale et la psychophysique. J'ai été intégré au sein de l'équipe animée alors par Reinoud Bootsma "Information et Dynamique du Mouvement". Cette équipe étudiait alors principalement le lien entre des mesures psychophysiques et les consignes motrices et sensorielles (de type "Tâche de Fitts"), selon une approche dite "écologique" de la perception et de l'action (Gibson, 1979)<sup>2</sup>.

Je suis resté au sein de ce laboratoire jusqu'en 2011. Tout en conservant ma thématique principale, mes recherches se sont orientées vers la question du rôle des dynamiques intrinsèques dans l'apprentissage moteur :

- Etude sur le comportement critique des réseaux aléatoires de taille finie, dans le cadre d'une co-direction d'ouvrage (Guillot & Daucé, 2002) (voir section 6.2.2 page 78). La perturbation d'un réseau récurrent par un signal externe à variations lentes

---

2. L'approche écologique fait reposer la commande motrice sur l'existence de motifs invariants dans la relation de l'agent à son environnement (comme le point de fuite) servant de support à la l'actionnabilité de cet environnement.

produit des transitions brusques, caractéristiques d'un comportement critique, support potentiel d'une catégorisation (discrète) de l'environnement (Beslon & Daucé, 2002).

- Extension des modèles de réseaux aléatoires équilibrés, avec couplages aléatoires dépendants de la distance permettant de développer des dynamiques méta-stables à “foyers d'activité” (Daucé, 2004b) (voir section 6.2.3 page 81 et annexe D). La combinaison de patrons de connexion à la fois topographiquement organisés et équilibrés (autorisant une activité périodique intrinsèque) permet d'implémenter une mémoire à court terme sur des neurones impulsionsnels, sans faire appel à des mécanismes cellulaires spécifiques.
- Apprentissage d'un contrôle moteur en boucle fermée, modèles Hebbiens de l'apprentissage par renforcement (Daucé, 2004a) (voir section 8.2.2 page 108 et annexe E)

### Neurones à impulsion (2004-2014)

**ACI “Temps et Cerveau” - projet “DYN” (2002-2007)** Le travail sur les neurones à impulsion a été initié dans le cadre du programme de collaboration de l'ACI en neurosciences intégratives et computationnelles “Temps et cerveau”, projet “DYN” (Dynamique des réseaux de neurones artificiels biologiquement plausibles) (Samuelides et al., 2007). Une des questions au coeur de ce projet était la possibilité d'étendre l'analyse de champ moyen des réseaux récurrents aléatoires au cas des neurones impulsionsnels, ainsi bien sûr que d'exploiter les possibilités dynamiques de ces modèles pour l'apprentissage et le contrôle. Les changements principaux, par rapport aux modèles développés durant la thèse, sont donc :

- sortie discrète (potentiels d'action),
- mise à jour par équation différentielle (et non plus équation aux différences finies),
- prise en compte explicite des délais et des constantes temporelles,
- spécialisation des neurones (excitation/inhibition).

Mes contributions dans ce domaine portent sur l'étude des changements d'échelle temporelle entre neurones fréquentiels et neurones à potentiels d'action, à la définition de patrons de connectivité séparant clairement neurones excitateurs et neurones inhibiteurs, et à la mise en place d'un formalisme architectural indépendant des modèles. Ce projet collaboratif a abouti à la publication d'un numéro spécial dans la revue “European Physical Journal – Special Topics” (Cessac et al., 2007).

J'ai également étudié dans ce cadre les conditions de l'émergence d'oscillations synchronisées *induites* par plasticité synaptique dans des réseaux équilibrés constitués d'unités binaires (autrement dit neurones impulsionsnels sans mémoire). Les principaux résultats concernant l'application de la plasticité sur ces réseaux sont détaillés dans l'article EPJ-ST de 2007 (Daucé, 2007) (voir annexe G).

**Apprentissage dans les réseaux équilibrés** L'extension à des modèles plus réalistes biologiquement a été réalisée, en co-direction avec Gilles Montagne, sur le projet de thèse de Frédéric Henry, étudiant du master de sciences cognitives à Lyon. Il a bénéficié d'une bourse ministérielle de l'école doctorale de sciences du mouvement (ED 463), entre sep-

tembre 2005 et juin 2008, puis a continué à travailler sur le sujet une année supplémentaire en tant qu'ATER à l'université de Lille (2008-2009) sans toutefois parvenir à achever son projet de thèse, malgré des publications et résultats significatifs.

L'enjeu du passage aux réseaux de neurones à potentiels d'action était la prise en compte plus fine des aspects temporels de l'activité neuronale, en considérant les délais de transmission (qui peuvent être variés) et l'ordre de tir sur des échelles de temps de l'ordre de la milliseconde.

- Une première étude, initiée dans le cadre de son stage de master sous la supervision d'Hedi Soula, et poursuivie et approfondie pendant sa première année passée à Marseille, exploite la dynamique d'un réseau récurrent aléatoire constitué de neurones impulsionnels, et montre comment l'utilisation alternée d'une règle Hebbienne et d'une règle anti-Hebbienne permet de former une réponse qui exploite la mémoire temporelle d'un substrat récurrent aléatoire (approche de type "réservoir"). Elle a fait l'objet de deux publications de conférence et d'un article (Henry & Daucé, 2006; Daucé & Henry, 2006; Henry et al., 2007) (voir section 7.2.4 page 99 et annexe F).
- Deux articles de conférences (Henry & Daucé, 2008a; Henry & Daucé, 2008b) étudient un protocole d'apprentissage où un réseau aléatoire est soumis à des motifs statiques aléatoires (comme dans (Daucé et al., 1998)), montrant un caractère régulateur de la règle de Hebb sur la dynamique du réseau (voir section 7.2.3 page 94 et annexe H).
- Un patron de connectivité plus réaliste a été proposé dans une étude datée de 2009 (Henry & Daucé, 2009), avec séparation fonctionnelle des neurones excitateurs et inhibiteurs, et utilisation de neurones à conductance (Kepler et al., 1992).

### Les projets MAPS et NOAT (2007-2010)

A travers deux projets ANR, j'ai pu aborder ces aspects de modélisation inspirés par des observations électrophysiologiques et psycho-physiques autour de l'étude des comportements d'orientation .

**MAPS** J'ai contribué début 2007 à l'écriture d'un projet ANR blanc intitulé "MAPS" (Mapping, Adaptation, Plasticity and Spatial Computation) qui a reçu un financement entre septembre 2007 et septembre 2010. Cette recherche collaborative impliquait entre autres Frédéric Alexandre, chercheur en neurosciences computationnelles au LORIA (Nancy), Hélène Paugam-Moisy, professeur à l'université de Lyon, et Laurent Goffart, chercheur en neurosciences à l'INCM (Marseille).

Ce projet portait sur la notion de codage topographique dans les modèles de réseaux de neurones biologiquement inspirés. Un des buts du projet ANR MAPS était de proposer une modélisation du comportement d'orientation visuelle, à travers l'étude du système saccadique.

J'ai participé dans ce cadre à une étude sur l'orientation visuelle à partir des données anatomiques et électrophysiologiques du tectum et des neurones prémoteurs du tronc cérébral, en collaboration avec Alain Guillaume, travaillant à l'Institut des Sciences du

Mouvement, et Anthony Mouraud, qui a effectué un séjour post-doctoral en 2010 sous notre supervision (Daucé et al., 2011) (voir section 8.2.4 page 115 et annexe L).

Deux études sur le rôle du codage topographique dans l'apprentissage d'une cinématique inverse, en collaboration avec Alain Dutech du LORIA, ont également été réalisées dans ce cadre (voir section 8.2.3 page 110 et annexes I et J).

**NOAT** J'ai participé au projet NOAT ("L'attention temporelle : modélisation neurale et substrats neuroanatomiques des processus attentionnels implicites et explicites") qui a été financé entre 2007 et 2010 dans le cadre d'un appel d'offre ANR "Neuro" datant de 2007 (ANR-07-NEURO-033-03). Ce projet collaboratif portait sur des mécanismes attentionnels liés à un contexte temporellement structuré.

Dans le cadre du projet ANR NOAT, en collaboration avec Jennifer Coull (LNC, Marseille) et Daniele Schön (INCM, Marseille), j'ai également travaillé sur des modèles génératifs du comportement d'orientation temporelle, en collaboration avec Gaurav Malhotra, chercheur post-doctorant en 2010-2011 sur cette question (Malhotra & Daucé, 2011) (voir section 8.2.4 page 115 et annexe K).

## Modèles stochastiques de l'apprentissage

**Délégation INRIA Lille Nord-Europe(2008-2009)** J'ai bénéficié entre septembre 2008 et aout 2009 d'une délégation INRIA au sein de l'équipe SequeL (Sequential Learning) dirigée par Rémi Munos au sein du centre INRIA Lille-Nord Europe. L'équipe de Rémi Munos travaille sur l'apprentissage statistique et plus particulièrement sur le dilemme exploration/exploitation, exprimé sous sa forme la plus simple comme un problème d'échantillonnage aléatoire de l'espace des choix (problème dit du "bandit manchot").

Ce séjour a été pour moi l'occasion de m'intéresser à la question de la stationnarité des processus d'apprentissage dans le cadre du contrôle moteur.

**Projet CO-ADAPT (2009-2014)** L'idée générale de cette stratégie "versatile" été le point de départ de ma participation à un travail de collaboration qui s'est concrétisé sous la forme d'un projet ANR "Défis", portant sur l'application de stratégies adaptatives dans le cadre des interfaces cerveau-ordinateur (projet CO-ADAPT - 2010-2014). Ce projet, porté par Maureen Clerc de l'INRIA Sophia-Antipolis, impliquait des expérimentateurs et des théoriciens, entre autres Rémi Munos, Boris Burle (LNC, Marseille), Jérémie Mattout (CRNL, Lyon) et Bruno Torresani (LATP, Marseille).

Le but de ce projet était de concevoir des interfaces dont la prise en main nécessite le moins d'assistance extérieure possible. L'idée du projet CO-ADAPT était d'étudier des méthodes et algorithmes adaptatifs pour faire en sorte que le logiciel utilise les réactions du sujet pour améliorer ses performances (temps de réponse, potentiels d'erreur, etc.).

- Une étude portant sur l'apprentissage par renforcement dans le cadre du "P300-speller" (section 8.2.5 page 119) a été entamée dans le cadre d'un stage de deuxième année de deux étudiants de Centrale Marseille : Timothée Proix et Arnaud Rachez (en 2010). Elle a fait l'objet d'une communication dans la conférence TOBI (TOwards Brain-computer Interfaces) en décembre 2010 à Rome (Rachez et al., 2011).

- Timothée a continué à travailler sur le sujet dans le cadre de son master recherche en 2011, en collaboration avec Liva Ralaivola du LIF (Laboratoire d’Informatique Fondamentale, Marseille), en étudiant les caractéristiques du problème dans le cadre plus général des classifieurs à noyaux.
- Ces travaux ont été publiés dans plusieurs conférences (Daucé & Ralaivola, 2012; Daucé & Proix, 2013; Daucé et al., 2013; Daucé et al., 2015) (section 8.2.5 page 119 et annexe O).

Dans une deuxième étude, j’ai travaillé dans le cadre d’une collaboration avec Maureen Clerc, Eoin Thomas, Dieter Devlaminck et Loïc Mahé, travaillant à Sophia-Antipolis, sur un modèle de gradient logistique dans la cadre de classifieurs probabilistes calculés sur un groupe de sujets, selon le principe proposé par (Kindermans et al., 2012) (section 8.2.5 page 119 et annexe N).

**Bornes d’erreur en classification multi-classes (2013-2016)** Enfin, sur un projet de thèse financé par le mécanisme CSC (China Scholarship Council), dans le cadre des collaborations entre Centrale Marseille et Centrale Pékin, et en co-direction avec Liva Ralaivola du LIF, nous travaillons jusqu’en février 2016 avec Hongliang Zhong sur l’estimation des bornes d’erreur d’algorithmes d’apprentissage en ligne de classifieurs multi-classes, lorsque l’erreur de classification n’est connue que par un signal binaire de type “vrai/faux”. Nous montrons, dans le cadre d’un algorithme fondé sur un principe d’optimisation quadratique, que les bornes d’erreur sont identiques à celles obtenues lorsque l’information est complète.

### Travaux à l’INS (Institut de Neurosciences des Systèmes) - 2011-présent

Le projet “The Virtual Brain” (TVB) (Leon et al., 2013), porté par Viktor Jirsa, vise à combiner la connaissance anatomique et connectomique pour produire des simulations de l’activité électrique à large échelle du cerveau. Le modèle proposé repose d’une part sur la diffusion isotropique de l’activité électrique sur la surface corticale (de type “Neural Field”), et d’autre part la prise en compte de la connectivité longue distance.

Dans le cadre d’un travail de co-direction de la thèse de Mathieu Golos à l’Institut de Neurosciences des Systèmes, j’ai contribué à une étude sur la multistabilité dans les systèmes neuronaux artificiels construits à partir du connectome (Golos et al., 2013; Golos et al., 2015). Nous sommes partis d’un modèle simplifié ignorant le terme de diffusion latérale. Chaque nœud du réseau est simplement identifié à une masse neuronale. Partant de l’analogie avec les réseaux de Hopfield (Hopfield, 1982), le but était d’identifier le nombre et l’organisation anatomique des attracteurs se développant au sein d’un tel réseau. L’étude de la dynamique de réseaux large échelle s’appuyant sur des patrons de connectivité réalistes met en évidence des régions de forte multistabilité, permettant à la fois d’analyser les affinités fonctionnelles entre les différentes régions anatomiques, et de simuler certains aspects de la dynamique large échelle (en particulier l’alternance entre différents patrons d’activité) (Golos et al., 2015) (voir section 6.2.5 page 84 et annexe P).

## 1.3 Organisation du document

Les études présentées dans ce rapport portent sur des algorithmes permettant de “programmer” des graphes constitués de nombreuses unités de calcul élémentaire. Nous nous plaçons dans une perspective de *modélisation biologique*. Nous cherchons donc à identifier les algorithmes “naturels” permettant au cerveau de s’adapter à son environnement. Les nœuds du réseau sont les neurones et les arêtes du graphe sont les axones transportant le signal d’un neurone à l’autre. La contrainte de plausibilité biologique impose de prendre en compte les observations et données expérimentales dans l’élaboration des modèles.

Les résultats ont été organisés autour de trois grandes thématiques :

1. La thématique des “capacités d’expression” des réseaux de neurones à couplages aléatoires, dans le cadre de leur interaction avec un signal extérieur.
2. La thématique de la plasticité synaptique et de l’apprentissage, via une *réduction de dynamique* induite par la plasticité de Hebb.
3. La thématique du contrôle moteur et de l’apprentissage moteur dans des espaces continus.

La première partie (“Vue d’ensemble”) offre un aperçu complet des problèmes et thématiques abordées au cours de mon activité de recherche. Il contient également un bilan critique, ainsi que différentes suggestions et projets pouvant servir à des développements futurs de ce champ de recherche.

- Le chapitre 2 présente certains grands principes liés à la modélisation biologiquement inspirée des réseaux de neurones. Il s’adresse principalement aux non-spécialistes du domaine.
- Le chapitre 3 présente les principes fondamentaux de l’apprentissage et du contrôle tels qu’on les trouve dans la littérature.
- Le chapitre 4 présente les principaux résultats de mon activité de recherche.
- Enfin, le chapitre 5 met en perspective ces différents résultats, et développe plusieurs pistes d’étude pouvant servir de base à des travaux futurs.

La deuxième partie (“Résultats détaillés”) offre donc un niveau de détail plus élevé sur les différents résultats et publications évoqués dans la première partie. Elle est organisée en trois chapitres :

- Le chapitre 6 aborde la question de l’expressivité des grands réseaux de neurones récurrents, c’est-à-dire de leur capacité à traiter et transformer des signaux.
- Le chapitre 7 aborde la question de la plasticité, en lien avec l’apprentissage de dépendances temporelles et de prédiction des signaux.
- Enfin, le chapitre 8 traite la question des interactions entre un contrôleur neuronal et son environnement, dans le cadre de l’apprentissage moteur.

Les annexes contiennent une liste de publications que je considère comme les plus significatives.

**Première partie**

**Vue d'ensemble**



# 2

## Préambule

Les travaux présentés dans ce mémoire sont centrés sur la modélisation de *circuits reconfigurables* dans le cadre du traitement et de la transformation des signaux sensoriels. L’approche que nous privilégions considère l’existence d’une *activité endogène* produite par le réseau de neurones *récurrent* qui traite le signal. Le but de nos travaux est de décrire, implémenter et simuler la reconfiguration des circuits neuronaux sous l’effet conjugué de la dynamique collective des neurones et de la plasticité synaptique. Le but est de proposer des modèles réalistes de l’activité neuronale et des mécanismes calculatoires qui se développent dans le cerveau.

### 2.1 Neurones et réseaux

Pour modéliser le cerveau nous utilisons l’analogie du circuit reconfigurable. Un circuit est ici un graphe (connexe) constitué de nœuds et d’arêtes. L’activité qui se développe dans le circuit repose sur la transmission de signaux sur les axones via les neurones qui en sont les relais.

Inspirés par le fonctionnement du cerveau, les réseaux de neurones sont des modèles de calculateurs dans lesquels le “programme” (la fonction de réponse) est décrit par un graphe. Ce graphe est constitué d’un ensemble de nœuds, qui sont les unités de calcul, et un ensemble d’arêtes, pondérées et orientées, qui transportent le signal entre les différentes unités de calcul.

Le signal produit par les neurones biologiques n’est pas une grandeur continue. Le système nerveux est donc caractérisé par l’émission et la réception de signaux discrets, qui présentent à première vue une certaine analogie avec les signaux numériques, à ceci près qu’il n’existe pas d’horloge centrale pour cadencer les opérations. C’est ce caractère discret des opérations neuronales qui avait inspiré les premiers modèles de neurones, au sein desquels les unités neuronales se comportaient comme des portes logiques (McCulloch & Pitts, 1943), par analogie avec les circuits logiques des architectures informatiques.

Il existe de nombreux modèles de neurones et de nombreux modèles de réseaux de neurones. La modélisation des processus neuronaux repose donc sur un choix du modélisateur, qui est fonction du mécanisme qu’il souhaite étudier, des outils d’analyse et/ou de la puissance de calcul dont il dispose.

Les propriétés et éléments de vocabulaire listés ci-dessous sont communs à la plupart des modèles de réseaux de neurones, indépendamment de leur utilisation en modélisation ou en apprentissage automatique.

### Activité et signal

Chaque unité de calcul (ou neurone) est modélisée comme une fonction de réponse  $f$  qui traite un jeu de *données d'entrée* multimodal  $s_1^{\text{in}}, \dots, s_n^{\text{in}}$ . On peut supposer sans perte de généralité que les données d'entrée sont indexées sur l'axe temporel. On parle alors de *signal d'entrée* où  $s_i^{\text{in}} = \{s_i^{\text{in}}(t)\}_{t \in \{t_0, \dots, t_f\}}$  représente un jeu de données indexé sur une trame temporelle et  $s_i^{\text{in}}(t)$  représente un point de mesure du  $i^{\text{ème}}$  signal à l'instant  $t$ .

La sortie du neurone au temps  $t_f$  est un scalaire :

$$s^{\text{out}}(t_f) = f(s_1^{\text{in}}, \dots, s_n^{\text{in}}) \quad (2.1)$$

(on parle aussi de *réponse* du neurone aux signaux d'entrée). Un neurone est donc une unité élémentaire de traitement des données.

Le signal de sortie du neurone  $s^{\text{out}} = \{s^{\text{out}}(t)\}_{t \in \{t_0, \dots, t_f\}}$  est constituée d'une succession d'états "hauts" et d'états "bas". Un neurone dont la sortie à l'instant  $t$  est dans l'état haut est dit *activé*. Un neurone dont la sortie à l'instant  $t$  est dans l'état bas est dit *inactif*.

Exemples :

- Dans le cas le plus simple des neurones binaires, le signal de sortie est une succession de 0 (inactif) ou de 1 (actif) (McCulloch & Pitts, 1943).
- Dans le cas plus complexe de neurones à impulsions, le signal de sortie est un train de potentiels d'action (PA), représenté par une somme de Diracs tels que  $s(t) = \sum_{\hat{t} \in \mathcal{T}_{\text{out}}} \delta(t - \hat{t})$ , où  $\mathcal{T}_{\text{out}}$  est un ensemble contenant les instants de décharge (voir par exemple (Gerstner & Kistler, 2002)).

Le signal de sortie (la suite d'états hauts et d'états bas) est aussi appelé l'*activité* du neurone. Ce signal est transporté sur un axone, qui se sépare à son extrémité en plusieurs branches se terminant par une (ou plusieurs) synapse(s).

Les synapses sont les canaux d'entrée des neurones. Chaque synapse traite un signal  $s_i^{\text{in}}$  émis par un autre neurone pré-synaptique  $i$  du réseau et produit une *entrée synaptique*  $e(s_i^{\text{in}})$ . L'état *interne* du neurone post-synaptique est donné par son *potentiel de membrane*  $V$ . La somme des entrées synaptiques agit sur la valeur de ce potentiel de membrane. L'intégration de la totalité des entrées peut être exprimée par une fonction de mise à jour du potentiel de membrane de la forme :

$$V = g(s_1^{\text{in}}, \dots, s_n^{\text{in}}) \quad (2.2)$$

où  $s_1^{\text{in}}, \dots, s_n^{\text{in}}$  sont les  $n$  signaux entrants.

Exemples :

- Dans le cas le plus simple,  $g$  est une combinaison linéaire des entrées synaptiques (McCulloch & Pitts, 1943; Rosenblatt, 1958).
- Les modèles plus détaillés prennent en compte la fonction de transfert des synapses ainsi que les interactions non-linéaires entre les influences excitatrices et les influences inhibitrices (Destexhe, 1997).

L'activation d'un neurone repose ensuite sur un mécanisme non linéaire de *passage de seuil*.

Exemples :

- Dans le modèle le plus simple (McCulloch & Pitts, 1943), l'état de sortie (0 ou 1) dépend d'un seuil d'activation  $\theta$  tel que  $s = 1$  (état haut) si  $V > \theta$  et  $s = 0$  (état bas) sinon.
- Dans les modèles plus détaillés (Lapicque, 1907; Hodgkin & Huxley, 1952), le passage à l'état haut déclenche un mécanisme actif de réinitialisation qui ramène le potentiel de membrane vers sa valeur de repos  $V_0 < \theta$ . Ce mécanisme de réinitialisation, qui interdit le maintien de la sortie dans l'état haut, rend l'état haut plus rare (et donc plus porteur d'information) que l'état bas. Il a également pour effet d'effacer la mémoire des données d'entrée antérieures au dernier potentiel d'action.

## Réseau de neurones

Plusieurs neurones connectés entre eux via des axones constituent un *réseau de neurones*. Le schéma de connexion, sous la forme d'un graphe orienté, caractérise l'*organisation structurelle* du réseau.

Les signaux transitant de neurone à neurone via les axones d'un réseau de neurones contiennent une succession d'états hauts et d'états bas qui présentent une ressemblance avec les signaux digitaux produits par les calculateurs numériques. Cette analogie de forme n'implique cependant pas que les principes de traitement et de transformation de ces signaux soient les mêmes :

- Dans les calculateurs numériques traditionnels,
  1. Un calcul est décrit comme une séquence d'opérations élémentaires réalisées dans un certain ordre à partir de données d'entrée. Ces opérations élémentaires peuvent être réalisées par les portes logiques d'un circuit électronique, ou par tout mécanisme automatique capable de ségréguer ses entrées sous la forme d'au moins deux états de sortie distincts.
  2. Les opérations entre chaînes de bits passent par une unité centrale qui réalise les opérations nécessaires aux algorithmes de calcul, reposant sur un *code* prédéfini qui attribue par exemple différentes puissances de 2 aux différents bits d'un octet, etc.
- Dans la cas des réseaux de neurones,
  1. La mise en œuvre d'un calcul repose sur la présence de neurones d'*entrée* soumis à des données extérieures, qui sont donc les opérandes du calcul. C'est le cas chez l'animal des cellules *sensorielles*. On note  $\mathbf{I}(t) = \{I_i(t')\}_{i \in \{1, \dots, N\}, t' \in [t_0, t[}$  ce signal extérieur. Si le neurone  $i$  est un neurone d'entrée, on a  $\exists t : I_i(t) \neq 0$ . Si le neurone  $i$  n'est pas un neurone d'entrée, on a  $\forall t : I_i(t) = 0$ .
  2. L'encodage des différentes composantes sensorielles a pour résultat un "patron d'activité"  $\mathbf{S}(t)$  constitué de l'ensemble des états de sortie des neurones depuis l'instant initial, soit  $\mathbf{S}(t) = \{s_i(t')\}_{i \in \{1, \dots, N\}, t' \in [t_0, t[}$ . Les échanges de signaux entre les différents neurones, via les liens récurrents, constituent le calcul.

**Graphe et fonction de réponse** Le réseau est décrit par une *fonction de couplage*, définie le plus souvent par un tableau à double entrée noté  $G$ . Chaque neurone se voit attribuer un indice  $i \in 1, \dots, N$  où  $N$  est le nombre de neurones dans le réseau. Le tableau  $G[i, j]$  contient alors les caractéristiques du lien du neurone  $j$  vers le neurone  $i$ . Un exemple de paramètres associés au lien  $(i, j)$  est le temps de propagation  $\tau_{ij}$  et poids synaptique  $J_{ij}$ . Lorsque le poids synaptique  $J_{ij}$  est non nul, on dit que les neurones  $i$  et  $j$  sont couplés. Si  $N$  est le nombre de neurones, le nombre de paramètres du réseau est en  $O(N^2)$ , correspondant à un principe de transmission du signal *de plusieurs à plusieurs* (un neurone envoie son signal de sortie vers plusieurs destinataires, un neurone reçoit des signaux de plusieurs émetteurs).

*Notations :*

- Le *vecteur d'activité*  $\mathbf{s}(t)$  (ou activité de population) est un vecteur de taille  $n$  contenant l'ensemble des sorties à l'instant  $t$  :

$$\mathbf{s}(t) = (s_1^{\text{out}}(t), \dots, s_n^{\text{out}}(t))$$

- Le *patron d'activité* est constituée de l'ensemble des signaux émis entre l'instant initial  $t_0$  et l'instant d'observation  $t$ , soit :

$$\mathbf{S}(t) = \{\mathbf{s}(t')\}_{t' \in [t_0, t]}$$

Les activités des différents neurones sont donc *interdépendantes* du fait des couplages. La *fonction de réponse* du réseau  $f$  est une fonction paramétrique décrite par le tableau de paramètres  $G$ , telle que l'activité  $\mathbf{s}(t)$  est solution de :

$$\mathbf{s}(t) = f(\mathbf{S}(t); G) \quad (2.3)$$

Il faut ici préciser, dans la mesure où les temps de transport sont supposés strictement positifs, que l'activité du réseau au temps  $t$  est dépendante de l'historique des activités précédent strictement l'instant  $t$ . L'activité d'un neurone post-synaptique  $i$  au temps  $t$  est dépendante des signaux pré-synaptiques produits aux instants  $t' < t$ , en tenant compte du *temps de transport* de ces signaux sur les axones. Plus précisément, si  $j$  est le neurone pré-synaptique et  $i$  le neurone post-synaptique, on a :

$$s_{ij}^{\text{in}}(t) = s_j^{\text{out}}(t - \tau_{ij}) \quad (2.4)$$

**Fonction de sortie du réseau** La *réponse* du réseau de neurones au signal d'entrée (la solution du calcul) est le patron d'activité induit par ce signal d'entrée.

En étendant la fonction de réponse aux entrées extérieures, l'activité instantanée  $\mathbf{s}(t)$  est la solution de :

$$\mathbf{s}(t) = f(\mathbf{S}(t), \mathbf{I}(t); G) \quad (2.5)$$

avec  $\mathbf{S}(t)$  l'activité développée dans le réseau jusqu'au temps  $t$ ,  $\mathbf{I}(t)$  le signal d'entrée et  $G$  le graphe. On dit également que le réseau de neurones *transforme* le signal d'entrée.

Il est possible de définir une *fonction de sortie* qui *décodes* ce patron d'activité. La sortie du réseau est alors

$$\mathbf{u} = h(\mathbf{s})$$

avec  $h$  fonction de décodage (ou de “*read-out*”).

Plus globalement, la fonction d’entrée/sortie (fonction de transfert) du réseau de neurones est :

$$\mathbf{u} = h(f(\mathbf{S}, \mathbf{I}; G)) \quad (2.6)$$

La sortie  $\mathbf{u}$  peut ainsi être interprétée comme le résultat du calcul réalisé à partir des données d’entrée  $\mathbf{I}$ . Dans la mesure où ce résultat est le produit de l’activité conjointe (et parallèle) des neurones du réseau, on se situe dans un contexte de *calcul distribué*, par opposition au calcul séquentiel centralisé réalisé par les ordinateurs traditionnels.

## 2.2 Traitement séquentiel et traitement récurrent

### 2.2.1 Traitement séquentiel

Une approche courante en modélisation est de stipuler l’existence d’une chaîne de traitement séquentielle et hiérarchique, des organes des sens vers des centres de traitement de plus en plus spécialisés, conduisant à l’identification d’un objet de l’environnement vis à vis duquel un programme moteur adéquat peut être déclenché. Selon cette approche, il y a un programme superviseur, un tableau de branchement qui associe les réponses motrices aux situations identifiées.

Dans une chaîne de traitement séquentiel, chaque couche de traitement est constituée d’un certain nombre de filtres sélectionnant certaines caractéristiques de la scène sensorielle. Puis les couches suivantes effectuent des combinaisons de ces filtres, puis les suivantes des combinaisons de combinaisons etc. On parle de “*pattern matching*”.

Le “*pattern matching*” (correspondance de forme) est l’opération consistant à mettre en correspondance un modèle  $\mathbf{w}$  (patron) avec une forme  $\mathbf{I}$  (signal), dans le cas le plus simple via un produit scalaire  $\langle \mathbf{w}, \mathbf{I} \rangle$ . On utilisera ici le terme “*appariement*” comme traduction de “*pattern matching*”.

Le principe de l’appariement appartient au monde de l’apprentissage automatique et du traitement du signal :

- Il prend historiquement ses racines dans la décomposition de Fourier, et s’est développé d’une part sous la forme de la compression de signal, comme par exemple la décomposition en des ondelettes, ou plus généralement les bases de filtres orthogonaux (ou pas), incomplètes, complètes ou “*surcomplètes*”.
- L’appariement est également le principe à la base des algorithmes d’apprentissage automatique, où des dictionnaires  $\{\mathbf{w}_1, \dots, \mathbf{w}_K\}$  se construisent par extraction d’invariants morphologiques (ou vecteurs caractéristiques) à partir des données présentées.

L’apprentissage de vecteurs caractéristiques est au cœur de l’apprentissage automatique, où les patrons caractéristiques sont déduits des données présentées (Vapnik, 1998). Différents algorithmes permettent, comme nous le verrons par la suite, de construire de telles chaînes de traitement de manière automatique à partir d’une simple base d’exemples, via

des mécanismes d’“auto-encodage”. Le calcul réalisé par les neurones est dans ce cas le degré d’appartenance (ou de correspondance) à un patron donné, sur un intervalle borné (en général entre 0 et 1). L’opération répétée sur de multiples neurones produit de nombreux degrés d’appartenance, cet ensemble de degrés d’appartenance pouvant lui-même être interprété comme une nouvelle forme pour un traitement ultérieur (sur le principe du “deep learning” (Hinton et al., 2006) par exemple), constituant un principe de calcul distribué efficace.

En modélisation neuronale, cette approche se traduit par le terme de “champ récepteur”, avec une réponse maximale de la cellule lorsque le signal correspond à la forme fixée (Hubel & Wiesel, 1962). Elle donne naissance à la reconnaissance par dictionnaire de formes, très utilisée à la fois en traitement d’images et pour les modèles de la perception visuelle (Rao & Ballard, 1999).

### 2.2.2 Traitement récurrent

Par opposition à ce traitement séquentiel, nous considérons une hypothèse alternative qui serait celle d’un branchement “non centralisé”. Cela revient à regarder des mécanismes de traitement dans lesquels l’appariement n’est pas réalisé au niveau des neurones, mais au niveau d’une population de neurones. Pour qu’une décision soit réalisée par une population, il faut que les neurones échangent de l’information pour aboutir à un “consensus”. Dans le langage des systèmes dynamiques, on parle de mécanisme de *relaxation* (voir plus bas). L’idée est donc que le cerveau, en présence d’un signal, construit sa réponse via une relaxation vers le circuit ayant la meilleure “correspondance” avec ce signal.

Le choix du mécanisme de relaxation, par opposition au produit scalaire, ne correspond pas uniquement à un niveau de description plus “fin”. Il correspond en effet à une architecture neuronale différente, dite architecture récurrente. Un réseau de neurones est dit récurrent lorsque le graphe qui le décrit contient des cycles. Cela signifie concrètement que le signal produit par les neurones peut circuler à l’infini, “en boucle”, à l’intérieur du réseau. On parle d’activité interne ou encore activité “endogène”. Le choix d’étudier des réseaux de neurones récurrents en traitement de l’information et apprentissage biologiquement inspirés se justifie par leur plus grande plausibilité biologique (Compte, 2006; Fiser et al., 2010).

Les réseaux de neurones récurrents sont capables de maintenir une activité persistante en l’absence de stimulation extérieure. Le cadre à considérer pour traiter l’appariement est celui des systèmes dynamiques “ouverts” dans lesquels l’activité endogène d’un réseau de neurones récurrent est “perturbée” par un signal extérieur.

**Systèmes dynamiques** La théorie des *systèmes dynamiques* repose sur un espace d’état  $\mathcal{X}$ , une trame temporelle  $\mathcal{T}$  et un flot  $\phi$  qui est une application de  $\mathcal{X} \times \mathcal{T}$  dans  $\mathcal{X}$  définissant pour tout couple  $(\mathbf{x}, t)$  l’évolution d’état comme :

$$\dot{\mathbf{x}} = \phi(\mathbf{x}, t) \tag{2.7}$$

La plupart des réseaux de neurones peuvent se modéliser sous cette forme, à partir du moment où l'état du réseau à l'instant  $t$  est entièrement spécifié, en tenant compte en particulier du potentiel de membrane et des différents temps de transport sur les axones.

La *trajectoire* du système sur la plage temporelle  $[t_0, t_f]$  est alors définie par l'intégration sur le flot de la condition initiale  $\mathbf{x}_0$ .

- Dans le cas d'un système dit "autonome", on a  $\phi(\mathbf{x}, t) = \phi(\mathbf{x})$  et la trajectoire du système  $\{\mathbf{x}(t)\}_{t \in [t_0, t_f], \mathbf{x}(t_0) = \mathbf{x}_0}$  est entièrement définie par les conditions initiales.
- Dans le cas d'un système dit "ouvert" (non-autonome), la dépendance temporelle est souvent modélisée sous la forme d'un signal externe  $\mathbf{I}(t)$ , soit  $\phi(\mathbf{x}, t) = \phi(\mathbf{x}, \mathbf{I}(t))$ . Ce signal peut être :
  - une donnée d'entrée dans le cas de modèles de traitement des données
  - un bruit externe dans le cas de modèles stochastiques

Dans le cas des systèmes autonomes dissipatifs, on appelle *attracteur* une région de l'espace d'état qui tend à attirer les trajectoires du système pour les temps longs. Un attracteur est formellement défini comme une région de l'espace d'états invariante par le flot, c'est à dire un ensemble de points  $\mathcal{A} \subset \mathcal{X}$  tel que si  $\mathbf{x}(t) \in \mathcal{A}$ , alors  $\mathbf{x}(t + \tau) \in \mathcal{A}$ , avec  $\tau > 0$ . Un attracteur est dit stable s'il existe un voisinage de  $\mathcal{A}$  tel que tout point de ce voisinage converge vers l'attracteur pour les temps longs. La trajectoire qui va de la condition initiale vers l'attracteur s'appelle la dynamique de *relaxation*. L'ensemble des conditions initiales convergeant vers un attracteur donné s'appelle le *bassin d'attraction*.

Dans le cas des systèmes non-autonomes, la réponse du réseau correspond à un compromis entre les contraintes internes, exprimées par le graphe, et les contraintes externes, exprimées par le signal. On pourra également parler de dynamique de *relaxation* pour caractériser la convergence vers un attracteur lorsque le signal extérieur est stationnaire, ou encore lorsque les changements du signal extérieur sont suffisamment "lents" pour laisser la dynamique converger vers une région de faible volume.

La généralisation des opérateurs d'appariement au cas des dynamiques de relaxation signifie essentiellement :

- (i) que la dynamique de relaxation doit permettre de séparer l'espace des signaux en régions distinctes. Les attracteurs atteints pour certaines classes de signaux doivent être *qualitativement différents* de ceux obtenus pour d'autres classes de signaux.
- (ii) qu'il existe des mécanismes de plasticité permettant de faire "bouger les frontières" entre les différentes régions, voire d'en créer de nouvelles et d'en faire disparaître.

## 2.3 Notre approche

Une des thèses principales développées ici est que :

*La plasticité de Hebb permet la mise en place d'une fonction d'**appariement conditionnel**, reposant sur les propriétés de l'activité persistante d'un système dynamique de grande dimension.*

où :

*Le mécanisme d'appariement est assimilé, dans le cas des réseaux de neurones récurrents, à une **réduction sélective de la complexité** de la dynamique.*

### 2.3.1 Séparation des signaux

La transformation des données d'entrée sous la forme d'un (ou plusieurs) vecteur(s) caractéristique(s) correspond à une *simplification* de ces données.

- Dans le cas des réseaux de neurones à couches, cette simplification correspond à une dimension de la couche de sortie plus faible que celle de la couche d'entrée.
- Dans le cas des réseaux récurrents, cette simplification se traduit par une *réduction du nombre de degrés de liberté* sur lesquels évolue la dynamique des neurones. L'activité du réseau accepte alors une description sur un espace de plus petite dimension.

L'équation de récurrence :

$$\dot{\mathbf{x}} = \phi(\mathbf{x}(t), I(t)) \quad (2.8)$$

permet de modéliser l'interaction entre une activité endogène (décrite par  $\mathbf{x}$ ) et un signal extérieur (décrit par  $I$ ).

**Réduction de la dynamique** L'évolution de l'état du système dynamique au cours du temps décrit une trajectoire dans un espace de dimension  $N$  (où  $N$  est le nombre de nœuds du réseau). On note  $n \leq N$  le nombre de degrés de liberté sur lesquels évolue *effectivement* la dynamique (la dimension de la variété sur laquelle évolue la trajectoire limite). La valeur de  $n$  s'interprète comme le *degré de réduction* opéré par le signal sur l'activité du réseau :

- dans le cas d'une entrée bruitée,  $n$  peut être proche de  $N$  ;
- lorsqu'au contraire l'entrée suscite l'activation d'une assemblée neuronale spécifique, la valeur de  $n$  tend à être faible (quelques unités).

De manière schématique,

- une activité à faible nombre de degrés de liberté correspond à un réseau de neurones "actif" ;
- une activité à haut nombre de degrés de liberté correspond à un réseau de neurones "inactif".

On parlera par la suite de *réponse de population* (ou réponse collective) lorsque le réseau "réagit" ainsi à une entrée particulière en réduisant son nombre de degrés de liberté.

**Appariement / rejet** Cette mesure de la dimension effective de l'activité du réseau permet donc de distinguer des signaux "stabilisateurs", tendant à corrélérer les neurones, des signaux déstabilisateurs, tendant à décorrélérer les neurones. Cette séparation entre les signaux d'entrée qui induisent une réduction de la dynamique et ceux qui ne le font pas correspond en outre à un principe de *sélection* tel que :

- la réduction signifie l'appariement avec une des formes stockées dans le réseau, et désigne les signaux connus ;
- l'absence de réduction signifie l'absence d'appariement avec les formes stockées dans le réseau, et désigne les signaux inconnus.

Ce type d'appariement dit conditionnel (c'est-à-dire avec rejet) autorise le réseau à *ne pas nécessairement choisir* parmi l'ensemble des choix possibles. Autrement dit, le réseau peut produire la réponse “je ne sais pas” lorsque le signal présenté n'est pas suffisamment proche d'une forme connue.

- Ce rejet des entrées inconnues correspond, dans la littérature statistique, au principe de “rejet de distance”, lorsqu'un phénomène observé est trop éloigné du modèle et classé comme “aberrant”.
- Une implémentation connue de ce principe dans le cadre des réseaux de neurones est le réseau ART (Carpenter & Grossberg, 1987), qui implémente un principe de “vigilance” correspondant à l'identification d'une entrée non-référencée par le réseau.
- Ce principe de sélection par la dynamique d'un réseau récurrent a été exprimé dans les travaux de Walter Freeman (Skarda & Freeman, 1987) dans le cadre d'un modèle de la perception des odeurs.

### 2.3.2 Plasticité

La synapse biologique est l'interface permettant la communication entre les neurones. Il s'agit d'une petite surface d'échange chimique (“bouton” synaptique) située à l'extrémité de l'arborisation terminale des axones. A l'arrivée d'un potentiel d'action, les synapses libèrent des neurotransmetteurs qui agissent sur les canaux ioniques des dendrites de la cellule post-synaptique (“libèrent” des ions), ce qui a pour effet de modifier le potentiel de membrane de la cellule post-synaptique.

L'efficacité d'une synapse dépend de plusieurs facteurs, comme la taille du bouton synaptique, la quantité de neurotransmetteurs disponibles, ainsi que la sensibilité de la cellule post-synaptique. Dans le cadre de ce mémoire, cet ensemble de facteurs est résumé sous la forme d'une valeur unique  $J_{ij}$  : le “poids” de la synapse, où  $j$  est l'index du neurone pré-synaptique et  $i$  l'index du neurone post-synaptique.

La *plasticité synaptique* est un mécanisme biologique qui modifie l'efficacité de la synapse au cours du temps. En reprenant les notations de l'équation (2.7), page 22 :

$$\dot{\mathbf{x}} = \phi(\mathbf{x}, \mathbf{J}, t) \quad (2.9)$$

$$\dot{\mathbf{J}} = \psi(\mathbf{x}, \mathbf{J}, t) \quad (2.10)$$

où  $\mathbf{x}$  est un vecteur représentant l'état du système dynamique et  $\mathbf{J}$  une matrice représentant le graphe de connexions.

Dans le cadre de ce mémoire, on considérera l'évolution des poids synaptiques comme “lente” par rapport à la dynamique d'activation (autrement dit, les poids synaptiques sont quasi-stationnaires sur de petits intervalles de temps). La plasticité (équation (2.10)) modifie donc les caractéristiques de la fonction de réponse des neurones (équation (2.9)) et vice-versa. Le mécanisme de plasticité introduit une interdépendance complexe entre le graphe, l'activité et le signal d'entrée s'il existe. Cette dynamique lente a un impact sur le comportement du réseau de neurones sur le long terme.

Dans le cadre proposé par Donald Hebb (Hebb, 1949), la plasticité est essentiellement un mécanisme *local* dépendant des échanges entre les cellules pré et post-synaptiques.

La règle de Hebb inscrit dans la structure du graphe les conjonctions d'activité pré et post-synaptique se produisant de façon répétée au cours du temps. Le poids  $J_{ij}$  est alors une quantité qui évolue au cours du temps sous la forme :

$$\dot{J}_{ij} = F(\mathbf{S}_i(t), \mathbf{S}_j(t), J_{ij}) \quad (2.11)$$

avec  $\mathbf{S}_j(t) = \{s_j(t)\}_{t \in [t_0, \dots, t]}$  l'activité pré-synaptique,  $\mathbf{S}_i(t) = \{s_i(t)\}_{t \in [t_0, \dots, t]}$  l'activité post-synaptique, et  $F$  la fonction de mise à jour des synapses.

Il s'agit essentiellement, selon l'idée initiale de Hebb, d'un mécanisme de sélection dans lequel *des activités corrélées tendent à se connecter, et des activités décorrélées à se déconnecter*. Ce modèle de plasticité synaptique, dite "potentiation à long terme" (Long Term Potentiation - LTP), a été confirmé à de nombreuses reprises par les observations (Bliss & Lomo, 1973; Bi & Poo, 1998).

Lorsque l'activité est elle-même induite par le signal d'entrée, le graphe reflète en partie les covariances présentes dans le signal. La règle de Hebb est une règle essentiellement additive et s'interprète comme "*plus de la même chose*". En pratique, le mécanisme facilitateur de Hebb est couplé avec un mécanisme stabilisateur (soustractif) qui évite la divergence du processus (via un principe de sélection compétitive) (Abbott & Nelson, 2000).

Les implémentations les plus connues du principe de Hebb sont :

- mémoire associative de Hopfield (Hopfield, 1982) ;
- réseau ART (Carpenter & Grossberg, 1987) ;
- cartes de Kohonen (Kohonen, 1982).

### 2.3.3 Assemblées neuronales

Le principe de plasticité par conjonction d'activité est articulé, dans le livre de Hebb, avec le concept d'*assemblée neuronale* (Hebb, 1949). Une assemblée neuronale<sup>1</sup> est un sous-ensemble de nœuds du réseau s'activant de manière coordonnée, et de façon réitérée au cours du temps. Toujours selon l'idée originale de Hebb, ce patron d'activité, qui réapparaît fréquemment, a été *sélectionné* (stabilisé) au cours du temps par la règle de plasticité.

Il est possible, par ailleurs, d'identifier l'activation d'une assemblée à la réalisation d'une tâche spécifique. On parlera dans la suite, pour désigner cette activité coordonnée, d'une activité *macroscopique* (par opposition à l'activité microscopique du neurone seul).

A un instant donné, une assemblée se trouve soit dans l'état inactif, soit dans l'état "activé". L'état du réseau est alors décrit, de manière macroscopique, par la (ou les) assemblée(s) particulière(s) qui est (sont) active(s) à cet instant.

La perspective des assemblées neuronales permet de voir les réseaux de neurones comme des *graphes reconfigurables* au sein desquels des circuits (assemblées) peuvent être activés ou désactivés au cours du temps.

---

1. La notion d'assemblée se distingue de la notion de population :

- les populations de neurones sont définies par des caractères anatomiques ou architecturaux
- les assemblées de neurones sont définies par la dynamique.

**Paramètre d'ordre** De manière plus quantitative, on trouve une idée voisine dans les travaux de Hermann Haken (Haken, 1983), qui a conceptualisé le principe de réponse par réduction de la dynamique collective à travers la notion de *paramètre d'ordre*, permettant de constituer une “variété centrale” sur laquelle le système dynamique évolue (*center manifold reduction*). Un paramètre d'ordre peut être vu comme une *unité de description* de l'activité d'un système dynamique, généralement de grande taille. Dans le cas de l'étude des transitions de phase dans les milieux désordonnés, permet de mesurer un type d'organisation spécifique. Il est “caractéristique” de cette organisation spécifique (Weisbuch, 1989).

**Champ moyen** Les modèles mathématiques permettant de déduire les propriétés “macroscopiques” (comportements moyens, corrélations d'activité, etc.) à partir de caractéristiques “microscopiques” s'appellent les modèles de “champ moyen” (Wilson & Cowan, 1972; Amari, 1972; Amit et al., 1985; Sompolinsky et al., 1988; Cessac, 1994; Amit & Brunel, 1997), etc. Les modèle de champ moyen considèrent l'activité instantanée de populations de neurones à travers un ou plusieurs paramètres d'ordre qui expriment une caractéristique d'ensemble de l'activité qui se développe dans le réseau. La variable d'ensemble la plus utilisée est la fréquence de décharge moyenne d'une population  $\nu(t)$ , qui exprime, à un instant donné, le nombre de neurones émettant un potentiel d'action divisé par la taille de la population.

## 2.4 Exemples

Le principe d'assemblée neuronale s'interprète principalement dans le cadre des réseaux de neurones récurrents. Nous avons vu que les réseaux de neurones récurrents sont capables de développer une activité propre (activité “endogène”), c'est-à-dire que les neurones peuvent demeurer dans un état “actif” en absence de stimulation extérieure. Cette activité interne repose sur la transmission “en boucle” de l'activité des neurones sur le graphe.

Dans ce cadre, une assemblée peut être vu comme une séquence d'activation interne stable au cours du temps. On parle parfois de “chaîne d'activation” – *synfire chain* (Abeles et al., 1993). Un certain nombre de modèles artificiels implémentent ce principe, comme le modèle des “groupes polychrones” postulé par (Izhikevich, 2006).

### 2.4.1 Réseau de Hopfield

Le modèle mathématique qui implémente le principe d'assemblée de la manière la plus simple est le modèle de Hopfield (Hopfield, 1982). Ce modèle est un réseau de neurones récurrent à unités binaires développant une dynamique de point fixe multistable. Malgré des limitations bien connues en terme de capacité de stockage (Amit et al., 1985), le réseau de Hopfield présente l'avantage de fonctionner avec des principes compatibles avec les idées de Hebb (caractère local de la règle de plasticité, analogie possible entre attracteur et assemblée), et donc de constituer un modèle plus “biologiquement plausible” que les réseaux de neurones traditionnels.

Un réseau de Hopfield est un graphe non-orienté et pondéré  $G$ , constitué de  $N$  nœuds et défini par une matrice de couplages  $\mathbf{J} = \{J_{i,j}\}_{i \in 1..N, j \in 1..N}$  ainsi que des seuils d'activation  $\boldsymbol{\theta} = \{\theta_i\}_{i \in 1..N}$ . La mise à jour de l'état du nœud  $i$  repose sur l'observation locale de ses voisins, selon :

$$x_i \leftarrow H\left(\sum_j J_{ij}x_j - \theta_i\right)$$

avec  $H$  fonction de Heaviside.

Le réseau de Hopfield implémente un principe de *mémoire associative*. Il apprend “par cœur” un certain nombre de scènes (ou images). Il est ensuite capable, lorsque certains constituants d'une scène apprise sont présentés, de compléter la scène avec les éléments manquants, de rectifier les valeurs erronées etc. Ainsi, si  $\mathbf{P}^{(1)}, \dots, \mathbf{P}^{(K)}$  est une base de vecteurs binaires (images) de taille  $N$ , le choix :

$$\begin{cases} \forall i, j, & J_{ij} = \sum_k (P_i^{(k)} - 0.5)(P_j^{(k)} - 0.5) \\ \forall i, & \theta_i = 0.5 \sum_j J_{ij} \end{cases}$$

permet d'organiser l'espace d'états en un nombre fini de *bassins d'attraction* associés aux différents motifs de la base. L'activité du réseau finit par converger vers un des ces patrons d'activité stationnaires, en fonction des conditions initiales. Chacun de ces attracteurs s'interprète ici comme une assemblée distincte, c'est-à-dire comme un circuit de renforcement mutuel aboutissant à la stabilisation d'un patron d'activité macroscopique.

## 2.4.2 Réseaux récurrents aléatoires

Un réseau de neurones aléatoire est un graphe orienté et pondéré  $G$  constitué de  $N$  nœuds défini par :

- la matrice des couplages  $\mathbf{J} = \{J_{i,j}\}_{i \in 1..N, j \in 1..N}$
- les seuils d'activation  $\boldsymbol{\theta} = \{\theta_i\}_{i \in 1..N}$
- les temps de propagation  $\boldsymbol{\tau} = \{\tau_{i,j}\}_{i \in 1..N, j \in 1..N}$
- etc.

tel que :

$$G \sim \mathcal{D}(\Xi) \tag{2.12}$$

où  $\mathcal{D}$  est une distribution de probabilité sur l'espace des graphes et  $\Xi$  est un ensemble de paramètres “macroscopiques” qui définissent cette distribution. La notion d'aléa ne porte donc pas sur la dynamique d'état mais sur le patron de connexions du réseau. On parle de bruit “figé” dans les connexions (“*quenched disorder*”).

L'ajout d'une fonction de réponse sur les nœuds définit un système dynamique déterministe, constitué de  $N$  neurones. La dynamique à temps discret et à temps de propagation constant est, en notation matricielle, définie par l'équation de récurrence :

$$\begin{aligned} \mathbf{x}(0) &= \mathbf{x}_0 \\ \forall t > 0 : \\ \mathbf{h}(t) &= \mathbf{J}\mathbf{x}(t-1) - \boldsymbol{\theta} + \mathbf{I}(t) \\ \mathbf{x}(t) &= f(\mathbf{h}(t); g) \end{aligned} \tag{2.13}$$

avec  $f$  fonction de réponse sigmoïde et  $g$  paramètre de contrôle (gain de la fonction de réponse du neurone).

L'indépendance statistique entre les différents liens est une caractéristique importante, qui permet d'analyser le comportement du modèle à la limite des grandes tailles, c'est-à-dire en faisant tendre la taille  $N$  vers l'infini. Avec une mise à l'échelle ("*scaling*") des poids appropriée (espérance, variance de la loi de tirage initial des poids en  $O(\frac{1}{N})$ ), on peut montrer que la dynamique (déterministe) tend à la limite thermodynamique vers un processus aléatoire (processus Gaussien), la corrélation entre les nœuds tendant vers 0 (Sompolinsky et al., 1988; Cessac et al., 1994; Moynot & Samuelides, 2002).

L'existence d'un paramètre de contrôle (le gain de la fonction de transfert des neurones) permet de reproduire, dans ces réseaux aléatoires, un comportement générique de déstabilisation, appelé route vers le chaos par "quasi-périodicité" (Berger et al., 1988; Doyon et al., 1993) (voir figure 6.2 page 78).

Le comportement d'une classe de réseaux (l'ensemble des réseaux définis par la même distribution) peut être décrit à la limite thermodynamique par des équations dites de "champ moyen". Ces équations définissent le processus aléatoire limite, décrit par la dynamique temporelle de grandeurs caractéristiques comme l'activité moyenne, le moments d'ordre 2 et la covariance dans le temps (Cessac, 1995).

Lorsque les poids synaptiques sont tirés indépendamment selon une loi de moyenne  $\frac{\mu_J}{N}$  et d'écart-type  $\frac{\sigma_J}{\sqrt{N}}$ , et les seuils selon une loi de moyenne  $\mu_\theta$  et d'écart-type  $\sigma_\theta$ , le processus Gaussien limite est défini au temps  $t$  par :

- sa moyenne  $\mu_h(t) = \langle h_i(t) \rangle$
- sa variance  $\sigma_h^2(t) = \langle (h_i(t) - \mu_h(t))^2 \rangle$

où  $h_i(t)$  est une variable aléatoire, selon :

$$\mu_h(t) = \mu_J m(t-1) + \mu_\theta \quad (2.14)$$

$$\sigma_h^2(t) = \sigma_J^2 q(t-1) + \sigma_\theta^2 \quad (2.15)$$

avec :

$$m(t) = \int_{-\infty}^{+\infty} Dh f(h\sigma_h(t) + \mu(t)) \quad (2.16)$$

$$q(t) = \int_{-\infty}^{+\infty} Dh f^2(h\sigma_h(t) + \mu(t)) \quad (2.17)$$

avec  $Dh = \frac{1}{\sqrt{2\pi}} \exp -\frac{h^2}{2} dh$ .

Les équations de champ moyen, qui forment un système à petit nombre de degrés de liberté, offrent ainsi une description "macroscopique" de l'activité d'un réseau, représentative de l'activité moyenne qu'on peut attendre dans un réseau de taille finie issu d'un tirage particulier. Elles constituent ainsi un cadre d'analyse adapté à une approche "multi-échelles".

### 2.4.3 Champ neuronal

La vocation des modèles de champ neuronal (Amari, 1977) est principalement la modélisation d'un substrat neuronal spatialisé vu à "large échelle" (continuum d'activité sur

la surface corticale). L'activité de larges populations de neurones s'apparente en effet, à la limite des grandes tailles, à un milieu continu dans lequel chaque portion de l'espace peut être décrit par la "densité" d'activité qui s'y développe.

Le modèle le plus connu est décrit par des équation intégral-différentielles. Soit  $x$  un point d'un espace (ou d'une portion d'espace) cartésien  $\mathcal{X}$ . Dans le modèle d'Amari (Amari, 1977), les neurones sont décrits par leur potentiel d'activation  $V$  et par la sortie binaire de la fonction d'activation  $f$  (fonction de Heaviside). Le potentiel d'activation d'un neurone est donné par l'équation :

$$\tau \frac{dV(x, t)}{dt} = -V(x, t) + I(x, t) - \theta + \int_{y \in \mathcal{X}} W(\|x - y\|) f(V(y, t)) dy \quad (2.18)$$

où  $I$  est un signal extérieur et  $W$  un noyau décrivant les interactions réciproques entre neurones en fonction de la distance  $\|x - y\|$ .

Selon le noyau choisi, le réseau développe un ou plusieurs foyers d'activité métastables interprétables au choix comme un patron d'activité ou comme un ensemble de grandeurs scalaires (en fonction des barycentres des foyers d'activité). Ce modèle partage avec le réseau de Hopfield (Hopfield, 1982) la capacité à conserver et entretenir une activité autonome indépendamment de son signal d'entrée. Il implémente néanmoins des opérations tout à fait spécifiques et intéressantes du point de vue computationnel. Un champ neuronal implémente simultanément une "mémoire de travail" (capacité à conserver la trace d'une stimulation passée), et un mécanisme de traitement (ici résumé à quelques opérations simples : émergence, extinction, fusion, poursuite) (Schöner et al., 1995). La combinaison de ces deux compétences (mémoire et calcul) au sein d'une architecture unique offre en principe un cadre computationnel complet, c'est à dire un dispositif de calcul ayant la même puissance d'expression qu'une machine de Turing (Siegelmann, 1999; Potthast et al., 2013). Le caractère continu de l'activité développée dans ces modèles suggérerait même une puissance d'expression supérieure (MacLennan, 1999), cette question n'étant à l'heure actuelle ni tranchée ni véritablement formalisée.

# 3

## Apprentissage et contrôle

L'étude des réseaux de neurones naturels suggère un mode de fonctionnement intermédiaire entre le "pattern matching" (reconnaissance de formes) et les systèmes dynamiques. Le mariage des deux mondes n'est pourtant pas évident d'un point de vue computationnel. C'est l'évidence d'un comportement massivement distribué et non strictement digital, tel qu'observé dans le cerveau, qui a poussé au développement de tels modèles.

- La démarche du pattern matching est efficace pour expliquer les mécanismes de plasticité et d'apprentissage, mais se révèle peu propice à la construction de schémas de contrôle moteur.
- La démarche des systèmes dynamiques explique bien les mécanismes de formation et de contrôle des mouvements, mais offre un cadre moins propice à la plasticité. Malgré de nombreuses tentatives, il manque encore, semble-t-il, un principe d'unification des deux approches.

Ce chapitre présente les différentes approches d'un point de vue mathématique, et indique quelques pistes de synthèse telles que proposées dans la littérature.

### 3.1 Apprentissage et plasticité

Selon une perspective biologique et développementale, l'apprentissage est le processus de changement comportemental, en relation avec ce mécanisme de plasticité. L'apprentissage est au sens large l'ensemble des processus épigénétiques (historiques) inscrivant dans l'animal ou l'individu les éléments d'expérience qui contribuent à accroître l'adaptation de ses réponses à son milieu.

Cette capacité à inscrire des événements ou des faits particuliers en vue de les exploiter dans le futur est une des propriétés essentielles du système nerveux des être vivants. Mieux comprendre les mécanismes biologiques qui sous-tendent cet apprentissage est un des défis majeurs pour les neurosciences computationnelles. Si la plasticité synaptique semble être le principe explicatif majeur de l'apprentissage, il reste encore de nombreuses zones d'ombre

- sur les caractéristiques précises de cette plasticité ;
- sur les mécanismes de choix qui vont sélectionner certains signaux et certains événements plus significatifs ;
- sur les déterminants macroscopiques des changements microscopiques et vice-versa.

### 3.1.1 Apprentissage automatique

Du point de vue computationnel, l'apprentissage automatique (le “Machine learning”) consiste à confier une tâche de programmation à un algorithme, dit algorithme d'apprentissage. En d'autres termes, il faut écrire un programme *capable de programmer*. Ce *méta-programme* s'appuie sur des méta-données (ou “contraintes”) qui sont des données servant à produire le programme. Cette conception de la programmation reposant sur l'optimisation d'un tableau de paramètres via un ensemble de contraintes remonte historiquement aux principes de la programmation dynamique (Bellman, 1956).

Le mécanisme fondamental sur lequel s'appuient la plupart des algorithmes d'apprentissage automatique est celui de l'*appariement* (“*pattern matching*”) entre un *dictionnaire de formes* et un vecteur (ou un ensemble de vecteurs) d'entrée (voir aussi page 21). Cet appariement se réalise :

- à l'aide d'un produit scalaire ;
- par un calcul de distance ;
- par un calcul probabiliste (vraisemblance, probabilité *a posteriori*,...);
- par une dynamique de relaxation (réseau de Hopfield) ;
- etc.

La plupart des algorithmes d'apprentissage construisent des dictionnaires de formes qui servent de clé d'interprétation des données d'entrée. Le choix de ces formes caractéristiques est bien sûr dépendant des capacités d'expression de la fonction de réponse :

- nombre de formes (taille du dictionnaire) ;
- types de dépendances pouvant être capturées (dépendances entre classes, dépendances spatiales, hiérarchies, dépendances temporelles, etc.)

**Remarque :** contrairement aux approches de traitement du signal classique (décomposition du signal sur des bases propices à l'analyse), les dictionnaires de formes sont ici entièrement *construits* à partir des contraintes, c'est-à-dire principalement le jeu de données fourni.

**Types de problèmes** La fonction de réponse  $h$  d'un réseau de neurones (voir équation (2.5), page 20) est en général une projection d'un espace d'entrée  $\mathcal{I}$  vers un espace de sortie  $\mathcal{U}$  de dimension plus petite que l'entrée. La fonction opère donc une *réduction* de son espace d'entrée, c'est-à-dire qu'elle réduit les données d'entrée à un petit nombre de caractéristiques. Il s'agit donc d'une transformation non-inversible du signal d'entrée, avec perte d'information.

La nature de cet espace de sortie (continu ou discret) définit deux grands types de problèmes d'apprentissage.

- Lorsque l'espace de sortie est discret, la réponse est une *catégorie*. On parle de catégorisation ou de classification des données d'entrée. Dans ce cas, la fonction de réponse repose sur un mécanisme d'*appariement sélectif* (ou “choix”), c'est-à-dire qu'en fin de compte, une forme unique est identifiée comme correspondant aux données d'entrée. Ce mécanisme est mieux connu sous le nom de “winner takes all” (le gagnant prend tout), implémenté par la fonction *argmax* :

$$u = \arg \max_{k \in 1..K} \langle \mathbf{w}_k, \mathbf{I} \rangle$$

- Lorsque l'espace de sortie est continu, on parle de problème de régression. La réponse résulte d'une *décomposition* des données d'entrée sur une base de plus petite dimension. L'appariement, dans ce cas, est dit non exclusif.

$$\mathbf{u} = h(\langle \mathbf{w}_1, \mathbf{I} \rangle, \dots, \langle \mathbf{w}_K, \mathbf{I} \rangle)$$

**Cadre général** Le but d'un algorithme d'apprentissage est de définir le jeu de paramètres  $\Theta$  de manière à produire une sortie conforme au calcul souhaité. Le jeu de paramètres est assimilable à un *programme*, par exemple  $\Theta = \{\mathbf{w}_1, \dots, \mathbf{w}_K\}$ , ou encore  $\Theta = G$  (graphe – voir aussi page 20), et le choix de ces paramètres correspond à la programmation du réseau.

L'apprentissage automatique s'exprime donc principalement sous la forme d'un problème d'optimisation dans des espaces vectoriels de grande dimension. Il s'agit d'optimiser une fonction de réponse  $h_\Theta$ , définie selon un tableau de paramètres  $\Theta$ , selon un critère d'optimalité défini par une fonction de coût  $\mathcal{H}$  et un jeu de contraintes  $\mathbf{D}$  :

$$\min_{\Theta} \mathcal{H}(\Theta, \mathbf{D}) \quad (3.1)$$

où  $\Theta$  est le “programme” et  $\mathbf{D}$  les “méta-données”.

Dans le cadre de l'apprentissage automatique, les contraintes  $\mathbf{D}$  sont les données d'apprentissage, et prennent la forme d'une base d'exemples organisée sous forme de couples  $\{(\mathbf{I}_1, u_1), \dots, (\mathbf{I}_n, u_n)\}$ . L'optimisation consiste à trouver une fonction de l'espace des entrées  $\mathcal{I}$  vers l'espace des sorties  $\mathcal{U}$  qui minimise la distance aux données.

L'optimisation par énumération est généralement impossible pour les problèmes de grande dimension. On a alors recours à des algorithmes dits d’"optimisation stochastique", ou encore des algorithmes utilisant des hypothèses simplificatrices (comme la séparabilité linéaire des données ou encore le caractère Gaussien de distributions génératrices). Le but des algorithmes d'apprentissage est alors d'extraire des données fournies des régularités non apparentes, des facteurs explicatifs, permettant de mieux prédire et classifier les données nouvelles.

### Exemples

- Les approches discriminatives (Vapnik, 1995) extraient des représentants caractéristiques des données (vecteurs supports - vecteurs les plus discriminants).
- Les approches génératives (Bishop, 2006) identifient la source cachée des observations (les facteurs causaux) selon des modèles probabilistes pré-définis. Elles définissent des vecteurs *caractéristiques* d'une classe.
- Les réseaux de neurones (voir section suivante), dont les caractéristiques propres (nombre de neurones, nombre de couches cachées, architecture et patron de connexions), déterminent des méthodes d'optimisation spécifiques.

**Modèles probabilistes** Un grand nombre de problèmes d'apprentissage peuvent s'exprimer de manière probabiliste. Dans ce cas, le problème d'apprentissage revient à déduire des données une distribution de probabilité sous-jacente (dite distribution génératrice). Les algorithmes d'apprentissage consistent alors à estimer cette distribution de probabilité en maximisant la vraisemblance des données, étant données un certain nombre de

contraintes ou hypothèses a priori de régularité (nombre de sources, type de distribution, etc.) (Bishop, 2006)

Dans ce cadre :

- Soient  $\mathcal{I}$  l'espace des observations,  $I$  une v.a. sur  $\mathcal{I}$ ,  $\mathcal{U}$  l'espace des sources,  $U$  une v.a. sur  $\mathcal{U}$  et  $f(I, U; \Theta)$  la distribution génératrice.
- L'optimisation porte sur les paramètres  $\Theta$  de la distribution génératrice (maximum de vraisemblance).

$$\max_{\Theta} \sum_{(I,u) \in \mathcal{D}} f(I, u; \Theta)$$

- La réponse est une distribution marginale  $f(U = u | I = \mathbf{I}; \Theta)$ , déduite de la distribution génératrice et des observations à partir de la formule de Bayes.

Les modèles probabilistes se distinguent des modèles d'appariement standard par le caractère probabiliste de la réponse. Autrement dit, le programme fournit une liste des probabilités de réponses et non une réponse unique. Le caractère combinable et facilement interprétable des probabilités donne à ces modèles une puissance supérieure aux classifieurs binaires standards.

## Réseaux de neurones en apprentissage automatique

Les réseaux de neurones et leurs méthodes de programmation sont un des piliers de la recherche en apprentissage automatique. Les premiers modèles remontent historiquement aux années 60-70 (Hubel & Wiesel, 1962; Grossberg, 1970). L'idée générale est celle d'un système dynamique de grande dimension, constitué d'un grand nombre de variables d'état, chaque variable d'état reflétant une composante de la scène sensorielle, les différentes composantes interagissant par des relations de voisinage excitatrices ou inhibitrices. De ces modèles sont issus les principales architectures neuronales, avec le principe d'un signal transitant par différentes "couches", et les notions de liens feed-forward, feedback et latéraux. Ces modèles ont connu un développement important dans les années 80, aboutissant aux modèles classiques (Hopfield, 1982; Kohonen, 1982; Carpenter & Grossberg, 1987).

Ils sont, à l'heure actuelle, principalement utilisés pour des tâches de catégorisation de leurs données d'entrée. Les méta-données sont alors constituées d'un jeu de données d'entrée ainsi que des sorties attendues, appelé *base d'apprentissage*. La plupart des algorithmes d'apprentissage s'affranchissent de la contrainte de localité définie par Hebb et implémentent des mécanismes de mise à jour reposant sur la propagation d'erreur (et non sur la conjonction d'activité).

**Modèles à couches** Les modèles de réseaux de neurones, tels qu'il se sont développés depuis le Perceptron (Rosenblatt, 1958), reposent majoritairement sur un principe de traitement séquentiel ("*feed forward*").

L'organisation d'un réseau à couches correspond à la transformation d'un signal d'entrée par étapes successives pour produire un signal de sortie. Les couches dites "cachées" correspondent à des étapes de calcul intermédiaires. La programmation des modèles à couches repose sur un principe de *propagation du signal d'erreur* dans le sens inverse de la propagation des données, de la sortie vers l'entrée (Rumelhart et al., 1986).

- Le modèle de Perceptron le plus simple, possédant une cellule de sortie unique, sépare son espace d'entrée en deux régions.
- Dans le cas du Perceptron multi-couches (Rumelhart et al., 1986), chaque neurone de la couche cachée est un séparateur permettant de distinguer deux régions de son espace d'entrée. Le concepteur fixe à l'avance le nombre de neurones à mettre dans la couche cachée, et donc la complexité des opérations de ségrégation produites par le réseau.
- Enfin, le modèle des séparateurs à vaste marge (SVM) (Vapnik, 1995) repose sur le modèle du Perceptron. Les données d'entrée sont plongées dans un espace de redescription de dimension élevée. La fonction de réponse est déterminée selon un principe d'optimisation quadratique. Une définition quantitative de la complexité d'un classifieur est proposée dans ce cadre : la dimension de Vapnik-Chervonenkis, correspondant dans ce cas précis au nombre minimal de vecteurs supports nécessaires pour correctement séparer les données d'entrée.

### 3.1.2 Lien entre le computationnel et le biologique

De nombreux points communs peuvent être établis entre l'apprentissage computationnel et l'apprentissage biologique. L'apprentissage automatique identifie en effet des *communautés de problèmes* :

- définissant des contraintes sur le type de solutions attendues ;
- nécessitant souvent des méthodes de résolution spécifiques.

**Remarque :** la communauté de problèmes n'implique pas nécessairement la *communauté de mise en œuvre*. Comme nous avons vu dans l'introduction :

- identifier correctement les problèmes et les contraintes des systèmes biologiques sont des enjeux à part entière ;
- il existe des contraintes spécifiques aux systèmes biologiques, telles que la contrainte de localité (plasticité de Hebb) ;
- le critère d'efficacité qui guide le développement des méthodes d'apprentissage automatique ne préjuge pas de leur plausibilité biologique.

Dans le cadre des réseaux de neurones, vus précédemment, la fonction de réponse est implémentée sous la forme d'un graphe, les nœuds sont les “neurones”, et les arêtes sont les “synapses”. La correspondance formelle entre l'apprentissage (au sens informatique) et la plasticité (voir équation (2.9), page 25) consiste à identifier :

- les paramètres de la fonction de réponse avec les poids synaptiques,
- et la plasticité avec un mécanisme de mise à jour des paramètres guidé par les données.

Nous aurons l'occasion d'étudier par la suite des modèles plus directement inspirés par la biologie. Avant cela, nous abordons l'autre grande approche utilisée en modélisation neuronale, à savoir les modèles issus de la théorie des systèmes dynamiques et du contrôle.

## 3.2 Systèmes dynamiques et contrôle

Les systèmes de contrôle en boucle fermée appartiennent à la tradition cybernétique, initialement formée autour de la commande à feedback négatif (Wiener, 1965). Les systèmes cybernétiques sont essentiellement des boucles de régulation, qui maintiennent un certain nombre de variables de contrôle dans un domaine de viabilité via un mécanisme de correction d'erreur. Dans ce schéma,

- la perception est principalement l'écart à la valeur nominale
- et la commande est un déplacement moteur qui corrige l'écart observé.

Les systèmes cybernétiques offrent un exemple élémentaire d'autonomie logicielle, avec une "prise de décision" reposant sur des échanges de signaux entre le contrôleur et son environnement.

### 3.2.1 Couplages sensori-moteurs

Le lien entre les actions et les entrées sensorielles est décrit, au niveau le plus général, par le modèle physique au sein duquel le logiciel est plongé.

Nous reprenons ici le formalisme des systèmes dynamiques proposé dans l'équation (2.7), page 22. On suppose que le milieu physique et le programme obéissent à une description d'état, où :

- $\mathbf{x}_{\text{out}}$  est un vecteur d'état décrivant le milieu physique (actuateurs compris) ;
- $\mathbf{x}_{\text{in}}$  est un vecteur d'état décrivant le contrôleur – ou encore "milieu" logiciel (entrées sensorielles comprises).

L'évolution du système au cours du temps est alors décrite par le jeu d'équations :

$$\dot{\mathbf{x}}_{\text{out}} = \phi_{\text{out}}(\mathbf{x}_{\text{out}}, h_{\text{out}}(\mathbf{x}_{\text{in}})) \quad (3.2)$$

$$\dot{\mathbf{x}}_{\text{in}} = \phi_{\text{in}}(\mathbf{x}_{\text{in}}, h_{\text{in}}(\mathbf{x}_{\text{out}})) \quad (3.3)$$

où  $h_{\text{in}}$  et  $h_{\text{out}}$  représentent (de manière simplifiée) les mécanismes de transduction entre le milieu physique et le circuit logiciel et  $\phi_{\text{in}}$  et  $\phi_{\text{out}}$  les fonctions d'évolution d'état.

Plus précisément :

- On note  $\mathbf{I}(t) = h_{\text{in}}(\mathbf{x}_{\text{out}}(t))$  le signal d'entrée du contrôleur.
- On note  $\mathbf{u}(t) = h_{\text{out}}(\mathbf{x}_{\text{in}}(t))$  la commande motrice.
- L'équation (3.2) décrit le lien de causalité entre les actions produites et les (futurs) sensations.
- L'équation (3.3) décrit les transformations appliquées à l'entrée sensorielle.

La théorie du contrôle classique décrit un problème de contrôle comme un problème d'optimisation reposant sur :

- une consigne, correspondant à un état nominal  $\mathbf{x}^*$ , dans lequel on souhaite voir le milieu physique,
- un contrôleur doté de capteurs et d'actuateurs,
- et un milieu physique situé dans l'état  $\mathbf{x}_{\text{out}}$ .

Le problème consiste à programmer une fonction de réponse  $\phi_{\text{in}}$  qui, étant donnée une consigne  $\mathbf{x}^*$ , conduit le milieu physique dans cet état désiré, soit :

$$\max_{\phi_{\text{in}}} \|\mathbf{x}_{\text{out}} - \mathbf{x}^*\|^2 \quad (3.4)$$

**Remarque :** La fonction solution  $\phi_{\text{in}}(\mathbf{x}^*, \cdot)$ , lorsqu'elle existe, s'apparente à un *programme moteur* qui conditionne la manière dont le contrôleur traite ses entrées sensorielles : pour deux consignes différentes, les mêmes entrées sensorielles ne produisent pas la même réponse.

Le point d'opération (le point où l'on retourne) d'un tel système est analogue formellement aux attracteurs des systèmes dynamiques contraints (voir page 23). Un changement de consigne induit un déplacement vers un nouvel attracteur (une nouvelle configuration mécanique ou spatiale) où l'appareil se stabilise. Concevoir un bon contrôleur nécessite donc de connaître l'ensemble du dispositif, autrement dit le circuit logiciel, les mécanismes de transduction et l'effet des commandes motrices sur le signal sensoriel (le “modèle du monde” ou “modèle forward”) (Wolpert & Ghahramani, 2000). Cet ensemble constitue un système dynamique, soumis à des influences extérieures (les facteurs non modélisés) face auxquels le système produit des réponses. Ces modèles reposent implicitement sur le traitement de grandeurs extensives (positions, vitesses) mesurées dans l'environnement.

**Remarque :** L'état de l'environnement n'est connu du contrôleur qu'à travers une projection  $h_{\text{in}}$  qui n'est pas nécessairement inversible. On dit dans ce cas que l'environnement (ou une partie de l'environnement) est “caché”.

### Estimation d'état

Le cas le plus courant en robotique et en modélisation biologiquement inspirée correspond à une situation où l'état de l'environnement n'est pas directement observable. Deux cas se présentent :

1.  $\mathbf{I} = h_{\text{in}}(\mathbf{x}_{\text{out}})$  est le signal issu des appareils de mesure. Le contrôleur doit, autant que possible, deviner la valeur de l'état, soit  $\tilde{\mathbf{x}} \simeq h_{\text{in}}^{-1}(\mathbf{I})$  pour produire la commande. On parle là encore de *modèle inverse* de la mesure.
2. Une autre approche consiste à comparer le signal observé au signal attendu  $\mathbf{I}^* = h_{\text{in}}(\mathbf{x}^*)$ . Ce signal attendu n'est pas directement observable. Il peut être *reconstruit* via un *modèle direct*.
  - On dit que le modèle effectue une prédiction sensorielle et la différence entre l'entrée prédite et l'entrée attendue est l'*erreur de prédiction*.
  - La sortie motrice doit être telle que cette erreur de prédiction sensorielle soit minimisée.

### Facteurs exogènes

Le problème de l'observation de l'état est d'une importance particulière en contrôle puisque la justesse de la commande en dépend.

L'estimation d'état dépend principalement de deux facteurs : le modèle physico-mécanique de l'environnement, et le modèle des capteurs. Ces modèles sont souvent incomplets et les facteurs non pris en compte (frottements, inertie, fatigue des composants, non-stationnarités, variations d'éclairage, bruit de mesure, fiabilité des capteurs, ...) sont mo-

délinés comme des signaux extérieurs venant moduler de façon imprévisible le comportement du système.

Ces facteurs exogènes sont modélisés comme des signaux venant influencer le système, mais n'étant pas influencés par lui. L'équation de mise à jour devient :

$$\dot{\mathbf{x}}_{\text{out}} = \phi_{\text{out}}(\mathbf{x}_{\text{out}}, \mathbf{u}, \boldsymbol{\xi}_{\text{out}}) \quad (3.5)$$

$$\mathbf{I} = h_{\text{in}}(\mathbf{x}_{\text{out}}, \boldsymbol{\chi}_{\text{out}}) \quad (3.6)$$

$$\dot{\mathbf{x}}_{\text{in}} = \phi_{\text{in}}(\mathbf{x}_{\text{in}}, \mathbf{I}, \boldsymbol{\xi}_{\text{in}}) \quad (3.7)$$

$$\mathbf{u} = h_{\text{out}}(\mathbf{x}_{\text{in}}, \boldsymbol{\chi}_{\text{in}}) \quad (3.8)$$

où les entrées  $\boldsymbol{\xi}_{\text{out}}(t)$ ,  $\boldsymbol{\chi}_{\text{out}}(t)$ ,  $\boldsymbol{\xi}_{\text{in}}(t)$ ,  $\boldsymbol{\chi}_{\text{in}}(t)$  représentent les écarts entre le modèle et la "réalité".

### 3.2.2 Approche probabiliste

Pour la simplicité de l'exposé, nous passons ici en notation discrète pour les indices temporels. La généralisation aux modèles à temps continu peut être trouvée dans (Kushner, 1964).

Les écarts entre le modèle et les observations sont fréquemment modélisés comme des bruits, issus de tirages aléatoires indépendants de la dynamique d'état, où :

—  $\boldsymbol{\xi}_{\text{out}}(t)$  est le bruit d'état et  $P$  le modèle de transition d'état externe, selon :

$$\mathbf{x}_{\text{out}}(t+1) \sim P(\mathbf{X}|\mathbf{x}_{\text{out}}(t), \mathbf{u}(t)) \quad (3.9)$$

—  $\boldsymbol{\chi}_{\text{out}}(t)$  est le bruit de mesure et  $Q$  le modèle de mesure, selon :

$$\mathbf{I}(t) \sim Q(\mathbf{I}|\mathbf{x}_{\text{out}}(t)) \quad (3.10)$$

Les équations (3.5)-(3.8) décrivent alors la réalisation d'un processus stochastique reposant sur, d'une part, des modèles d'état (les variables "porteuses" de signal – ou d'information) et d'autre part des modèles de bruit (variables "porteuses" de bruit).

Étant donné l'état connu  $\mathbf{x}_{\text{out}}(t)$ , l'estimation d'état suivant peut être produite à l'aide de la formule de Bayes :

$$\mathbb{P}(\mathbf{x}|\mathbf{I}(t+1)) = \frac{P(\mathbf{x}|\mathbf{x}_{\text{out}}(t), \mathbf{u}(t))Q(\mathbf{I}(t+1)|\mathbf{x})}{\int_{\mathcal{X}} P(\mathbf{y}|\mathbf{x}_{\text{out}}(t), \mathbf{u}(t))Q(\mathbf{I}(t+1)|\mathbf{y})d\mathbf{y}} \quad (3.11)$$

où :

- $P(\mathbf{x}|\mathbf{x}_{\text{out}}(t), \mathbf{u}(t))$  est la probabilité *a priori*.
- $Q(\mathbf{I}(t+1)|\mathbf{x})$  est la vraisemblance de l'observation  $\mathbf{I}$  étant donné  $\mathbf{x}$ .

### Filtrage

Le problème de filtrage consiste principalement à séparer le signal du bruit, à l'aide de la connaissance (ou *a priori*) représentée par les modèles d'état et les modèles de bruit.

Soit  $\mathbf{I}(t)$  le signal disponible sur les capteurs. Il s'agit d'estimer le mieux possible l'état extérieur – ou état "caché"  $\mathbf{x}_{\text{out}}(t)$ . Autrement dit il s'agit de soustraire du signal les bruits ( $\boldsymbol{\xi}_{\text{out}}(t), \boldsymbol{\chi}_{\text{out}}(t)$ ), connaissant leurs caractéristiques statistiques.

Le filtrage utilise un *modèle interne*, qui est un processus miroir de l'environnement, dans lequel la variable d'état *interne*  $\mathbf{x}_{\text{in}}$  évolue en “imitant” l'environnement. On note dans ce cas  $\mathbf{x}_{\text{in}} = \hat{\mathbf{x}}_{\text{out}}$  cette estimation, sur laquelle le contrôleur s'appuie pour choisir sa commande.

Le problème est principalement celui de compenser la dérive de l'estimation d'état dans le cas de processus aléatoires avec bruit additif (cas où l'écart à la moyenne augmente au cours du temps). Il s'agit donc de produire la meilleure estimation à *chaque instant* de l'état extérieur  $\hat{\mathbf{x}}_{\text{out}}$ .

**Dynamique interne** Dans le cadre des modèles de filtrage, l'état interne tel que défini dans l'équation (3.3), page 36, est le résultat d'un processus *récurrent* reposant :

- sur des facteurs internes représentés par la probabilité a priori  $P(\mathbf{x}_{t+1}|\mathbf{x}_t, \mathbf{u}_t)$  ;
- sur des facteurs externes représentés par la vraisemblance  $Q(\mathbf{I}_t|\mathbf{x}_t)$  ;

où les indices représentent le temps en notation discrète.

L'évolution de l'état repose sur la formule de Bayes (3.11), soit le maximum a posteriori en tenant compte des estimations précédentes. En prenant en compte les deux critères, on peut définir une dynamique sur la distribution comme :

$$f(\mathbf{x}_0) = f_0(\mathbf{x}_0)$$

$$g(\mathbf{x}_{t+1}|\mathbf{x}_{0:t}, \mathbf{I}_{1:t}, \mathbf{u}_{1:t}) \propto P(\mathbf{x}_{t+1}|\mathbf{x}_t, \mathbf{u}_t)f(\mathbf{x}_t|\mathbf{x}_{0:t-1}, \mathbf{I}_{1:t}, \mathbf{u}_{1:t-1}) \quad (3.12)$$

$$f(\mathbf{x}_{t+1}|\mathbf{x}_{0:t}, \mathbf{I}_{1:t+1}, \mathbf{u}_{1:t}) \propto Q(\mathbf{I}_{t+1}|\mathbf{x}_{t+1})g(\mathbf{x}_{t+1}|\mathbf{x}_{0:t}, \mathbf{I}_{1:t}, \mathbf{u}_{1:t}) \quad (3.13)$$

où les termes de normalisation ont été omis, avec  $\mathbf{x}_{0:t}$  représentant une trajectoire particulière du processus stochastique et les densités de probabilité  $f$  et  $g$  portent sur l'espace d'états  $\mathcal{X}$  du processus externe.

Le jeu d'équations (3.12)-(3.13) forme une équation d'*observation* où la nouvelle estimation  $f(\mathbf{x}_{t+1}|\mathbf{x}_{0:t}, \mathbf{I}_{1:t+1}, \mathbf{u}_{1:t})$  se déduit de l'estimation précédente  $f(\mathbf{x}_t|\mathbf{x}_{0:t-1}, \mathbf{I}_{1:t}, \mathbf{u}_{1:t-1})$  et de l'observation des nouveaux signaux  $\mathbf{u}_t$  et  $\mathbf{I}_{t+1}$ . L'espace  $\mathcal{X}$  est appelé l'espace des *sources* de l'observation.

**Erreur de prédiction** L'équation (3.12) représente la première étape de l'estimation aboutissant à une prédiction “interne” sur l'espace des sources. Cette prédiction est combinée (eq. 3.13) à la vraisemblance  $Q$ , qui peut être vue comme une prédiction “inverse” sur l'espace des sources. L'écart entre les deux distributions est l'*erreur de prédiction*, également appelée *innovation* (Kalman, 1960). Les différences de variance entre les distributions permettent de privilégier l'un ou l'autre selon que le bruit d'état ou le bruit de mesure dominant le processus d'observation.

**Remarque :** Un cas particulier important est le modèle sans transition d'état ( $\forall t, \mathbf{x}_{\text{out}}(t) = \mathbf{x}_{\text{out}}$ ). Ce cas correspond à des mesures bruitées multiples opérées sur les mêmes données. Le fait de multiplier les observations permet de réduire à chaque mesure l'incertitude sur l'état “réel” observé. Ce mécanisme, est appelé *accumulation d'évidence*. Dans le cadre de campagnes de mesures, il permet de décider à chaque mesure s'il est nécessaire d'en effectuer une nouvelle, en fonction d'un seuil de fiabilité fixé (Wald, 1947).

**Exemples :**

- Le filtre de Kalman (Kalman, 1960) utilise des modèles de bruit Gaussiens, définis par une moyenne et une matrice de variance-covariance. L’environnement est modélisé comme un processus Gaussien soumis à une commande extérieure connue.
- Les modèles de Markov cachés (Baum & Petrie, 1966) utilisent une représentation d’état discrète. L’environnement est une chaîne de Markov connue. Chaque état de la chaîne génère une mesure selon une distribution connue. Le problème consiste principalement, à partir de la mesure, à identifier les probabilités de transition d’état (algorithmes forward-backward, de Viterbi,...).
- Les filtres particuliers (Doucet et al., 2001) utilisent une approche de Monte Carlo pour échantillonner l’espace d’état du processus externe à chaque instant. Chaque “particule” suit la dynamique stochastique du processus externe  $P(\mathbf{x}_{t+1}|\mathbf{x}_t, \mathbf{u}_t)$  et se voit attribuer un score de vraisemblance selon  $Q(\mathbf{I}_{t+1}|\mathbf{x}_{t+1})$ .

### 3.3 Approche mixte : modèles et problèmes

Nous abordons ici les problèmes d’apprentissage dans le contexte de la commande motrice. La plupart des modèles s’inscrivent dans le cadre de la commande et de la mesure décrits précédemment, c’est-à-dire qu’ils reposent sur des modèles de l’environnement physique. L’apprentissage, en ce sens, se justifie lorsque, d’une manière ou d’une autre, les modèles ne sont pas considérés satisfaisants et doivent être “améliorés”.

Nous cherchons également à montrer ici le caractère problématique du pattern matching lorsqu’il s’agit de contrôler un corps dans un environnement, et le caractère problématique de l’approche systèmes dynamiques lorsqu’il s’agit de projeter un feedback moteur sur des espaces sensoriels complexes (modèle forward). Peut-on contrôler efficacement un corps sans faire appel à la notion d’attracteur et de feedback négatif? Peut-on s’adapter à un environnement autrement qu’en extrayant des “dictionnaires” d’invariants perceptifs et moteurs de l’environnement?

#### 3.3.1 Approche probabiliste

Dans le cadre du contrôle tel que présenté précédemment, l’apprentissage consiste donc à modifier le contrôleur décrit par la fonction  $\phi_{\text{in}}$  (voir l’équation (3.3), page 36, et l’équation (3.7), page 38) dans un sens qui améliore un critère de performance prédéfini. La mise à jour d’un modèle ou d’une fonction de réponse consiste en général à modifier un jeu de paramètres  $\Theta$  qui décrit le modèle.

Cette approche repose sur un modèle *génératif*  $P(X|\mathbf{x}, \mathbf{u}, \Theta)$  et  $Q(I|\mathbf{x}, \Theta)$ , où  $\Theta$  est le jeu de paramètres définissant les deux distributions. Dans le cas des modèles de filtrage, il s’agit par exemple d’améliorer les modèles  $P$  (processus externe) et  $Q$  (mesure), afin de mieux prendre en compte les facteurs exogènes non modélisés et de mieux prédire les entrées. Dans le cadre que nous avons regardé jusqu’à présent (modèles de la mesure, filtrage), l’état extérieur  $\mathbf{x}_{\text{out}}$  est l’état caché “latent” à estimer. Le produit du processus

de mesure est la distribution *a posteriori*  $\hat{f}$  portant sur l'espace des variables latentes, paramétré par  $\Theta$ , i.e.  $\hat{f}(\mathbf{x}) = f(\mathbf{x}|\mathbf{I}, \mathbf{u}, \Theta)$ . L'amélioration du processus repose alors sur le fait de trouver de nouveaux paramètres  $\Theta'$  en estimant la vraisemblance  $L(\mathbf{x}, \Theta')$ <sup>1</sup> des paramètres étant donné l'état latent  $\mathbf{x}$ , en maximisant :

$$\max_{\Theta'} \int L(\mathbf{x}, \Theta') \hat{f}(\mathbf{x}) d\mathbf{x} \quad (3.14)$$

ce qui permet, de manière récurrente, de produire une nouvelle distribution *a posteriori*  $\hat{f}'$  qui elle-même produit un nouveau jeu de paramètres  $\Theta''$  etc. Cette succession d'estimations et de maximisations donne son nom à l'algorithme EM (Dempster et al., 1977) qui, étant données les observations, converge vers une distribution limite  $\hat{f}^*$ .

**Contrôle avec consigne** Dans le cas du contrôle avec consigne (voir section 3.2.1), l'amélioration consiste à faire en sorte que le contrôleur obéisse mieux à la consigne, autrement dit minimise l'*erreur motrice*  $\mathbf{x}^* - \mathbf{x}_{\text{out}}$ . On considère ici que le jeu de paramètres  $\Theta$  définit la fonction de transfert du contrôleur.

Ce problème s'exprime comme :

$$\min_{\Theta'} \int \|\mathbf{x}^* - \mathbf{x}\|^2 \hat{f}(\mathbf{x}) d\mathbf{x}$$

**Contrôle via un signal de renforcement** Dans le cas du contrôle sans consigne, l'apprentissage peut se baser sur un signal qualitatif issu d'une mesure effectuée sur le dispositif physique. Cette mesure, est assimilée à un signal de récompense ("reward"  $r$ ) ou de punition ("loss"  $l$ ). Le contrôleur ne connaît pas  $\mathbf{x}^*$ , et se base donc sur cette information qualitative pour définir sa réponse motrice, afin de maximiser le gain (et/ou minimiser la perte), soit :

$$\max_{\Theta'} r \quad \left( \text{ou } \min_{\Theta'} l \right)$$

Un cas particulier très étudié en contrôle est celui de la récompense différée (temporal credit assignment), auquel cas ce n'est pas la récompense directe que l'on optimise mais une valeur estimée de son espérance, soit :

$$\max_{\Theta'} \mathbb{E}(r)$$

- Si l'environnement est connu, ce problème est celui de la *programmation dynamique* (Bellman, 1956).
- Si l'environnement est inconnu, on parle d'apprentissage *par renforcement* (Sutton, 1988).

**Contrôle sans consigne** Une idée récente, proposée par (Friston, 2010), est de définir à la fois l'apprentissage et le contrôle comme un problème d'optimisation dans un cadre conceptuel lié à la prédiction des signaux. Il s'agit de façonner l'activité interne de manière

---

1. égale à  $Q(\mathbf{I}|\mathbf{x}, \Theta')$  quand il n'y a pas de dynamique d'état et à  $Q(\mathbf{I}_{t+1}|\mathbf{x}_{t+1}, \Theta')P(\mathbf{x}_{t+1}|\mathbf{x}_t, \mathbf{u}_t, \Theta')$  quand il y a une dynamique d'état.

à rendre le signal arrivant à l’instant  $t$  le plus prédictible possible, d’être le moins surpris, et d’utiliser l’erreur de prédiction pour améliorer autant que possible les prédictions futures, soit par un mécanisme de plasticité, soit par le déclenchement d’une action visant à réduire l’écart à la prédiction.

L’absence de consigne conduit à effectuer l’optimisation à la fois sur les paramètres du modèle et sur la commande, soit, en reprenant les notations précédentes :

$$\max_{\Theta', \mathbf{u}'} \int L(\mathbf{x}, \Theta', \mathbf{u}') \hat{f}(\mathbf{x}) d\mathbf{x} \quad (3.15)$$

où la distribution a posteriori  $\hat{f}(\mathbf{x})$  est générée selon un principe de “codage par population” (Georgopoulos et al., 1986; Averbek et al., 2006), autrement dit via un échantillonnage de  $P_{\Theta}$ .

Ce cadre offre une interprétation relativement nouvelle au déclenchement des actions. Le point d’opération du système de contrôle est construit sur une quantité abstraite qui est la distance entre la prédiction et l’observation. La perte de la notion de consigne *rend non nécessaire la référence à un milieu physique extérieur*. L’action sert à produire une scène sensorielle plus conforme aux attentes exprimées par le modèle forward, selon un principe de *persévération*. L’activité interne suggère une scène sensorielle, qu’il s’agit de réaliser (ou de produire) à l’aide du déplacement moteur (ou plus généralement de l’envoi d’un signal sur une interface) approprié. La question n’est plus la capacité à reproduire fidèlement le monde extérieur, mais la capacité à prédire ses manifestations sensorielles.

### 3.3.2 Modèles bio-inspirés

**Observations** Par rapport au cadre classique de la commande, la difficulté en modélisation biologique repose sur la multiplicité des réalisations motrices accomplies par le couplage cerveau-environnement. L’étude du système nerveux central suggère en effet une organisation parallèle, avec de nombreux circuits dédiés à des tâches distinctes. L’étude d’organismes relativement primitifs permet d’identifier les tâches élémentaires, celles qui sont présentes à tous les niveaux du règne animal et qui structurent l’organisation du système nerveux central. On distingue ainsi classiquement :

- l’orientation (déplacement les organes des sens vers les stimuli d’intérêt, éloignement et évitement des stimuli aversifs),
- la locomotion (déplacement du corps),
- et l’ensemble des comportements consistant à utiliser les ressources de l’environnement (activité alimentaire par exemple) (Brembs, 2003; Grillner et al., 2008; Stephenson-Jones et al., 2011).

Ces tâches globales sont réalisées par l’activité coordonnée de différents circuits s’étendant à l’ensemble des structures de l’encéphale :

- un circuit parmi les plus primitifs est celui de la sélection de l’action organisé autour des structures sous-corticales (ganglions de la base) fonctionnant par désinhibition sélective de choix d’actions concurrents, comme mis en évidence par (Grillner et al., 2008).
- La dérivation de la commande motrice à travers le cervelet conduit par ailleurs à des ajustements posturaux associés à un contrôle fin et adaptatif des commandes motrices (Wolpert et al., 1998; De Zeeuw et al., 1998).

- D’autres circuits plus récents, comme ceux de l’hippocampe, sont impliqués dans la navigation spatiale, et la mémorisation (McClelland et al., 1995), exprimant en particulier la situation de l’animal dans un environnement connu mémorisé (son territoire) d’un point de vue extérieur, implémentant un mécanisme d’orientation dans un référentiel extéro-centré (O’Keefe & Nadel, 1978).
- Le cortex, région anatomiquement la plus récente, est principalement lié au traitement sensoriel et aux commandes motrices fines. Les boucles entre le cortex, le striatum et le cervelet semblent également associées chez certains mammifères à la simulation motrice précédent le déclenchement de l’action (Jeannerod, 2001), etc.

**Modèles** Le principe de contrôle par rétroaction négative (*negative feedback control*) correspond à une commande directement proportionnelle à l’erreur motrice, autrement dit  $\mathbf{u}(t) \propto \mathbf{x}^* - \mathbf{x}_{\text{out}}(t)$ . En tenant compte de l’inertie et des temps de réponse, il permet d’obtenir un contrôleur homéostatique stable (résistant aux perturbations).

Il existe de nombreux modèles utilisant les principes du contrôle adaptatif pour modéliser l’activité nerveuse. On citera parmi les plus connus (Robinson, 1975; Wolpert et al., 1998). Les schémas de guidage par correction d’erreur sont par ailleurs présents au niveau des structures sensori-motrices profondes, comme dans le cadre du contrôle vestibulaire (Ito, 1982).

On oppose généralement au contrôle par feedback négatif (également appelé *poursuite*) le contrôle dit *balistique* dans lequel l’erreur motrice n’est pas mesurée. Le contrôle repose sur la consigne  $\mathbf{x}^*$  seule. On parle aussi de contrôle en position.

Plusieurs méthodes permettent d’implémenter ce type de contrôle :

- Dans le cadre du contrôle linéaire, la commande balistique repose sur l’inversion de la fonction de transfert de l’environnement. Si  $\mathbf{x}^*$  est la position finale désirée, on a  $\mathbf{u} = \phi_{\text{out}}^{-1}(\mathbf{x}^*)$ . On parle de commande par modèle inverse.
- Le cas des modèles non-linéaires peut être traité via la combinaison de “primitives motrices” dans l’espace de la tâche (Mussa-Ivaldi & Solla, 2004).

Le contrôle balistique présente plusieurs propriétés intéressantes :

- la commande  $\mathbf{u}$  reste constante pendant toute la durée d’exécution.
- la commande est insensible aux conditions initiales. Quelle que soit la position initiale du contrôleur, la commande conduit “mécaniquement” le système contrôlé vers son point fixe  $\mathbf{x}^*$ .
- le contrôleur est insensible à ses entrées sensorielles. Autrement dit la conséquence sensorielle de la commande n’est pas utile pour le contrôle lui-même. Le signal mesuré sur les capteurs peut néanmoins conduire le contrôleur à *réviser la consigne* si le signal d’entrée n’est pas conforme au signal attendu.

Une version “dynamique” du contrôle balistique a été proposée par (Bicho & Schöner, 1997), dans laquelle :

- la commande motrice est implémentée sous la forme d’un champ de potentiel dans l’espace de la tâche (combinaison de champs de force) ;
- la consigne évolue continuellement en fonction des entrées sensorielles.

Enfin, certaines activités motrices (périodiques en particulier) semblent reposer sur un mécanisme de feedback *positif*, dans lesquels le milieu physique et la commande motrice forment un couple qui s’auto-renforce, comme dans le cas de la locomotion. Cette activité rythmique spontanée peut être modulée par divers facteurs (contacts avec le sol ou avec des obstacles, orientation visuelle, etc.).

- La rétroaction positive correspond à la résonance entre des schémas moteurs internes concurrents, la disposition des effecteurs et l’environnement.
- Les transitions entre différentes tâches correspondent à la déstabilisation d’un schéma moteur, suivi d’une relaxation sur un autre schéma moteur, suite à une perturbation<sup>2</sup>. Un mouvement balistique simple peut être vu par exemple comme la transition d’une posture inadaptée vers une posture plus adaptée aux contraintes exprimées par l’environnement.

Les mécanismes de contrôle par feedback positif ont jusqu’à présent principalement été étudiés sous la forme d’oscillateurs simples ou couplés (Haken et al., 1985; Sternad, 2008), servant de support à des schémas moteurs périodiques, comme celui de la locomotion. L’idée de la sélection d’un canal de réponse par résonance avec les signaux de l’environnement de trouve néanmoins déjà dans des articles plus anciens (Greene, 1962). Les mécanismes neuronaux à la base des oscillateurs comportementaux, présents au niveau de la moëlle épinière, commencent également à être bien connus (Grillner et al., 2008). La rétroaction positive, et les effets de mémoire qu’elle permet, est probablement un composant essentiel des opérations réalisées dans le cerveau (Compte, 2006). Le mécanisme dit de “réverbération” (Hebb, 1949) a été modélisé de longue date (Wilson & Cowan, 1972; Amari, 1977; Hopfield, 1982), mais reste peu utilisé dans les architectures de contrôle.

Le manque de développements applicatifs vient probablement

- d’une difficulté pratique à doser le seuil de réponse aux changements de l’environnement, ce seuil dépendant du modèle choisi et du problème posé. En particulier, avec une sensibilité trop faible, le réseau devient résistant au changement (autrement dit ignore le signal d’entrée).
- Les mécanismes de feedback positif introduisent une dose d’autonomie dans les architectures de contrôle, rendant la réponse du réseau dépendante du contexte. Cet effet de réponse contextuelle (ou de rémanence) et offre bien sûr une plus grande panoplie de comportements, comme par exemple la mémoire à court terme, mais augmente la difficulté d’analyse du contrôleur (cadre non-Makovien), et donc de la connaissance des risques de dysfonctionnement.

---

2. La question est de savoir si le mécanisme de déstabilisation/restabilisation définit (ou non) un lien causal entre l’ancienne configuration et la nouvelle configuration. L’ancienne configuration est-elle oubliée pour laisser place à la nouvelle, ou existe-t-il une relation de similarité ou de proximité qui favorise le passage de l’une à l’autre? Le fonctionnement par feedback positif décrit-il une activité versatile, constituée de “modes comportementaux” successifs détachés les uns des autres, ou une activité séquentielle, dans laquelle le nouveau mode est “dédruit” de l’ancien (où l’ancien mode forme un contexte favorable à l’apparition du nouveau; où l’ancien mode persiste dans le nouveau)?

# 4

## Panorama des principaux résultats

Les travaux présentés dans ce mémoire offrent des points d'entrée vers le développement d'architectures de contrôle utilisant les principes d'appariement dynamique évoqués dans les deux chapitres précédents. Pour les développer, nous avons utilisé un certain nombre de parti pris méthodologiques :

- Utiliser le cerveau comme source d'inspiration. Nos travaux reposent largement sur la littérature et les observations en neurosciences. Nous considérons en particulier la plasticité de Hebb comme modèle de référence pour nos algorithmes d'apprentissage.
- Nous adoptons l'hypothèse d'architectures de contrôle reposant sur l'opérateur d'appariement. En particulier nous cherchons à généraliser le principe d'appariement ("pattern matching") à l'œuvre en apprentissage automatique au cas des systèmes dynamiques non-linéaires bouclés.
- Nous partons d'une hypothèse de calcul *distribué* implémenté par les réseaux de neurones récurrents, c'est-à-dire que l'ensemble du calcul est réparti à parts égales sur les différents nœuds du réseau. Le "temps de calcul" correspond au temps de relaxation vers un attracteur. On parle aussi de réponse de population (ne dépendant pas d'un neurone seul).

Par ailleurs :

- Les études sur la plasticité synaptique reposent largement sur la simulation numérique. Les simulations numériques visent à établir des comportements d'ensemble difficilement déductibles du modèle, en particulier lorsque les modifications locales introduites par la plasticité synaptique ont des conséquences sur la dynamique d'ensemble.
- Nous validons principalement les modèles d'apprentissage proposés sur des tâches liées au contrôle moteur. Nous cherchons à établir les conditions dans lesquelles un réseau de neurones peut acquérir des compétences de contrôle sur l'appareil moteur et/ou sur des périphériques de sortie. Ces validations impliquent :
  - des problèmes "jouets" dont nous maîtrisons les paramètres,
  - des architectures neuronales directement inspirées par les observations biologiques,
  - des applications relevant du domaine de l'ingénierie (interfaces cerveau-machine

en particulier).

Ces travaux sont présentés en détail dans la partie II. Je liste ici quelques résultats que je considère comme significatifs :

**a. Un réseau récurrent aléatoire en régime critique et soumis à un signal extérieur catégorise arbitrairement son environnement sensoriel.**

L'étude proposée dans (Beslon & Daucé, 2002), également rapportée dans (Daucé, 2007), porte sur un réseau récurrent aléatoire comportant une seule population (voir section 2.4.2, page 28). Le paramètre de contrôle choisi place le réseau dans une région chaotique, proche de la transition vers le cycle limite. Le réseau est soumis à une stimulation extérieure quasi stationnaire, évoluant à une vitesse très lente par rapport à la vitesse de mise à jour du réseau. Une succession de régimes dynamiques distincts, séparés par des transitions brusques, est mise en évidence. Chaque régime est caractérisé par une distribution d'activité différente, qui se manifeste par la présence de fréquences caractéristiques distinctes dans le spectre de Fourier du signal moyen (voir section 6.2.2, page 78). La différence qualitative entre les différents régimes observés (différence de fréquence principale) peut servir de support à une catégorisation de l'espace des entrées.

**b. L'hétérogénéité du graphe de connexions défavorise le développement d'oscillations synchronisées dans les grands réseaux équilibrés.**

Les réseaux de neurones dits "équilibrés" (balanced neural networks) sont des réseaux constitués d'une population de nœuds inhibiteurs et d'une population de nœuds excitateurs. Le terme équilibre se réfère ici à l'idée d'équilibre dynamique entre deux forces opposantes : la dynamique du réseau est amenée à converger vers un point d'équilibre traduisant localement (et globalement) le "poids" des deux influences. Une configuration des poids se traduisant par une activité endogène soutenue traduit une domination de l'influence excitatrice. A l'inverse, une configuration où l'activité intrinsèque disparaît après un certain temps traduit une domination de l'activité inhibitrice. Comme dans le cas des réseaux aléatoires simples, il existe un paramètre de contrôle correspondant au rapport entre l'influence excitatrice et l'influence inhibitrice. De manière intéressante, une grande variété de comportements dynamiques peuvent être observés à la frontière, et entre l'extinction et l'activité soutenue, comme les oscillations synchronisées (Brunel, 2000a) et/ou le chaos (Van Vreeswijk & Sompolinsky, 1998).

En étendant les équations de champ moyen au cas des populations multiples, nous étudions dans (Daucé et al., 2001) un régime de faible activité (où l'inhibition domine l'excitation). Nous regardons le comportement de deux paramètres de contrôle : l'amplification et l'hétérogénéité d'un faisceau d'axones (son coefficient de variation), jouant sur la nature des régimes développés (extinction, activité chaotique non spécifique et oscillations synchronisées). Nous mettons en évidence le caractère régulateur du paramètre d'hétérogénéité dans les régimes de forte amplification (dits aussi de "haute conductance") : l'hétérogénéité des liens défavorise le développement de dynamiques d'ensemble simples (type oscillations synchronisées) et favorise des signaux individuels fortement indépendants (voir section 6.2.1 page 76 et annexe B).

**c. Les graphes de connexions à la fois topographiquement organisés et équilibrés permettent d’implémenter une mémoire à court terme sur des neurones impulsionsnels, sans faire appel à des mécanismes cellulaires spécifiques.**

La question abordée par le modèle publié en 2004 (Daucé, 2004b) est celle de l’implémentation d’un modèle de champ neuronal sur un support composé d’unités discrètes. Sur les modèles jusqu’alors proposés, dans un cas l’activité focale était soit uniquement réactive (Hansel & Sompolinsky, 1996), dans l’autre cas, l’activité persistante reposait sur un mécanisme de bistabilité cellulaire (Camperi & Wang, 1998) ou encore sur une plasticité synaptique à court terme (Compte et al., 2000). Notre modèle est une des premières implémentations d’un champ neuronal sur support discret reposant sur des mécanismes de réseau uniquement.

Le modèle proposé dans (Daucé, 2004a; Daucé, 2004b) (voir section 6.2.3 page 81 et annexe D) met en évidence la présence simultanée d’une activité de fond chaotique, de grandes oscillations lentes et de foyers d’activité de type champ neuronal, sur un réseau de neurones composé d’unités neuronales binaires. Comme dans le cas du champ neuronal continu (Amari, 1977), l’activité persistante repose sur l’activité excitatrice locale, avec un effet de lissage lié à l’hétérogénéité des connexions. Les oscillations lentes reposent sur l’existence de délais différenciés. Ce sont les oscillations qui donnent au réseau la capacité de répondre aux changements de l’entrée, permettant des basculements d’activité lorsque ces changements sont forts, ou encore une mémoire à court terme (rémanence) lorsque le signal d’entrée s’éteint.

**d. Les graphes de connexion intra-cérébraux, tels qu’observés en imagerie fonctionnelle, peuvent servir de support à des dynamiques multistables.**

L’étude proposée dans (Golos et al., 2015) (voir section 6.2.5 page 84, et annexe P) étudie la dynamique de relaxation d’un réseau de neurones dont le graphe de connexions est issu de l’analyse des faisceaux de matière blanche à l’échelle du cerveau entier (Hagmann et al., 2008).

La dynamique, inspirée par (Hopfield, 1984), est définie par l’équation différentielle :

$$\tau \dot{x}_i = -x_i + \frac{1}{2} \sum_{j=1}^N J_{ij} (1 + \tanh(g(px_j - \theta_j))) \quad \text{avec} \quad \theta_i = \frac{1}{2} \sum_{j=1}^N J_{ij}$$

où  $g$  est le gain,  $p$  une constante de proportionnalité (ratio excitation/inhibition) et  $\mathbf{J}$  la matrice des couplages issue de l’imagerie fonctionnelle humaine.

Ces travaux ont permis de mettre en évidence un nombre très important d’attracteurs dans certaines gammes de paramètres. L’analyse des ensembles d’attracteurs obtenus permet d’identifier une huitaine de “modes” différents, chaque mode correspondant à l’activation d’un réseau anatomique particulière. Ces modes peuvent être comparés aux modules régionaux mis en évidence par une analyse structurelle (Hagmann et al., 2008) ainsi qu’avec les composants indépendants identifiés sur les signaux fMRI de l’état de repos (Damoiseaux et al., 2006). Ils permettent de simuler certains aspects de la dynamique large échelle (en particulier l’alternance entre différents patrons d’activité au cours du temps).

### e. La durée d'une commande balistique peut être contrôlée par un mécanisme d'intégration sous-liminaire

L'orientation visuelle est un comportement moteur très spécifique, qui s'apparente plus au monde de la perception qu'au monde de la motricité, reposant sur des circuits de contrôle distincts des circuits de commande de l'appareil moteur. La saccade oculaire est un mouvement présent chez l'ensemble des vertébrés consistant à orienter la partie centrale de la rétine (la fovea) vers un point d'intérêt de l'espace visuel. Le comportement d'orientation saccadique, par opposition aux comportements locomoteurs ou alimentaires, est constitué de mouvements brefs (discrets), permettant d'explorer la scène visuelle. La saccade oculaire est le prototype des mouvements dits "balistiques", partant d'une posture initiale et s'achevant sur une posture finale (par opposition à des mouvements d'ajustement ou de contrôle guidé par un feedback comme la poursuite oculaire par exemple).

La structure au carrefour de la transformation visuo-motrice est une carte rétinotopique située sur la partie supérieure du tectum, appelée le colliculus supérieur. L'étude présentée dans (Daucé et al., 2011) (voir page 115 et annexe L) montre l'existence d'une activité autonome contrôlée dans le temps et dans l'espace sur la durée de la saccade (100 ms environ), c'est-à-dire sur un intervalle temporel précis, "programmé". Notre étude, en implémentant de façon détaillée des populations de neurones à impulsions,

- simule l'activité autonome d'une carte colliculaire bidimensionnelle constituée de plusieurs milliers de neurones,
- implémente un mécanisme de décision par "levée d'inhibition" ,
- contrôle l'activité sur la carte dans le temps et dans l'espace via des mécanismes d'intégration,
- utilise, pour le contrôle de la durée (100 ms environ), des neurones à constante de membrane lente, faisant l'économie d'un mécanisme d'intégration par population.

Nous montrons donc que des mécanismes d'intégration sous-liminaires (en dessous du seuil de tir) peuvent contrôler l'intégralité du circuit de la saccade oculaire, et plus généralement la durée de mouvement balistiques dont la durée n'excède pas la centaine de millisecondes.

### f. La plasticité de Hebb, appliquée à des réseaux de neurones récurrents aléatoires, provoque une réduction de complexité des dynamiques spontanée et induites

L'appariement par relaxation/réduction, sous sa forme la plus simple, s'exprime comme une transition depuis une activité à haut degré de liberté (chaos) vers une activité plus simple (cycle limite) en présence d'un stimulus statique (Daucé et al., 1998; Henry & Daucé, 2008b).

Les règles de plasticité Hebbiennes définissent une dynamique sur les liens synaptiques (voir les équations (2.9), (2.10) et (2.11), page 26), selon les valeurs des activités pré et post-synaptiques. Dans la forme simple étudiée ici, la plasticité repose soit sur la conjonction des activités pré et post-synaptiques, i.e :

$$\dot{J}_{ij} = F(s_i^{\text{out}} \times s_j^{\text{in}}, J_{ij}) \quad (4.1)$$

(on parle alors de règle d'ordre 0), soit sur des différences d'activité, i.e.

$$\dot{J}_{ij} = F((s_i^{\text{out}} - m_i^{\text{out}}) \times (s_j^{\text{in}} - m_j^{\text{in}}), J_{ij}) \quad (4.2)$$

où  $m_i^{\text{in/out}}$  est une activité de référence, par exemple l'activité moyenne. On parle dans ce cas de règle d'ordre 1. Ces règles de plasticité ne sont pas nécessairement convergentes. Elles sont généralement appliquées sur des durées brèves, en fonction des études (par exemple jusqu'à l'apparition d'un changement de régime dynamique).

Dans le cadre de systèmes dynamiques ouverts (c'est-à-dire soumis à un signal d'entrée, voir l'équation (2.8) page 24), nous mettons en évidence deux mécanismes de réduction distincts :

- L'article (Daucé et al., 1998) (voir section 7.2.1 page 88 et annexe A) met en évidence une simplification systématique de l'activité intrinsèque dans le cas d'une règle d'ordre 0. La réponse au motif d'entrée change sous l'effet de la plasticité : en partant d'un régime chaotique, la dynamique évolue vers un cycle limite, puis, de manière asymptotique, vers le point fixe. On parle de "réduction" de la dynamique, la plasticité produisant une "route inverse" de la complexité vers le point fixe (figure 7.1, page 89).
- Dans le cas des règles d'ordre 1 (par exemple des différences temporelles (Daucé, 2007), ou la règle de la covariance (Daucé et al., 2002)) la plasticité produit dans les mêmes conditions une transition depuis un régime irrégulier chaotique vers un régime périodique, tendant asymptotiquement vers de grandes oscillations sur la dynamique d'ensemble.

La plasticité a donc pour effet de capturer (ou plutôt accentuer) un *mode dynamique* particulier. Elle a également pour effet de séparer l'espace des entrées entre régions "connues" et régions "inconnues" : à l'issue de l'"apprentissage", le réseau est capable de discriminer les entrées similaires au motif appris, en produisant une réponse *spécifique*, et les entrées non-similaires, en produisant une réponse *générique*. Le réseau peut donc être vu comme un "réservoir" de réponses qualitativement distinctes, construites par le mécanisme de plasticité en présence de signaux présentés de manière répétée.

En détail, nous faisons les propositions suivantes :

- Un réseau récurrent (en régime "persistant") se comporte comme un "réservoir" de patrons d'activité (ou "assemblées" de neurones). Chaque patron d'activité correspond à l'activation d'un circuit spécifique sur le graphe.
- Le calcul repose sur l'appariement entre l'entrée (le signal extérieur) et un des patrons d'activité activables sur le graphe. On parle de *reconfiguration* du circuit.
- Cet appariement se traduit par une réduction de complexité correspondant à la relaxation du système dynamique vers un attracteur de dimension faible.
- Cette réduction sélective de complexité repose sur l'action de la plasticité de Hebb sur le graphe en présence de signaux contenant "suffisamment" de régularités.

La principale difficulté rencontrée pour ces réseaux est la contamination de la réaction induite à des patrons non appris. Si des motifs voisins du motif initialement appris induisent également une réduction de la dynamique, la répétition de l'apprentissage sur plusieurs motifs différents conduit invariablement à un effondrement généralisé de la complexité, étendant la réduction de dynamique à l'ensemble des motifs. La capacité d'un

réseau est définie comme le nombre de motifs pouvant être appris sans provoquer cet effondrement. Cette capacité s'avère assez faible en pratique. Dans notre cas, comme dans le cas du réseau de Hopfield, elle semble dépendre linéairement de la taille du réseau.

**g. La plasticité de Hebb inscrit, dans des réseaux récurrents à deux couches, des fonctions de réponse auto-associatives sur des signaux d'entrée spatio-temporels périodiques.**

Les travaux présentés dans (Daucé, 2000; Daucé et al., 2002) implémentent un modèle de la perception de signaux spatio-temporels périodiques, sur un réseau récurrent aléatoire à deux couches, constitué d'une couche primaire sans dynamique intrinsèque et une couche associative possédant une dynamique intrinsèque, analogue aux réseaux récurrents aléatoires simples présentés précédemment (figure 7.2, page 92).

- Sur des réseaux récurrents de grande taille, les signaux périodiques sont “capturés” par la dynamique interne du réseau par “accrochage de fréquence” (Arnold, 1978). Ils activent spontanément un circuit spécifique sur la couche secondaire (figure 6.4, page 80).
- La plasticité, en présence d'un signal d'entrée spatio-temporel, régularise ce circuit d'activité interne. Du fait des correspondances statistiques entre cette activité interne et la couche d'entrée, la plasticité a également pour effet d'intensifier le motif spatio-temporel d'entrée via les liens de feedback (figure 7.3, page 93)
- L'aller-retour du signal entre la couche primaire et la couche secondaire conduit à une *résonance* (ou réverbération) entre l'entrée sensorielle et la couche interne.
- L'activité de réverbération constitue une réponse de population. La réponse est “activée” en présence de signaux concordants avec les formes spatio-temporelles apprises, et “désactivée” dans le cas contraire, conformément au principe d'appariement conditionnel (voir page 24).

La réponse induite par la plasticité correspond fonctionnellement à un mécanisme de filtrage (voir page 38), autrement dit une *prédiction dans le présent*. Le feedback correspond à la prédiction actuelle, tandis que l'entrée sensorielle correspond à l'observation effective. Tout écart entre la prédiction et l'observation fait entrer le système dans un conflit qui peut se résoudre de deux manières :

- Si l'écart est faible, le système tend à ignorer ce changement et persévère dans sa prédiction. Si une partie du signal disparaît par exemple, l'activité interne est capable de compléter les morceaux manquants, selon un mécanisme auto-associatif qui s'apparente à celui des réseaux de Hopfield (Hopfield, 1982). La persévérance correspond également à un mécanisme de mémoire à court terme, conservant une mémoire de l'activité passée dans l'activité présente.
- Si l'écart est important et/ou réitéré, le système met à jour son activité interne (son “modèle”) afin de mieux anticiper l'entrée. L'irruption de composants sensoriels dissonants favorisent la désagrégation de l'activité résonante. Le réseau retourne alors à son état de base (non réverbérant).

Au cours du temps, les caractéristiques du signal présenté en entrée conduisent à des alternances de décomposition et de recombinaison de l'activité réverbérante. On peut également assimiler ces transitions au passage d'un état “bas” (non spécifique)

à un état “haut” (spécifique), puis au retour vers un état “bas”, etc., un peu comme dans le modèle de (Brunel, 2000b) (voir aussi étude **i**).

### **h. Dans les réseaux récurrents aléatoires de neurones à impulsions, la STDP permet de construire une fonction de réponse arbitraire sur un espace d’entrée de séquences d’activation.**

La plasticité dépendante des instants de tir (*Spike-timing dependent plasticity* – STDP) est une règle de plasticité dépendant de l’ordre de tir des neurones pré-synaptiques et post-synaptiques (Bi & Poo, 1998). Cette règle est facilitatrice pour les synapses où l’activité pré-synaptique précède causalement l’activité post-synaptique, et, inversement, suppressive pour les synapses où l’activité post-synaptique précède (de manière anti-causale) l’activité pré-synaptique. Cette règle tend donc à favoriser les conjugaisons d’activité causalement liées, mais surtout, de manière symétrique, à empêcher les activités non causalement liées.

Une formulation possible de la règle est la suivante :

$$\dot{J}_{ij} = F((s_i^{\text{out}} \times m_j^{\text{in}}) - (s_j^{\text{in}} \times m_i^{\text{out}}), J_{ij})$$

où  $s_j^{\text{in}}$  est l’activité pré-synaptique,  $s_i^{\text{out}}$  est l’activité post-synaptique,  $m_j^{\text{in}}$  une trace exponentielle de l’activité pré-synaptique, et  $m_i^{\text{out}}$  une trace exponentielle de l’activité post-synaptique. Cette règle peut formellement être reliée aux règles de plasticité d’ordre 1 (voir page 49).

La discrimination entre les différentes séquences temporelles repose sur l’entretien au niveau de la couche récurrente d’une activité qui permet de conserver la mémoire de faits passés, sur le principe du réservoir déjà évoqué (Maass et al., 2002). Le caractère très désordonné des activités individuelles dans les réseaux de neurones à impulsions rend difficile la capture de dépendances aux échelles temporelles fines. Dans le cas d’une activité fortement irrégulière (chaos profond), la population de neurones semble en effet garder une mémoire assez faible des faits passés (autocorrélation du même ordre de grandeur). Les réponses induites par la STDP sur des entrées statiques s’apparentent à des “bursts” (bouffées) d’activité, qui impliquent une grande partie des neurones du réseau (Henry & Daucé, 2008b) (voir section 7.2.3 page 94 et annexe H), peu spécifiques et peu différents des comportement obtenus avec des neurones fréquentiels.

L’étude présentée dans (Henry et al., 2007) (voir section 7.2.3 et annexe F) montre néanmoins le caractère séquentiel de l’activité induite par la STDP, support ici de la mémoire à court terme. L’activité auto-entretenu permet la prise en compte de dépendances entre des événements séparés par plus de 100 ms. L’application alternée des plasticités Hebbienne et anti-Hebbienne ouvre par ailleurs la voie à l’implémentation de mécanismes d’apprentissage par renforcement dans les réseaux de neurones récurrents, présentés plus loin.

La présence de séquences d’activation spatio-temporelles reproductibles (dites “groupes polychrones”), telles que postulées par (Izhikevich, 2006) en tant qu’unités opérationnelles élémentaires, peut également être mise en évidence dans les réseaux aléatoires. Une étude que j’ai publiée récemment, utilisant des unités neuronales plus détaillées (Daucé, 2014) (voir annexe M) montre que la réponse de population à des stimuli spatio-temporels suit

une séquence reproductible (mais bruitée) qui peut être amplifiée par la plasticité (voir figure 7.5, page 98). La STDP identifie bien des séquences d’activation reproductibles dans le signal d’entrée, rendant possible l’apprentissage et la ségrégation par le réseau des séquences d’activité spécifiques, entre les signaux produisant une réponse de type bouffée, et ceux qui ne le font pas. La STDP induit donc bien une structuration de l’activité interne en amplifiant certaines séquences d’activation présentées de manière répétée (mais non périodiques) en entrée du réseau, favorisant l’émergence de séquences d’activation stables, comme postulé par (Izhikevich et al., 2004; Izhikevich, 2006) (et contrairement aux observations de (Morrison et al., 2007)).

### **i. L’apprentissage “par cœur” d’appariements sensori-moteurs spécifiques met en évidence, en milieu ouvert, des couplages transitoires entre la dynamique d’un réseau récurrent aléatoire et les trajectoires motrices.**

Certaines activités motrices (périodiques en particulier) semblent reposer sur un mécanisme de feedback positif, dans lesquels le milieu physique et la commande motrice forment un couple qui s’auto-renforce, comme dans le cas de la locomotion. Cette activité rythmique spontanée peut être modulée par divers facteurs (contacts avec le sol ou avec des obstacles, orientation visuelle, etc.) (voir aussi page 43).

L’expérience présentée dans (Quoy & Daucé, 2000; Daucé & Quoy, 2000; Quoy et al., 2001; Daucé et al., 2002) a été conduite sur la plateforme de contrôle robotique PerAc (Gaussier & Zrehen, 1995). Le robot, placé dans un environnement ouvert, apprend, par plasticité Hebbienne, à associer un ensemble d’amers visuels avec des réponses motrices imposées (voir figure 8.2 page 107). Pendant le test, les petits décalages angulaires liés aux frottements introduisent un décalage progressif de la scène visuelle, et une perte de correspondance entre la prédiction et l’observation. On observe alors, au niveau comportemental, une alternance entre un comportement périodique, correspondant à la tâche apprise (“exploitation”), et un comportement irrégulier sans correspondance avec la tâche (“exploration”).

Cette expérience robotique offre un exemple de couplage entre activité mécanique (déplacements moteurs) et l’activité endogène du réseau de neurones. La commande ne repose pas sur la correction d’erreur, comme dans le contrôle traditionnel, mais sur l’appariement de la dynamique endogène avec des séquences spatio-temporelles apprises, et sur la persistance de cet appariement pour des versions dégradées du signal appris. L’erreur (la non-correspondance) ne produit pas une correction mais une transition vers une activité irrégulière qui s’apparente à une exploration aléatoire de l’espace des couplages sensori-moteurs jusqu’à atteindre un nouvel attracteur comportemental.

- Le mécanisme de synchronisation transitoire, ou de résonance, au sein de systèmes constitués de nombreuses unités couplées, est interprété comme favorisant la *communication entre les différentes échelles* (Daucé, 2002). Ce type de mécanisme n’est pas encore, à ma connaissance, implémenté dans des architectures de contrôle, mais son importance a déjà été souligné dans (König & Engel, 1995) et (Varela et al., 2001).

- La *non-reconnaissance* correspondrait à une situation où le milieu neuronal produit une activité à haut degré de liberté, sans effet macroscopique par absence de coordination.
- La *reconnaissance* correspondrait à une situation où le milieu neuronal “renonce” transitoirement à ses degrés de liberté et se coordonne, favorisant le couplage avec la dynamique macroscopique. La coordination de la population permettrait le passage à une description “grande échelle” de l’activité neuronale, perceptible au niveau de description supérieur, caractérisé par une échelle spatiale plus large et une échelle temporelle plus lente, le même mécanisme pouvant en principe se décliner à différentes échelles.

L’importance de la désynchronisation, comme complément de la synchronisation, est également souligné dans (Rodriguez et al., 1999). Le fonctionnement en boucle fermée par couplage positif (par renforcement mutuel) contient en effet fondamentalement le risque de l’enfermement dans un attracteur comportemental unique, correspondant à la répétition indéfinie de la même tâche. Au mécanisme de synchronisation doit donc correspondre un mécanisme de désynchronisation, qui peut reposer sur des facteurs déstabilisants liés à une mauvaise correspondance internes ou externe, mais également sur des mécanismes actifs de désynchronisation, comme la fatigue neuronale par exemple.

**j. Dans des espaces d’état et d’action continus, l’apprentissage par renforcement peut être implémenté via la plasticité Hebbienne et le principe de trace synaptique. La dynamique endogène joue le rôle d’un bruit d’exploration.**

L’apprentissage par renforcement définit un cadre formel à l’apprentissage de règles de contrôle moteur selon un principe d’essais/erreur. Le bruit moteur (réponse motrice aléatoire) joue un rôle important dans cette famille d’algorithmes. On parle d’*exploration motrice*. La fonction d’appariement probabiliste  $\pi_{\Theta}(\mathbf{u}|\mathbf{I})$ , qui associe des probabilités de réponses aux différentes sorties en fonction de l’entrée, est appelée la “politique”. Dans un système de contrôle en boucle fermée, la réponse effective  $\mathbf{u}$  induit alors une nouvelle observation  $\mathbf{I}'$  via l’échantillonnage du modèle inconnu  $P$ , qui elle-même induit une nouvelle réponse  $\mathbf{u}'$ , etc. A l’instant  $t$ , un signal de récompense  $r_t$  indique si cette séquence de choix était correcte ou pas. La séquence de décisions  $\mathbf{u}_{1:t}$  vise ici à maximiser l’espérance de  $r(t)$ , qui joue le rôle d’une “vraisemblance” sur l’espace des couplages sensori-moteurs  $(\mathbf{I}, \mathbf{u})$ .

Les algorithmes classiques utilisent un signal de renforcement pour guider l’apprentissage. L’algorithme du “policy gradient” (Williams, 1992; Peters & Schaal, 2008), par exemple, met à jour la fonction d’appariement par “apprentissage direct” selon une descente de gradient qui s’apparente à la recherche du “maximum de vraisemblance”.

Dans un contexte de réseau de neurones récurrents aléatoires, le caractère irrégulier de la dynamique intrinsèque sert de support à l’exploration motrice.

L’utilisation de récompenses ponctuelles nécessite de prendre en compte (et produire les modifications synaptiques liées à) la totalité de la trajectoire motrice (problème du “temporal credit assignment” (Sutton & Barto, 1998)). Ce principe peut être mis en place via un mécanisme de trace synaptique, c’est-à-dire une mémoire locale des

co-activations récentes pour appliquer la modification synaptique. Le principe de trace synaptique consiste à stocker l'historique long des corrélations d'activités pré et post-synaptiques dans une mémoire spécifique, selon le principe des “traces d'éligibilité” (Sutton, 1988), soit :

$$\dot{Z}_{ij} = F_Z(\mathbf{S}_i(t), \mathbf{S}_j(t), Z_{ij}) \quad (4.3)$$

$$\dot{J}_{ij} = F_J(Z_{ij}, \mathbf{y}(t), J_{ij}) \quad (4.4)$$

Un environnement continu classique (contrairement à un environnement discret) présente peu d'occasions d'interactions périodiques avec l'environnement. Dans un contexte d'environnement continu, la dynamique externe est lente par rapport à la dynamique neuronale. Dans ce cadre, les règles Hebbiennes d'ordre 0 (reposant sur la conjonction d'activité) se sont révélées plus efficaces pour la résolution des tâches, bien que moins propices aux effets de synchronisation et de résonance. Les résultats principaux sont les suivants :

- La règle d'apprentissage proposée dans (Daucé, 2004a) est strictement Hebbienne, mais distingue les différents faisceaux d'axones arrivant sur les neurones, autrement dit identifie les différentes sources de l'activité arrivant sur les neurones, conduisant à favoriser, selon les cas, les signaux externes, ou, au contraire, les signaux générés en interne (page 109 et annexe E).
- Le modèle proposé dans (Daucé, 2007) alterne règle Hebbienne et anti-Hebbienne, et ajoute un principe d'alternance entre l'activité perceptive et l'activité motrice (page 110 et annexe G).
- Le lien entre le policy gradient et la plasticité de Hebb est établi dans (Daucé, 2009) (page 110) sur un modèle de neurones à impulsions, et dans (Daucé & Dutech, 2010), (page 113 et annexe J) sur des neurones fréquentiels.
- Ces modèles systématisent de plus l'usage du codage topographique sur les couches sensorielles et proprioceptives, et mettent l'accent sur la formation de champs récepteurs.
- De nombreux modèles de neurones prennent en compte le caractère stochastique de la réponse des neurones biologiques (Tuckwell, 1988; Gerstner & Kistler, 2002). Dans ces modèles, le bruit d'exploration est un bruit de membrane “lent” ayant la même constante de temps caractéristique que l'erreur motrice.

Contrairement aux modèles de la perception fondés sur la résonance perceptive, les modèles mis en place dans ce cadre fonctionnent selon des principes d'appariement “classiques”, c'est-à-dire par décomposition de la tâche complète en sous-tâches élémentaires, en appariant des régions de l'espace sensoriel avec des réponses motrice spécifique, correspondant à des “cartes motrices” (voir figure 8.5, page 112).

#### **k. L'appariement par “accumulation d'évidence” implémente un mécanisme de révision des croyances dans un espace de choix concurrents.**

Les modèles probabilistes fondés sur des processus stochastiques Markoviens (voir les équations (3.12) et (3.13), page 39) sont très populaires en modélisation biologiquement

inspirée du fait de leur capacité à rendre compte des profils de réponse extrêmement bruités à la fois de l’activité neuronale et des réponses comportementales. Les processus Markoviens peuvent également implémenter, au même titre que les modèles déterministes, les phénomènes d’activité persistante, de rémanence, de mémoire à court terme, etc.

Nous considérons ici le cadre restreint où le modèle d’état  $P$  n’est pas connu. Nous considérons le cas où l’*observation*  $\mathbf{I}_{1:t} = (\mathbf{I}(1), \dots, \mathbf{I}(t))$  est le résultat de l’échantillonnage de ce modèle caché  $P$ . Le calcul de la distribution a posteriori  $\hat{f}$  repose sur l’observation  $Q_{\Theta}(I|X)$  uniquement (voir l’équation (3.13), page 39). L’évolution de  $\hat{f}$  au cours du temps décrit le processus de mise à jour et de révision des croyances sur la source des données.

Dans le cadre des modèles de l’appariement, l’évolution temporelle de  $\hat{f}$  modélise un mécanisme de décision par délibération où, comme dans le cas des réseaux de neurones récurrents, plusieurs “passages” sont nécessaires pour aboutir à un compromis. L’estimation faite à l’instant  $t$  sert de prior pour affiner l’estimation à l’instant  $t + 1$ .

- Le modèle présenté dans (Malhotra & Daucé, 2011) (voir page 118 et annexe K) reproduit des données psychométriques sur les temps de réponse à des “périodes préparatrices” de longueur variable (“variable foreperiod”). Plusieurs modèles génératifs  $P_1, \dots, P_k$  sont mis en concurrence. Le défilement du temps au cours d’un essai permet d’éliminer certaines hypothèses, réduisant l’incertitude et les temps de réponse au stimulus.
- Une application de ce mécanisme concerne les interfaces cerveau-machine de type “clavier virtuel P300” (Daucé & Thomas, 2014) (voir page 119 et annexe N). Le programme repose sur un modèle générique de la réponse EEG au stimulus cible (dite “onde P300”). Le mécanisme d’accumulation permet ici d’optimiser les temps de réponse du dispositif. Un seuil de décision définit l’instant où suffisamment d’évidence a été accumulée pour afficher la lettre courante et passer à la suivante.

De manière générale, ces modèles rendent compte du passage du temps dans l’élaboration d’une réponse, autrement dit modélisent un processus de délibération. De nombreux modèles de la perception utilisent ces principes pour rendre compte des temps de réponse à des stimuli plus ou moins ambigus (Gold & Shadlen, 2001; Simen et al., 2011; Khoei et al., 2013). Les modèles mis en place dans le cadre de l’orientation temporelle (Malhotra & Daucé, 2011) se placent sur des échelles de temps un peu plus grandes (plusieurs secondes), qui les rapprochent des modèles probabilistes du contrôle moteur (Wolpert & Ghahramani, 2000; Haruno et al., 2001) suggérant un lien possible entre l’estimation temporelle et la simulation motrice (Jeannerod, 2001).

Par ailleurs, ces algorithmes reposent sur des modèles paramétriques de la mesure, qui permettent la mise à jour du modèle après chaque série de mesures, selon le principe du maximum de vraisemblance, implémenté via une descente de gradient ou via l’algorithme EM (voir l’équation (3.14), page 41).

- Dans le cas du clavier virtuel P300, la mise à jour du modèle de la mesure permet, en partant d’un modèle d’utilisateur générique, d’adapter l’interface à l’utilisateur, ce qui améliore la réactivité de l’interface au fur et à mesure de son utilisation (Daucé & Thomas, 2014)
- Par ailleurs, de manière analogue à (Kivinen et al., 2004), l’utilisation d’un terme de régularisation dans la mise à jour réalise un modèle de mesure “glissant” qui permet

de s'adapter aux non-stationnarités de l'environnement (Daucé et al., 2013; Daucé et al., 2015).

### **l. En présence de signaux de renforcement binaires instantanés, l'apprentissage par exploration non-stochastique a les mêmes bornes d'erreur que l'apprentissage supervisé.**

L'analogie entre les algorithmes d'apprentissage par renforcement et la recherche du maximum de vraisemblance peut être poussée un peu plus loin dans le cas où le problème d'apprentissage consiste à appairer une entrée  $\mathbf{I}_t$  à une catégorie  $\mathbf{u}_t$  sur la base d'un signal de renforcement instantané binaire  $r_t$  de type "bien/pas bien". Une correspondance formelle peut alors être établie entre l'algorithme du policy gradient (sans trace) et l'algorithme du gradient logistique qui suit le maximum de vraisemblance sous hypothèse Gaussienne. Avec un choix judicieux des valeurs de récompense, la mise à jour du policy gradient *rejoint en espérance* le gradient logistique, ce qui permet d'appliquer une politique de *choix non stochastique* (politique du maximum a posteriori).

- Cette approche est mise en oeuvre dans le cadre d'un clavier virtuel P300 guidé par une touche "backspace" (Daucé et al., 2013; Daucé et al., 2015) (voir page 119 et annexe O).
- Nous montrons qu'avec une décision non-stochastique et en l'absence de modèle, le classifieur converge vers un taux de réponse comparable à l'état de l'art en classification BCI.
- Nous montrons également, dans le cas d'environnements non-stationnaires, que l'utilisation d'un terme de régularisation dans la règle de mise à jour définit un modèle "glissant" qui s'adapte aux changements de l'environnement.
- Dans le cadre voisin de l'optimisation quadratique (Vapnik, 1998; Crammer et al., 2006), nous montrons également, dans le cadre d'un guidage binaire (dit de type "bandit" - travail de thèse d'Hongliang Zhong), que la borne de l'erreur cumulée est identique à celle obtenue avec un guidage complet. Contrairement à l'algorithme du Banditron généralement utilisé dans ce cadre (Kakade et al., 2008), cette propriété conduit également à la possibilité de politiques d'exploration avec un bruit arbitrairement petit, autrement dit à une politique de choix non-stochastique.

# 5

## Bilan et projets

Ce chapitre présente enfin un certain nombre de questions laissées ouvertes par ces travaux, et présente un cadre conceptuel général pour des développements possibles en neurosciences computationnelles. Deux projets de recherche plus concrets sont également proposés, d’une part dans le cadre de la simulation large échelle des fonctions cérébrales, et d’autre part dans le cadre des interfaces cerveau-machine.

### 5.1 Bilan

#### 5.1.1 Résumé

Les résultats présentés dans la section précédente peuvent être regroupés sous un postulat commun : dans les architectures neuronales définies par un graphe de connexions récurrent, l’apprentissage et le contrôle s’implémentent selon le mécanisme de la *rétroaction positive*.

Le principe de l’appariement par rétroaction positive correspond à une recherche de correspondance entre un signal présenté en entrée et une des “formes” possibles que peut prendre l’activité du réseau (étude **a**). Le nombre de formes possibles (le “vocabulaire”) semble être maximal lorsque le système est en régime critique, c’est-à-dire proche d’un point de bifurcation dans l’espace des paramètres de contrôle (études **a** et **d**). L’activité est également plus riche (moins stéréotypée) lorsque le graphe de connexions est fortement hétérogène (étude **b**).

La réponse du programme n’est pas instantanée, mais s’inscrit dans une trame temporelle. Elle s’apparente à un mécanisme d’*intégration* (études **g**, **h** et **k**). L’implémentation neuronale peut reposer sur des mécanismes cellulaires seuls (étude **e**), ou sur une relaxation à l’échelle d’une population (études **a**, **c** et **g**)

Le mécanisme de l’appariement, au centre des algorithmes d’apprentissage, prend la forme d’une réduction de complexité de la dynamique intrinsèque (étude **f**). Deux grandes familles de plasticité sont abordées :

- Les plasticités d’“ordre 0”, guidées par un signal de renforcement, identifient les relations d’entrée-sortie les plus favorables, selon un principe qui s’apparente au gradient logistique (études **j** et **l**).

- Les plasticités d’“ordre 1” amplifient, au sein du réseau, des séquences d’activation récurrente, permettant de construire des réponses relatives à des événements temporellement séparés (études **g** et **h**).

Le couplage entre une activité persistante et un signal extérieur conduit de plus à des mécanismes de mémoire à court terme, se traduisant parfois sous la forme d’une “résistance au changement”, c’est-à-dire une persévération de l’activité interne malgré des changements au niveau des signaux sensoriels (études **c** et **g**). La “perception du changement” se traduit par des basculements d’activité, faisant évoluer la relation entre le système et l’environnement (études **g** et **i**). Enfin, nous avons montré que cette activité intrinsèque peut servir de support à l’innovation dans les rapports à l’environnement, c’est-à-dire principalement de proposer de nouvelles réponses comportementales, éventuellement stabilisées dans le cas où elles apportent une amélioration (étude **j**).

De manière un peu plus générale :

- Dans un réseau récurrent, le caractère hétérogène du graphe de connexions permet, en régime critique, de constituer un ensemble varié de réponses à des signaux extérieurs. Le hasard (l’aléa des connexions) constitue ici un principe structurant qui sépare l’environnement en régions distinctes, autrement dit qui “digitalise” l’environnement, selon un principe proche de l’“auto-encodage” tel qu’il est mis en œuvre dans les réseaux profonds (Hinton et al., 2006) (études **a**, **b**, **c** et **d**).
- Il semble possible d’établir un lien entre la plasticité d’ordre 1 (études **g** et **h**) et le principe d’“accumulation d’évidence” dans un environnement incertain (étude **k**). Sur des échelles de temps relativement courtes, la plasticité dépendante des instants de tir contribue à renforcer certaines séquences d’activation spécifiques, et donc à établir des liens causaux entre événements séparés. Ce mécanisme conduit également à une réponse de population via la synchronisation et la réduction de complexité (études **f**, **g** et **h**), qui pourrait correspondre à la formation d’un “canal de communication” transitoire entre l’échelle microscopique et l’échelle macroscopique, autorisant la mise en œuvre de programmes moteurs spécifiques.
- Nous avons également contribué, selon un principe de codage fréquentiel (non lié aux instants de tir), à établir une correspondance entre la règle de Hebb, dans sa variante “guidée”, et la recherche du maximum de vraisemblance dans un espace de couples (perception, action), où la combinaison de recherches d’optima locaux s’apparente à la recherche d’un optimum global en l’absence de modèle de l’environnement (études **j** et **l**).

### 5.1.2 Discussion

Le travail de compilation et de relecture/réexamen de résultats parfois anciens liés à la rédaction de ce rapport a été l’occasion de développer quelques rapprochements qui ne sont pas présents en tant que tels dans les publications.

Il m’a permis de mettre à jour ce qui semble être une difficulté fondamentale liée au fait de vouloir appliquer les principes de l’apprentissage automatique à l’acquisition de compétences motrices. Cette difficulté nécessiterait probablement un examen fouillé dépassant le cadre de ce rapport.

Comme nous l'avons vu, la plupart des algorithmes d'apprentissage fonctionnent selon un principe d'*appariement* ("pattern matching"), c'est-à-dire sur la construction, à partir des données, d'un ensemble de vecteurs "caractéristiques".

- Ces vecteurs caractéristiques, vus comme un dictionnaire de formes, constituent une base de description synthétique des données d'entrée.
- La projection des données d'entrée sur cet espace de redescription facilite ensuite leur traitement et leur interprétation.

L'apprentissage revient donc à la constitution d'un vocabulaire de base qui sert de clé d'interprétation des données d'entrée.

Il existe cependant peu d'exemples dans la littérature où ces familles d'algorithmes, dits d'auto-encodage, aboutissent à l'apprentissage d'une loi de contrôle dans un espace d'actions continu. Cette difficulté semble liée à la séparation conceptuelle et pratique entre, d'une part, l'approche systèmes dynamiques et, d'autre part, la reconnaissance de formes.

- La simplification, qui consiste à considérer un milieu intérieur (intensif) et un milieu extérieur (extensif) couplés comme un seul et même système dynamique, se heurte ainsi au caractère inhomogène de ce couplage.
- En particulier les distances, telles qu'elles sont définies dans l'espace Euclidien de la commande, ne sont pas conservées dans les espaces de redescription utilisés en reconnaissance de formes.
- Le caractère dissocié des ensembles de caractéristiques induits par la reconnaissance de formes conduit à poser le problème de leur "assemblage" pour produire des réponses adaptées aux dispositions des actionneurs et des capteurs dans l'espace physique.

## Disparités d'échelle

Dans domaine biologique, les échelles de temps et d'espace des comportements macroscopiques (mouvements, déplacements, activité motrice) se distinguent des échelles de temps et d'espace observées au niveau de l'activité nerveuse. On a d'un côté des échelles spatiales macroscopiques et des échelles temporelles lentes, et de l'autre des circuits microscopiques, constitués d'unités neuronales échangeant rapidement des signaux discrets. Les actions produites par le corps sont à une échelle temporelle lente en comparaison des potentiels d'action produits par les neurones. On parle de *Disparité d'échelle*.

Ce problème a été abordé de longue date dans les sciences cognitives. L'hypothèse du *passage d'échelle* suggère des moments transitoires où l'activité microscopique rejoint l'activité macroscopique lors des épisodes de synchronisation, conduisant à une réduction de complexité selon le principe déjà évoqué suggéré par la synergie (Haken, 1983). L'hypothèse est ici que ce couplage transitoire entre activité motrice et activité neuronale repose sur la réduction du nombre de degrés de liberté, au sein d'assemblées neuronales spécifiques, se traduisant une synchronisation des activités et des corrélations à grande échelle (Gray et al., 1989). Cette théorie n'a cependant pas encore, comme nous l'avons vu, abouti à des réalisations pratiques dans le domaine du contrôle moteur.

En plus de la disparité d'échelle, les lois régissant les deux systèmes ne sont pas de même nature. Le monde physique est assez bien décrit par l'approximation Newtonienne des masses en mouvement dans un milieu tri-dimensionnel. A l'inverse, l'activité logicielle s'inscrit dans un réseau où l'intensité des interactions est faiblement liée à la disposition spatiale, essentiellement produite par la connectique et les transports d'activité électrique.

Cette dichotomie implique un mode de fonctionnement composite, fondé d'une part sur des rapports entre quantités (monde physique), et d'autre part sur des rapports entre formes (monde digital). D'une certaine façon, la plupart des modèles en neurosciences computationnelles reposent sur un compromis, sur un "bricolage" entre les deux approches, ne prenant pas toujours en compte le caractère inhomogène de ces deux "mondes".

### Espace métrique et espace sensoriel

Dans le cadre des algorithmes d'apprentissage, l'universalité de l'apprentissage (sa capacité à apprendre n'importe quelle application) passe par l'utilisation de fonctions non-linéaires à une étape au moins du traitement, soit des fonctions seuil, sigmoïdes ou encore des fonctions à support compact (Haykin, 1999).

Cette transformation non-linéaire, imposée par la condition d'universalité (capacité à séparer des données non-linéairement séparables par exemple) a pour contrepartie de rendre difficile la définition d'une distance entre deux signaux. Les distances au sein de l'espace initial n'ont pas de traduction au sein de l'espace de redescription. L'approche par pattern matching ne garantit donc pas la proportionnalité des traitements entre signaux proches et signaux éloignés, comme le font les transformations linéaires traditionnelles. Autrement dit la distance définie au sein de l'espace de redescription (l'espace contenant les caractéristiques du signal) ne respecte pas nécessairement les distances ou la topologie de l'espace initial. On parlera de manière générale d'espace "qualitatif" pour décrire cet espace de redescription, ou encore d'espace "digital" pour décrire un tel ensemble de caractéristiques.

De manière similaire, le caractère distribué de l'activité nerveuse suggère la distribution d'une même composante sensorielle sur de nombreux neurones, chaque neurone étant sensible à une manifestation particulière de cette composante. L'exemple le plus extrême est celui de la rétine dont l'activité met en jeu plusieurs millions de neurones sensoriels, chaque neurone étant la manifestation d'une valeur de luminance (et de longueur d'onde) en un point précis de la scène visuelle.

### Problème de l'assemblage ("*Binding problem*")

Les modèles classiques de la commande motrice reposent implicitement sur l'hypothèse d'un espace physique extensif où prennent place des transformations linéaires<sup>1</sup>. Le

---

1. où les prédictions s'apparentent à des opérations isométriques, soit des translations et des rotations. Par exemple, les objets visibles sont translatés vers la gauche du champ visuel lorsque le regard tourne vers la droite. Nous avons donc un référentiel (celui de la direction du regard) qui est translaté, redéfinissant les coordonnées des différents objets qui composent la scène visuelle. Le référentiel visuel est lui-même défini par rapport au tronc (le support stable des mouvements de la tête), et l'axe de la direction du regard se définit par rapport à un autre référentiel qui est le référentiel du tronc. Les objets de la même scène sont également définis (en tant que point final des mouvements) dans ce même référentiel du tronc.

caractère distribué (“digital”) de l’activité neuronale rend difficile, comme nous l’avons vu, la réalisation d’opérations simples comme l’addition ou la différence de deux quantités, comme cela est pourtant nécessaire lorsque les déplacements doivent prendre en compte des sources sensorielles multiples issues de référentiels variés. Le problème se pose plus généralement lorsqu’il s’agit d’estimer, au sein de l’espace de redescription, des quantités *relatives* à différents points de l’espace pré-moteur ou de l’espace sensoriel (et non des quantités absolues). Si chaque modalité sensorielle de l’espace métrique environnant est décomposée en constituants élémentaires (intervalles ou “pixels” d’environnement), activant ou désactivant différentes populations de neurones primaires, alors les transformations aussi simples que les translations, rotations, voire même l’identité, etc. doivent être apprises “par cœur” (Pouget & Sejnowski, 1997) et ne peuvent résulter de la combinaison d’activités unitaires, c’est-à-dire d’un calcul.

L’apprentissage, par exemple, de la conséquence visuelle d’une commande d’orientation visuelle par “pattern matching” nécessiterait un apprentissage par cœur de la conséquence visuelle de toutes les combinaisons de dispositions d’objets et de tous les changements de référentiel tronç-tête, ce qui paraît computationnellement très coûteux ((Pouget & Sejnowski, 1997; Pouget et al., 2002)). Il est néanmoins nécessaire que le “lien” praxo-visuel entre les différents objets de la scène ne soit pas perdu au cours des déplacements moteurs (un raisonnement symétrique peut être appliqué pour le cas des déplacements du corps dans un environnement visuel stable, comme dans le cas du réflexe vestibulo-oculaire (Paillard & Beaubaton, 1978)). Ce problème du lien entre les différentes manifestations sensorielles d’un même objet s’appelle le problème de l’“assemblage” (“binding problem”). Il a été identifié de longue date dans les sciences cognitives sans qu’une solution définitive ne puisse lui être apportée (Von der Malsburg, 1986)<sup>2</sup>.

---

2. Il existe trois grandes catégories d’approches pour traiter de cette question : (1) l’activité tierce, (2) l’assemblage dynamique et (3) la mémoire (ou les transformations) extérieure(s).

**Activité tierce** La structure algorithmique la plus simple est celle de l’activité tierce (ou projection inverse), c’est à dire une zone d’activité (ou un neurone) instancie un objet complet, répondant de manière sélective à une combinaison spécifique de ses différents constituants. Cette approche est celle dite du “neurone grand-mère”. S’il apparaît peu probable qu’un seul neurone soit le support de la liaison, rien n’exclut qu’une population de neurones puisse répondre de façon sélective à toutes les configurations d’interaction avec un objet particulier (par exemple notre grand-mère) (Quiroga et al., 2005), voir aussi (Dehaene & Changeux, 2011). Une telle activité est donc la continuation à travers toutes ses transformations (déplacement du corps, mouvement des yeux) d’un même objet de l’environnement, autrement dit une “présence de l’objet”. Dans sa forme la plus élaborée, l’activité tierce est un modèle de l’environnement, constituée des principales quantités relatives aux objets d’intérêt, dans un espace extérocentré. Malheureusement, cette reconstruction de la structure externe de l’environnement par modèle inverse suppose implicitement des transformations géométriques au sein du substrat permettant de projeter l’objet invariant sur la scène prémotrice ou visuelle, la réalisation computationnelle de telles projections dans un substrat distribué n’étant pas garantie (à l’heure actuelle).

**Assemblage dynamique** L’approche “distribuée”, également appelée approche “dynamique”, stipule le caractère non exclusif des activités des population : il n’existe pas de neurone ou de population de neurones répondant exclusivement à un invariant de haut niveau. Le lien doit reposer sur des canaux de communication transitoires, ou encore des résonances, s’établissant à travers le substrat entre différentes activités constituants partiels de la scène. L’invariance est ici la susceptibilité à “résoner ensemble” (la persistance de la résonance à travers les changements ; l’existence de canaux le permettant) qui distingue ce qui est possible de ce qui ne l’est pas (les associations de constituants plausibles et celles qui ne le sont pas). L’hypothèse de l’assemblage dynamique est celle qui, à l’heure actuelle, semble la plus susceptible

L'approche de la cognition par correspondance de forme offre donc un support efficace à la plasticité, mais pose en revanche quelques problèmes lorsqu'il s'agit d'utiliser ce mode de codage pour contrôler un appareil moteur, c'est-à-dire schématiquement effectuer la transformation inverse du qualitatif vers le quantitatif.

## 5.2 Perspectives et projets

Le projet de recherche que je cherche à défendre vise à approfondir la compréhension de certains des problèmes calculatoires rencontrés dans le cadre de ces efforts de modélisation biologiquement inspirée.

Nous formons l'hypothèse que cette difficulté à faire apprendre des compétences motrices à un dispositif artificiel repose sur l'opposition entre deux grands principes fonctionnels : le feedback positif et le feedback négatif.

- Le mécanisme d'appariement, central dans l'apprentissage, repose majoritairement sur une logique de feedback positif (ou résonance)
- La théorie du contrôle, telle qu'elle est définie en sciences de l'ingénieur, repose majoritairement sur une logique de feedback négatif (correction de l'erreur motrice)

Si l'on souhaite construire des dispositifs artificiels capable d'apprendre des tâches motrices de la même manière qu'ils apprennent à classifier des signaux et des images, il faut opérer un renversement conceptuel en théorie du contrôle afin de rendre compatible le contrôle moteur avec l'apprentissage, en particulier *utiliser les principes du feedback positif et de l'appariement pour définir les tâches de contrôle moteur*.

Cette approche implique, comme dans le cas de l'analyse des signaux, une *décomposition de l'espace des tâches* en tâches élémentaires, autrement dit la capacité à constituer de très nombreux circuits de commande associés à des configurations précises de l'environnement.

Appliquée dans un cadre de contrôle classique, nécessitant des transformations isométriques, cette approche conduit, comme nous l'avons vu, à une explosion combinatoire des

---

d'implémenter le lien praxo-visuel, sans faire appel à une activité centrale spécifique ou à un modèle interne. Elle évite en particulier le coût combinatoire des changements de référentiels. Le mécanisme le plus populaire est celui qui a été proposé initialement par (König & Schillen, 1991) (voir également (Rodriguez et al., 1999; Wang, 2010)) reposant sur les résonances entre oscillateurs. Il reste néanmoins de nombreuses questions relatives à l'implémentaton concrète (ou opérationnelle) de ces mécanismes, ainsi que sur leur caractère (ou non) prédictif.

**Mémoire extérieure** Une troisième approche, celle de la “mémoire extérieure” (O'Regan & Noë, 2001) reprend la tradition de l'approche écologique impulsée par Gibson (Gibson, 1979). L'existence de motifs invariants dans la relation de l'agent à son environnement (comme le point de fuite) sert de support à la manipulation de cet environnement. Le lien instrumental avec les objets (le mode d'utilisation des objets) institue une forme d'invariance qui est reproduite à travers le temps à chaque rencontre avec le même type d'objets. L'approche Gibsonienne met donc au premier plan l'activité motrice en tant que soubassement de la constitution d'invariants dans le rapport au monde. D'un point de vue algorithmique, on peut imaginer par exemple des activités “template” correspondant à des successions caractéristiques d'actions (actions stéréotypées). Dans le cadre de l'exploration visuelle, on peut imaginer des séquences caractéristiques de saccades selon le type de scène visuelle à explorer, ce “plan de visite” constituant le socle d'interprétation des éléments de la scène. La mise en relation entre les éléments se réduit ici à un schéma d'exploration (ou d'utilisation) caractéristique de ces différents éléments, constitutif d'une unité d'action.

combinaisons sensori-motrices à apprendre. Pour éviter ce problème, il faut donc établir des règles de contrôle *ne nécessitant pas* de comparaison entre des grandeurs de l'espace environnant.

### 5.2.1 Construction d'un espace métrique péripersonnel

“le monde visuel est constitué par l'action.”

Varela, FJ

Un premier projet que je souhaiterais développer dans ce cadre est d'identifier les principes de construction de l'espace métrique tridimensionnel courant tel que nous l'expérimentons dans nos interactions quotidiennes avec l'environnement.

Si nous regardons l'activité sensori-motrice telle qu'elle se déploie à l'échelle macroscopique, on constate que l'activité humaine est fortement guidée par la composante visuelle. Si on regarde de plus près l'interrelation entre l'exploration visuelle et l'action simple consistant à manipuler des objets situés à portée de bras, deux grands domaines attentionnels peuvent être définis :

- d'une part le domaine visuel, organisé autour de la direction du regard ;
- d'autre part, l'espace de travail, organisé autour des opérations sur des objets attrapables ou manipulables.

Il est intéressant de constater que ces deux types d'activité (exploration visuelle, utilisation) sont rarement simultanées (Paillard & Beaubaton, 1978). De manière un peu schématique,

- l'espace de travail reste invariant pendant l'exploration visuelle,
- et, réciproquement, le regard reste statique (focalisé) durant l'activité de manipulation.

En particulier, le déplacement du corps dans l'environnement se fait généralement dans un environnement visuel stable grâce au réflexe vestibulo-oculaire qui maintient le regard en direction d'un même objet pendant le déplacement (“ancrage” visuel).

Cette stabilité sélective d'une partie des composantes de l'environnement apparaît comme un élément important de l'apprentissage du référentiel extérieur. C'est également un bon modèle pour questionner les opérations nécessaires à la projection de l'espace pré-moteur vers l'espace visuel et vice-versa. Si on fait appel aux notions Piagétienne de permanence de l'objet (Piaget, 1973), les différents objets de la scène varient dans l'espace pré-moteur et restent statiques dans l'espace visuel lorsque le corps bouge, et inversement varient dans l'espace visuel et restent statiques dans l'espace pré-moteur lorsque le regard bouge.

- Etant supposée par exemple la permanence des objets dans l'espace pré-moteur, il est possible de prédire leur position dans le champ visuel étant donnée la direction du regard.
- Inversement, étant donnée la permanence des objets dans l'espace visuel, il est possible de prédire leur position dans l'espace pré-moteur étant donné le déplacement du corps (il est difficile de bouger à la fois le corps et le regard).

Il existe donc une application qui à partir de la position des objets dans un premier référentiel définit la position des mêmes objets dans un autre référentiel. La bijection entre les deux espaces offre un bon support à l'invariance ainsi qu'à la prédiction sensorielle.

Le mécanisme de construction de l'espace métrique environnant reste à l'heure actuelle non-élucidé. Nous avons vu que le caractère "digital" de l'activité neuronale rend difficile, en interne, la construction d'opérateurs sur cet espace. Ce mécanisme semble pourtant pouvoir être implémenté sur la base de principes d'appariement très simples, mais n'a, à ma connaissance, pas encore été proposé en modélisation. Deux principes semblent s'imposer dans ce cadre :

1. la commande par champ de vecteurs,
2. le principe de l'alternance visuo-motrice.

**Commande par champ de vecteurs** Des stimulations directes du cortex moteur, sur des échelles de temps comparables à la durée d'un déplacement (500 ms), suscitent des mouvements complexes convergeant toujours vers la même posture finale, quelle que soit la position de départ (Graziano et al., 2002). Une composante de direction peut également être mise en évidence, aboutissant à des possibilités de reconstruction de la trajectoire métrique à partir de l'activité de différents neurones (Georgopoulos et al., 1986; Wessberg et al., 2000).

La commande balistique par combinaison de champ de vecteurs, telle qu'elle a été proposée par (Mussa-Ivaldi & Solla, 2004), permet de rendre compte de cette organisation spatiale de la commande motrice. Selon cette approche :

- les commandes motrices balistiques correspondent à la définition d'un attracteur dans l'espace de la tâche. Cet attracteur est produit comme combinaison de différents champs de vecteurs (appelés "primitives motrices"). Ce type de commande ne nécessite pas de référence à un point d'opération externe (voir aussi page 43). La dynamique de relaxation et les constantes de temps propres à l'effecteur contrôlé conduisent les différentes articulations vers leur point d'équilibre.
- L'ensemble des points que le bras peut atteindre forme une région de l'espace tridimensionnel. Le cortex moteur reproduit cette organisation spatiale : le cortex est organisé sous la forme d'une carte motrice bidimensionnelle où chaque région de la carte correspond à une combinaison différente de primitives motrices conduisant la main vers un point de l'espace tridimensionnel environnant (péripersonnel).

Ainsi :

- L'espace péripersonnel se définit comme l'ensemble des points fixes que la main peut atteindre.
- Seule compte la position finale. Il y a donc une commande unique quel que soit le point de départ.

**Alternance visuo-motrice et permanence de l'objet** Une idée que je souhaiterais développer dans le cadre de ce projet est que l'ancrage visuel, tel qu'il est observé dans le cadre des déplacements moteurs et de l'activité de manipulation, a pour contrepartie symétrique un ancrage postural qui peut servir de support au principe de *permanence de l'objet*.

L'hypothèse que je souhaite développer est que ces deux principes d'ancrage permettent de constituer un espace tridimensionnel péripersonnel, par appariement prédictif, sans faire appel à des transformations complexes entre référentiels (voir note de bas de page 1).

Le principe d'alternance entre un ancrage pré-moteur et un ancrage visuel suppose l'existence de deux systèmes de commande non-simultanés correspondant, pour le premier, à l'orientation du regard (commande d'orientation), et pour le second à la manipulation de l'environnement proche ("utilisation" de l'environnement). Ces deux systèmes de commandes définissent deux espaces : l'espace de la direction du regard  $\mathcal{U}_v$  (constitué de l'ensemble des points focaux où le regard se pose), avec  $\mathbf{u}_v$  la commande d'orientation, et l'espace péripersonnel  $\mathcal{U}_p$  (espace atteignable/manipulable par la main), avec  $\mathbf{u}_p$  une posture de la main. A ces deux espaces s'ajoute un troisième, qui est celui de la *scène visuelle*  $\mathcal{I}$ , correspondant à l'ensemble des points de la rétine.

Dans ce cadre, l'apprentissage de l'espace (le fait de pouvoir situer dans l'espace péripersonnel l'ensemble des points de la scène visuelle) peut s'organiser autour de deux mécanismes d'appariement :

- Appariement par ancrage visuel. L'ancrage visuel signifie que la direction du regard  $\mathbf{u}_v$  est fixe. L'activité de manipulation permet, dans le contexte défini par  $\mathbf{u}_v$ , d'apparier différentes scènes visuelles  $\mathbf{I}_1, \mathbf{I}_2, \dots$  avec différentes positions de la main  $\mathbf{u}_{p,1}, \mathbf{u}_{p,2}, \dots$
- Appariement par ancrage postural. Un point de l'espace péripersonnel s'apparente ici à un "objet", c'est-à-dire un élément de l'environnement atteignable par la main. L'ancrage postural correspond ici à la position de la main sur (ou à la préparation d'un mouvement de la main vers) ce point de l'espace. Si on suppose un ancrage postural fixe  $\mathbf{u}_p$ , l'activité d'orientation (changement de point focal du regard) permet, dans le contexte défini par  $\mathbf{u}_p$ , d'apparier différentes scènes visuelles  $\mathbf{I}_1, \mathbf{I}_2, \dots$  avec différentes directions du regard  $\mathbf{u}_{v,1}, \mathbf{u}_{v,2}, \dots$ , ce qui permet de constituer la *permanence de l'objet* à travers différentes scènes visuelles.

Il reste à déterminer dans ce cadre (1) l'appariement à mettre en œuvre (par produit scalaire, probabiliste ou par relaxation vers un attracteur), (2) de préciser le mécanisme spécifiant l'alternance entre les deux principes d'ancrage, et (3) le principe qui guide le choix et l'actualisation de la commande motrice.

Dans ce cadre,

- Le simple fait de produire une commande arbitraire (aléatoire) permet de tester la conséquence sensorielle de la commande de manière *non-supervisée*. L'apprentissage des contingences sensori-motrices a récemment fait l'objet d'un projet européen (eSMC project - extended Sensori-Motor Contingencies). Ce type d'apprentissage est par ailleurs évoqué dans le cadre applicatif de la robotique développementale (Andry et al., 2001; Revel & Andry, 2009).
- Dans un contexte donné, l'appariement consiste simplement à faire des associations point à point, sous forme de bijection. Si  $N$  est le nombre de points à apparier entre deux scènes, il y a  $O(N)$  appariements à apprendre par contexte.
- La prédiction des conséquences sensorielles de l'action, dans un contexte donné, découle simplement des appariements. Une préparation motrice prédit, *en contexte*, sa conséquence sensorielle attendue.

- Le principe d’ancrage nécessite qu’une activité persistante soit maintenue au cours des changements respectifs de posture ou d’orientation.
- Chaque contexte définit (spécifie) un ensemble d’appariements distinct. Si ce principe est facile à définir mathématiquement (fonction indexée par le contexte, vraisemblance conditionnée au contexte, etc.), le mécanisme neuronal grâce auquel il peut être implémenté reste à découvrir.

### 5.2.2 Vers une approche computationnelle de l’activité à large échelle

L’environnement Marseillais offre des possibilités de collaborations nombreuses dans le domaine de la modélisation biologiquement inspirée. Nous avons sur le campus de la Timone (faculté de Médecine) deux laboratoires de neurosciences directement intéressés par les problématiques de modélisation : l’INT (Institut de Neurosciences de la Timone), laboratoire CNRS en pointe sur la modélisation du système visuel, du système moteur ainsi que sur l’analyse des données d’imagerie optique, et l’INS (Institut de Neurosciences des Systèmes), laboratoire INSERM en pointe sur l’analyse et la modélisation de l’épilepsie ainsi que sur la modélisation dite à large échelle de l’activité nerveuse, à travers le projet TVB (The Virtual Brain) porté par Viktor Jirsa et Randy McIntosh (Université de Toronto). L’INS héberge également un centre de calcul principalement dédié à la mise en oeuvre des calculs coûteux liés à la simulation détaillée des processus de diffusion à la surface d’une reconstruction 3D du cortex, dans le cadre de l’interface de simulation du projet TVB.

Cet environnement donne accès à des données d’observation sur l’animal, mais également chez l’humain dans le cadre des chirurgies pré-opératoires consistant à implémenter des électrodes stéréotaxiques profondes chez les patients épileptiques pour identifier les régions épileptogènes à retirer. Nous avons donc accès à des données de première main concernant l’activité neuronale telle qu’elle se développe chez le sujet épileptique. Les signaux obtenus dans le cadre de l’analyse pré-opératoire correspondent à une activité “large échelle”, se développant sur des échelles de temps et d’espace macroscopiques. L’objectif clinique (opérer le patient sans produire de lésions incapacitantes) nécessite une compréhension fine des grands mécanismes de réseau à l’oeuvre dans le cerveau. L’un des objectifs de l’ANR “Vibrations”, financée à partir de 2014, et à laquelle je participe, est de produire des modèles large-échelle personnalisés pour chaque patient en observation. Il s’agit en particulier d’améliorer le diagnostic (identifier la zone épileptogène) ainsi que prédire les conséquences de la chirurgie, pour minimiser les séquelles post-opératoires.

Mes apports dans ce type de projet, et plus généralement dans le cadre de la modélisation large échelle, concernent principalement l’analyse et la modélisation des mécanismes mettant en jeu une activité intrinsèque et des interactions entre grandes populations de neurones, dans une des perspectives fonctionnelles évoquées précédemment. Une idée que je cherche en particulier à développer est celle de l’identité (ou de la simultanéité) de l’activité intrinsèque avec certaines grandeurs caractéristiques de l’environnement.

L’étude que nous avons conduite sur l’orientation visuelle (étude e) met en évidence une caractéristique importante de la commande motrice : la simultanéité entre l’activité des structures motrices ou prémotrices et le déplacement moteur correspondant (co-

activité sur un intervalle de l'ordre de 100 ms dans le cas de la saccade). Les données de la littérature indiquent que l'activité électrique motrice est le "reflet" d'une activité mécanique prenant place dans l'environnement physique externe. Des études portant sur le cortex moteur montrent également des corrélations entre l'activité mesurée sur les neurones du cortex moteur et le mouvement effectif, confirmant leur coextensivité temporelle (l'activité s'arrête lorsque le mouvement s'arrête).

Le thème de la commande motrice nous offre ainsi un bon point d'appui pour caractériser certains aspects de l'activité se développant *au cours du temps* dans le cerveau. Ce thème peut avoir d'intéressants développements, à la fois théoriques et pratiques.

D'un point de vue théorique, l'activité motrice est la forme la plus élémentaire de prédiction relative à des grandeurs de l'environnement physique, plus précisément l'activité motrice prédit que l'appareil moteur est en train de bouger. Cette correspondance terme à terme entre une activité électrique et un déplacement moteur n'est pas une nécessité fonctionnelle, et traduit une relation de fort couplage dont la raison d'être reste à établir. Une idée possible est que cette activité sert de point de départ pour établir des prédictions plus complexes sur l'espace métrique environnant, selon une modalité qui reste à déterminer. Une autre idée est que cette activité motrice définit l'échelle temporelle caractéristique sur laquelle se déroulent les opérations réalisées par le cerveau (de 100 à 500 ms, voir aussi (Petitot et al., 2002)).

Une étude réalisable serait d'identifier chez des patients implantés les activités électriques se développant de manière simultanée avec des déplacements moteurs mesurables par des dispositifs externes (comme une caméra ou des dispositifs accélérométriques). Les outils (capteurs, caméras) existent à l'INS, de même que les outils de visualisation permettant d'identifier les régions anatomiques sur lesquelles l'activité se déploie. Un projet en cours à l'INS se propose d'ailleurs de caractériser certaines crises du lobe frontal, se manifestant par des activités motrices complexes, à l'aide de tels capteurs accélérométriques dans le but de mieux quantifier et classifier ces crises.

Modéliser une telle activité électrique implique par ailleurs des mécanismes computationnels à implémenter, par exemple sous la forme de programmes moteurs, tels qu'observés dans le cadre de la saccade oculaire. De tels programmes mettent probablement en jeu les circuits sous-corticaux (ganglions de la base, thalamus et peut-être cervelet) qui ne sont pas encore implémentés sur l'outil de simulation TVB.

En plus de ces mécanismes computationnels, l'analyse de l'activité spontanée du cerveau à l'état de repos, comme nous l'avons déjà évoqué, montre des alternances marquées entre différents patrons d'activité distribués, mettant en jeu des régions similaires d'un sujet à l'autre. Selon l'outil de mesure utilisé, ces alternances peuvent prendre place sur des échelles de quelques centaines de millisecondes (Van de Ville et al., 2010) (analyse EEG) ou quelques dizaines de secondes (Fox et al., 2005) (analyse fMRI). L'alternance rapide de l'activité semble être le support de traitements cognitifs élémentaires. Le caractère rapide de ces alternances laisse supposer un mécanisme actif d'assemblage et de désassemblage des assemblées neuronales successivement actives, comme proposé dans (Rodriguez et al., 1999). Une interprétation possible est celle d'une alternance active, au sein du cerveau, entre une activité principalement perceptive, préparatrice de l'action, et une activité principalement motrice, préparatrice de la perception. Cette alternance rapide entre préparation motrice et préparation perceptive est un cadre possible à la réa-

lisation du *predictive coding*, comme stipulé par (Rao & Ballard, 1999) (voir aussi (Wurtz, 2008; Rolfs et al., 2011)). Les alternances à plus longue constante temporelle pourraient quant à elles reposer sur l'alternance déjà évoquée entre les activités d'exploration sensorielle au sens large et les activités liées à la manipulation et à l'utilisation des objets de l'environnement (sous la forme éventuelle de la simulation motrice), qui peuvent s'apparenter, dans une certaine mesure, aux processus internes et externes stipulés dans (Fox et al., 2005). L'analyse des patrons d'activité développés à large échelle nécessite plus généralement une analyse fonctionnelle et non simplement structurelle. Il s'agirait ici de sortir des approches quantitatives pour poser les bases d'une interprétation plus computationnelle des activités large-échelle observées, ouvrant la voie à des modèles informatiques plus précis.

### 5.2.3 Interfaces informatiques inspirées par la biologie

Nous avons également à Marseille une forte tradition en informatique théorique, portée historiquement par le LIF (Laboratoire d'Informatique Fondamentale de Marseille), avec une expertise poussée dans le domaine des langages fonctionnels, mais qui développe également une recherche de pointe en apprentissage automatique (Machine Learning). L'I2M (Institut de Mathématiques de Marseille) est également impliqué sur des thématiques proches du machine learning, en particulier dans le cadre du développement de modèles probabilistes en traitement du signal. Le centre INRIA de Sophia-Antipolis est également impliqué dans des collaborations régulières avec les laboratoires de neurosciences Marseillais, sur des thématiques en lien avec le traitement du signal et la construction de modèles inverses en imagerie.

Cet environnement facilite le développement de projets mixtes, à l'interface du traitement du signal et de la modélisation biologique dans le cadre de l'étude des processus décisionnels classiques.

Les interfaces cerveau-machine, déjà abordées, offrent un cadre applicatif intéressant et potentiellement utile principalement pour les déficiences motrices et la réhabilitation motrice. L'activité décelée dans le cadre des interfaces non-invasives (de type EEG) ne permet pas de reconstruction précise des activités internes à l'origine du signal de surface. La plupart des interfaces proposées se basent sur une démarche de classification automatique, ne permettant pas une interprétation de l'évolution spatio-temporelle de l'EEG.

L'idée générale d'un projet applicatif de ce type serait de construire une interface dont les principes de fonctionnement devraient être aussi proches que possible, étant données les connaissances actuelles, du fonctionnement du système nerveux. Il s'agit, d'un côté, de bien comprendre l'activité large-échelle pour mieux interpréter les données issues des capteurs non invasifs (EEG, voire MEG). De l'autre, il s'agit de s'inspirer des modes de fonctionnement du cerveau pour élaborer l'interface qui traite le signal. L'interface est en effet conceptuellement dans une situation similaire à celle d'un sujet humain : elle fait face à un signal peu fiable, incomplet, bruité et doit apprendre à interpréter parfois compléter des données manquantes, pour produire des réponses (possiblement des déplacements moteurs dans le cas d'une interface de contrôle). Une avancée importante dans le domaine des interfaces cerveau-machine pourrait donc être la mise en place dans

un cadre conceptuel bien contrôlé d'une véritable boucle de rétroaction entre l'interface et le sujet, en traitant le signal en temps réel sur un principe d'anticipation et de prédiction pour compenser autant que possible les délais de traitement.

L'introduction de cette démarche "en boucle fermée", utilisant éventuellement des connaissances acquises lors de mesures invasives, devrait en principe permettre de traiter, sous une forme peut-être élémentaire, ces changements d'état rapide, et améliorer ainsi l'interprétation du signal et donc la fiabilité de l'interface.

Une partie centrale du projet devrait être la prise en compte des mécanismes adaptatifs, en particulier dans le cadre d'une interface impliquant des stimuli visuels, pour "caler" autant que possible les instants d'apparition des stimuli avec les rythmes et les fréquences intrinsèques au cerveau, en respectant en particulier, s'ils peuvent être identifiés, les alternances entre les moments de préparation motrice et les moments de préparation perceptive. D'autre part, le modèle devrait prendre en compte autant que faire se peut les caractéristiques à la fois structurelles et dynamiques du cerveau du sujet, en identifiant la localisation des circuits large échelle ainsi que les rythmes ou résonances propres de son activité cérébrale.

En dernier lieu, il faut envisager des mécanismes adaptatifs agissant en direct pendant l'interaction avec le sujet. On parle dans ce cas d'algorithmes d'apprentissage "en ligne". Il s'agit d'une part d'en apprendre chaque jour un peu plus sur le sujet au travers des interactions répétées, pour affiner le modèle utilisé par l'interface. D'autre part, il s'agit d'adapter "en direct" le modèle en fonction d'événements imprévus (comme des électrodes qui se déplacent ou se décollent) ou de variations lentes de l'état du sujet (suite à la fatigue, à des baisses d'attention etc.). Dans ce cadre, le développement d'algorithmes adaptatifs spécifiques serait à envisager.

Plus généralement, les connaissances en plein essor sur le fonctionnement du cerveau à large échelle nous permettent d'envisager de nouvelles méthodologies dans le développement d'objets interagissant de manière "naturelle" avec leur utilisateur. C'est à travers une connaissance plus approfondie de la logique et des mécanismes computationnels à l'œuvre dans le cerveau (en particulier à large échelle) qu'on pourra peut-être repenser la manière de concevoir des objets communicants (robots mobiles mais également interfaces vocales, jeux, etc.) à travers des modes de communication plus proches de ceux qu'un sujet expérimente quotidiennement avec son entourage.



Deuxième partie  
Résultats détaillés



# 6

## Dynamique des grands réseaux de neurones aléatoires

Les résultats présentés dans ce chapitre s'appuient en grande partie sur les comportements de systèmes dynamiques de grande dimension, c'est dire constitués par l'interaction de nombreuses unités couplées :

- Le but est d'exploiter des comportements d'ensemble simples qui peuvent apparaître dans les systèmes de grande dimension, selon le principe de *réduction* proposé dans le cadre de la physique des systèmes hors d'équilibre (Haken, 1983).
- Les comportements d'ensemble sont alors interprétés comme des *implémentations* d'opérateurs plus simples.

### 6.1 Réseaux de neurones récurrents et modélisation biologique

Les réseaux de neurones récurrents sont définis par le fait que leur graphe contient des cycles. Ces cycles conduisent à des dépendances réciproques entre les unités du réseau. Le formalisme des systèmes dynamiques permet de décrire ces interactions croisées et leur effet sur les patrons d'activité produits par le réseau.

#### Activité induite, modes et capacité

L'activité du réseau au temps  $t$  est le résultat de l'intégration des potentiels d'action émis aux instants précédant  $t$ , en tenant compte des délais de transmission, et de la valeur courante du signal externe (voir l'équation (2.5), page 20, et l'équation (2.7), page 22).

Dans le cas non autonome, la trajectoire dépend à la fois des conditions initiales et du signal extérieur.

- On parle dans ce cas d'une activité *induite* par le signal d'entrée.
- Du point de vue computationnel, cela signifie que le même signal peut recevoir un traitement différent si les conditions initiales sont différentes. Sans précision des conditions initiales, de multiples opérations sont donc possibles pour un même signal d'entrée. On parlera des différents *modes* de l'activité du réseau.

*Exemples :*

- Un exemple minimaliste est un réseau de neurones récurrent à deux neurones mutuellement connectés. Un tel réseau pourrait présenter deux modes : un mode “inactif” dans lequel les deux neurones sont dans l’état “bas”, et un mode “actif” dans lequel une activité “haute” locale est transmise indéfiniment d’un neurone à l’autre. Chacun de ces deux patrons d’activité est une solution du graphe de contraintes.
- L’exemple des réseaux de Hopfield (Hopfield, 1982) illustre bien la notion de modes multiples. L’expressivité du réseau est vue sous l’angle du nombre de patrons distribués (patterns) pouvant être obtenus en tant qu’état final (attracteur) atteint par le système dynamique décrit par les poids (arêtes) du réseau. La capacité s’exprime ici en nombre de modes (configurations atteignables) et non en nombre de nœuds/vecteurs supports etc.

### Réseaux équilibrés

L’activité nerveuse est caractérisée par la présence fréquente d’oscillations collectives, souvent transitoires, observables à différentes échelles (Wang, 2010).

Les comportements oscillants ont été analysés de manière approfondie en modélisation à travers l’étude des réseaux de neurones dits “équilibrés” (*balanced neural networks*). Il s’agit d’une classe de réseaux de neurones constitués d’une population de neurones inhibiteurs et d’une population de neurones excitateurs. Le terme “équilibre” se réfère ici à l’idée d’équilibre dynamique entre deux forces opposantes : la dynamique du réseau est amenée à converger vers un point d’équilibre traduisant localement (et globalement) le “poids” des deux influences.

- Une configuration des poids se traduisant par une activité endogène soutenue traduit une domination de l’influence excitatrice.
- A l’inverse, une configuration où l’activité intrinsèque disparaît après un certain temps traduit une domination de l’influence inhibitrice.

Comme dans le cas des réseaux aléatoires simples, il existe un paramètre de contrôle correspondant au rapport entre l’influence excitatrice et l’influence inhibitrice.

De manière intéressante, une grande variété de comportements dynamiques peuvent être observés à la frontière, et entre l’extinction et l’activité soutenue, comme les oscillations synchronisées (Brunel, 2000a) et/ou le chaos (Van Vreeswijk & Sompolinsky, 1998).

**Rôle computationnel des oscillations synchronisées** Plusieurs hypothèses ont été proposées sur le rôle des oscillations synchronisées,

- comme mécanisme permettant de “lier” (assembler) les différentes modalités (constituants) d’un même phénomène (Gray et al., 1989). La cohérence de phase observée à large échelle est vue comme la signature d’une activité coordonnée entre des aires séparées par plusieurs centimètres (Rodriguez et al., 1999).
- comme porteuse permettant une forme de cadencage et de séquençage des stimuli dans les couches sensorielles primaires (MacLeod & Laurent, 1996).
- Différents régimes oscillatoires sont classiquement identifiés, classifiés selon la bande de fréquence qu’ils occupent : oscillations theta (3-8 Hz), oscillations

alpha (10-15 Hz), oscillations beta (15-25 Hz) et oscillations gamma ( $> 30$  Hz) (Bremer, 1949).

- L’anticipation de phase (phase précession) ou le retard de phase observés au niveau cellulaire dans un contexte d’oscillations theta (Skaggs et al., 1996) sont interprétés comme caractéristiques d’une activité au sein de laquelle l’ordre de tir des potentiels d’action individuels joue un rôle important.

### Codage topographique et champ neuronal

Le codage topographique est la manière la plus naturelle de représenter les signaux sensoriels dans les architectures neuronales. Il correspond également aux observations sur l’organisation des couches sensorielles primaires (à l’exception de l’olfaction).

Le codage topographique est une des manière d’exprimer des grandeurs continues dans un substrat neuronal. Il s’apparente au calcul par champ (“Field computation” (MacLennan, 1999)) consistant à étudier des mécanismes de calcul à partir de modèles d’interaction reposant sur la physique des milieux continus. Le débat n’est pas tranché à l’heure actuelle pour savoir si l’activité neuronale, par nature discrète, traduit les grandeur continues de manière *qualitative* (différents neurones étant assignés à différentes valeurs - ou intervalles) ou *quantitative* (chaque neurone pouvant exprimer plusieurs valeurs par sa fréquence de décharge). Le système visuel primaire, organisé autour de la topographie de la rétine, semble par exemple obéir à une organisation de type qualitatif où les différentes coordonnées du champ visuel sont assignées à des populations de neurones distinctes (Hubel & Wiesel, 1962). L’étude des zones motrices primaires plaide plutôt pour un caractère quantitatif de l’activité motrice, la contraction des muscles étant proportionnelle au taux de décharge.

Le modèle fondateur du codage topographique est le champ neuronal (“Neural Field”) d’Amari (Amari, 1977), déjà mentionné page 75. L’avantage du codage topographique est sa plus grande expressivité. Un foyer d’activité unique sur un substrat bidimensionnel exprime ainsi une grandeur vectorielle (typiquement la position d’une cible sur une rétine). Les patrons d’activité plus complexes observés dans les couches primaires (Bonhoeffer et al., 1991) caractérisent des relations spatiales entre des contours difficiles à exprimer à l’aide de simples neurones fréquentiels. Les caractéristiques du codage topographique suggèrent par ailleurs des mécanismes de contrôle de type “point fixe” où la commande motrice est définie comme le point d’équilibre entre différentes postures dans l’espace des tâches (Schöner et al., 1995; Mussa-Ivaldi & Solla, 2004; Flash & Hochner, 2005). Cette théorie permet d’expliquer comment les variations d’activité, en agissant linéairement sur les différentes forces, ne déplacent pas le point d’équilibre.

Le codage topographique suggère également un apprentissage “par coeur” des relations métriques entre les différents référentiels sensoriels (extéroceptifs ou proprioceptifs) (Pouget & Sejnowski, 1997). Dans ce schéma, la conséquence proprioceptive d’une commande motrice visuelle (excentricité finale de l’œil dans le globe oculaire suite à une saccade oculaire par exemple) doit être appris par essai/erreur.

Le codage topographique joue par ailleurs un rôle qui s’apparente à celui des fonctions noyau utilisées en apprentissage automatique (Schölkopf & Smola, 2002). Le plonement des données d’entrée dans un espace de fonctions à noyau reproduisant garantit

en effet l'existence d'un hyperplan séparateur permettant la discrimination des données. Cette meilleure séparabilité des données est obtenue au prix d'une augmentation de la complexité du classifieur (dimension de Vapnik-Chervonenkis). Dans le cas des architectures neuronales, le codage topographique repose sur l'activité de larges populations de neurones, et représente donc un coût plus élevé que le codage fréquentiel, en principe implémentable sur un neurone isolé.

### Neurones impulsionnels

**Intégration sous-liminaire** Les neurones biologiques ont un comportement fortement non-linéaire, caractérisé par le franchissement d'un seuil d'activation. La cellule neuronale intègre les multiples potentiels post-synaptiques issus de ses synapses afférentes et émet un bref potentiel d'action lorsque le potentiel total atteint le seuil d'activation (on dit que la membrane se dépolarise). Après une brève période de repos ("remise à zéro"), le neurone recommence à intégrer les signaux entrants etc.

**Action et contre-action** Les neurones biologiques sont également caractérisés par le type de neurotransmetteurs qu'ils émettent au niveau des synapses de leur arborisation terminale. On distingue les neurones inhibiteurs produisant des synapses GABAergiques et les neurones excitateurs produisant des synapses excitatrices variées (glutamate, acétylcholine, dopamine, ...). L'action synaptique inhibitrice est une "contre-action" qui vient s'opposer à l'effet des synapses excitatrices.

L'action et la contre-action, c'est-à-dire l'équilibre dynamique entre deux forces opposées, est au coeur des opérations réalisées par le système nerveux.

**Rôle computationnel** Les neurones à potentiel d'action disposent en principe d'une expressivité supérieure à celle des neurones à fréquence de décharge (Izhikevich, 2006). Ils offrent des possibilités d'opérations (tels le codage par rang, la polychronisation,...) ainsi qu'un support pour un traitement parallèle distribué ("multiplexage") par décalage de phase ou par bande de fréquences (Skaggs et al., 1996; Wang & Buzsáki, 1996; Wang, 2010). Un des enjeux connus est la possibilité d'émuler un mécanisme de synchronisation/coordination longue distance supposé à l'oeuvre dans le cerveau pour l'intégration multimodale (Rodriguez et al., 1999).

## 6.2 Contributions personnelles

### 6.2.1 Réseaux aléatoires multi-populations

Une des premières études auxquelles j'ai participé portait sur les transitions de phase dans les réseaux de neurones aléatoires. Les différentes architectures neuronales étudiées ont permis d'identifier des comportements caractéristiques, décrits sous la forme de différents régimes dynamiques.

L'étude présentée dans (Daucé et al., 1999; Daucé et al., 2001) (voir annexe B) étudie le comportement d'un réseau dit "équilibré" (voir section 6.1). Cet article étend pour la première fois le modèle de champ moyen pour le cas des réseaux à connectivité aléatoire

à populations multiples (auquel cas les tirages aléatoires sont effectués sur des “faisceaux d’axones” reliant une population de neurones  $p \in 1..K$  à une population de neurones  $q \in 1..K$ , où  $K$  est le nombre de populations). Ce modèle renouvelle l’approche large-échelle classique (Wilson & Bower, 1989) en prenant en compte l’hétérogénéité des faisceaux d’axones reliant différentes régions du réseau. Il sert de base à la définition d’architectures neuronales fondées sur des substrats à connectivité aléatoire.

L’étude numérique proposée dans l’article fixe le rapport excitation/inhibition à  $\frac{1}{2}$ , ce qui correspond à un régime de faible activité (Amit & Treves, 1989). Nous étudions le rôle contrasté de deux paramètres de contrôle (voir Figure 6.1).

- Le premier paramètre,  $J$ , représente l’amplification du signal. L’augmentation continue de ce paramètre tend à produire une bifurcation entre un régime simple (point fixe) et une activité complexe (de type chaotique).
- Un second paramètre, nommé  $d$ , représente l’inverse du *coefficient de variation* des liens, autrement dit le degré d’homogénéité. Augmenter ce paramètre tend à rendre les poids plus homogènes.

De manière intéressante, l’augmentation continue de ce second paramètre (augmentation de l’“ordre”) produit une transition entre un régime stationnaire et un régime non-stationnaire périodique. Nous mettons en évidence une carte de bifurcation à 4 régions, comprenant un régime stationnaire de point fixe, un régime stationnaire complexe, un régime d’oscillations simple, et un régime d’oscillations complexes analogue à processus stochastique cyclostationnaire, ce dernier régime étant à notre connaissance mis en évidence pour la première fois dans ce type de réseaux.

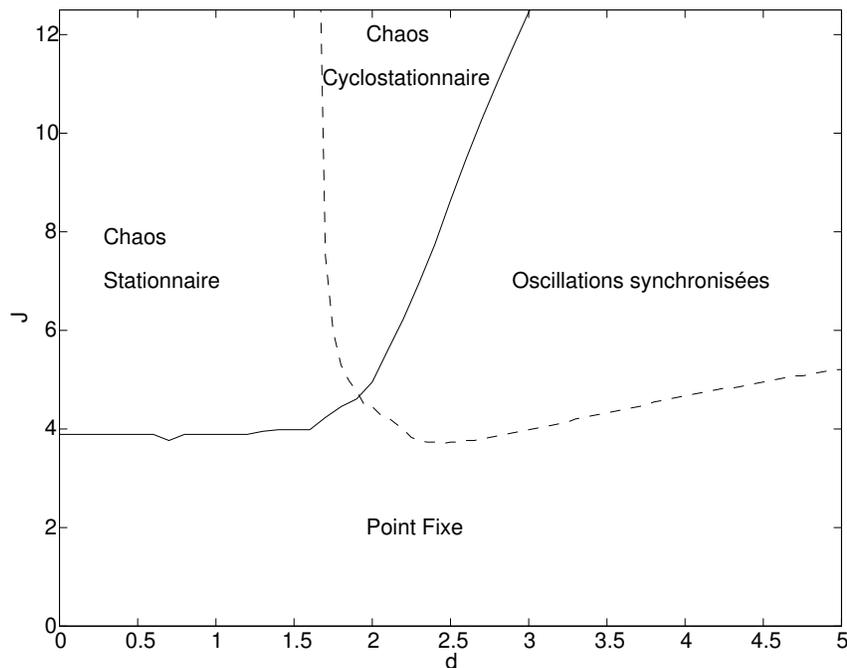


FIGURE 6.1 – Diagramme de bifurcation à la limite thermodynamique sur un réseau aléatoire équilibré à 2 populations, en fonction de  $d$  (homogénéité) et  $J$  (amplification).

Notre étude a donc ouvert la voie à l'étude multi-population systématique des réseaux de neurones à connectivité aléatoire, voir par exemple (Soula et al., 2006; Faugeras et al., 2009; Cabana & Touboul, 2013).

### 6.2.2 Criticalité dans les réseaux aléatoires

Un des intérêts principaux des réseaux à connectivité aléatoire est leur capacité à générer des dynamiques complexes endogènes (donc sans stimulation extérieure). Ces dynamiques, bien que reposant sur des équations déterministes, se comportent comme des processus stochastiques indépendants sur chaque neurone. Ce comportement n'est pourtant pas analogue à un simple bruit passif.

Dans le cas des réseaux de neurones récurrents aléatoires (voir page 28), il est intéressant de constater que le processus stochastique limite, tel que décrit par les équations de champ moyen, a une traduction à taille finie sous la forme d'une activité de type *chaos déterministe*. Les réseaux étudiés étant normalisés, leur comportement est réglé par deux paramètres : le gain  $g$  de la fonction de réponse neuronale (assimilable à l'inverse d'une température) et la variabilité  $\sigma_\theta$  du seuil d'activation. L'existence de ces paramètres de contrôle permet de reproduire, dans ces réseaux aléatoires, un comportement générique de déstabilisation, appelé route vers le chaos par "quasi-périodicité" (Berger et al., 1988; Doyon et al., 1993) (voir figure 6.2). L'existence d'une telle frontière entre l'ordre et le chaos est également prouvée à la limite thermodynamique, mais dans ce cas

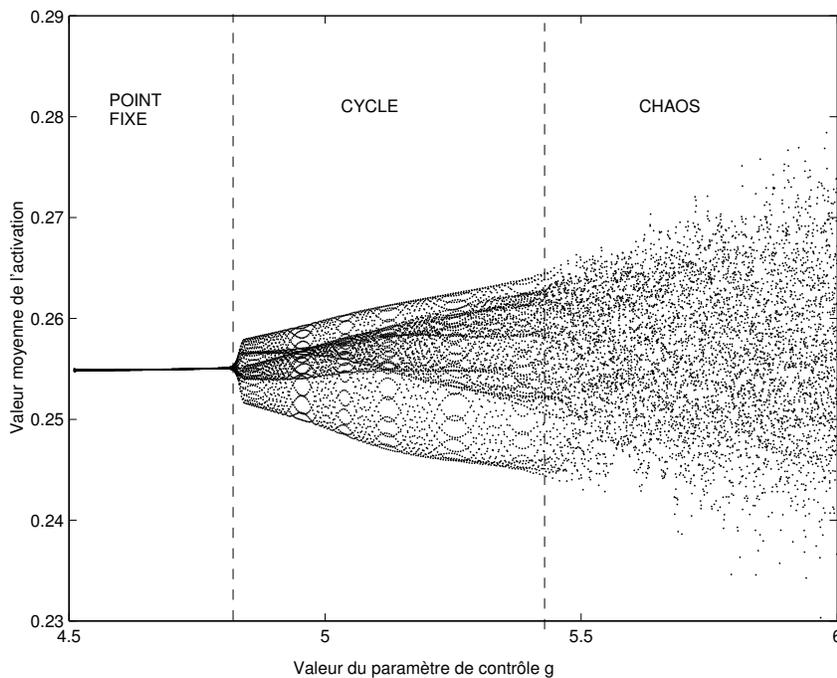


FIGURE 6.2 – Transitions de phase. La figure représente l'activité moyenne d'un réseau récurrent aléatoire de 200 neurones sur 14 000 pas de temps, tandis que l'on augmente lentement le gain entre  $g = 4$  et  $g = 6$  (Beslon & Daucé, 2002)

la transition se fait brutalement, avec un passage discret d'un processus de type point fixe à un processus stochastique (Cessac, 1995). Cette frontière entre ordre et chaos est par principe intéressante puisqu'elle indique le lieu où la variabilité de la dynamique, et donc potentiellement l'expressivité du réseau, est la plus importante.

Le signal présenté en entrée du réseau est un vecteur aléatoire qui vient se surimposer au seuil d'activation (rendant donc les neurones plus ou moins excitables). Le réseau soumis à un tel signal converge vers un nouvel attracteur (autrement dit se reconfigure). Une manière simple de caractériser le comportement non-linéaire (ou critique) d'une population de neurones consiste à introduire de petites perturbations dans la dynamique du réseau.

Une étude proposée dans le 3ème chapitre de l'ouvrage que nous avons co-dirigé avec Agnès Guillot (Beslon & Daucé, 2002) illustre le propos. Cette étude porte sur le modèle le plus simple, un réseau récurrent aléatoire comportant une seule population. Le paramètre de contrôle choisi place le réseau dans une région chaotique, proche de la transition vers le cycle limite. Le réseau est soumis à une stimulation extérieure quasi stationnaire, évoluant à une vitesse très lente par rapport à la vitesse de mise à jour du réseau. Une succession de régimes dynamiques distincts, séparés par des transitions brusques, est mise en évidence (voir figure 6.3). Chaque régime est caractérisé par un patron d'activité différent, qui se manifeste par la présence de fréquences caractéristiques distinctes dans le spectre de Fourier du signal moyen.

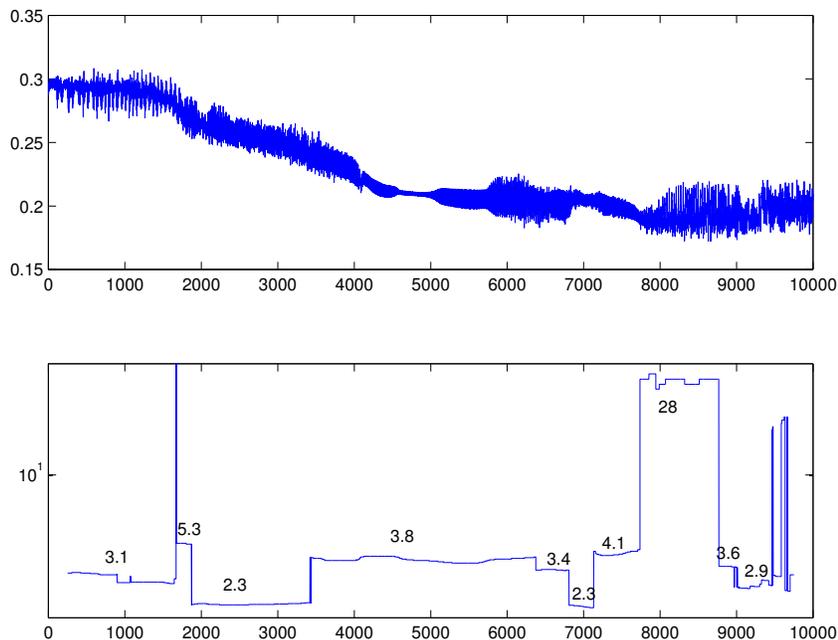


FIGURE 6.3 – Activité d'un réseau récurrent aléatoire de 200 neurones, sur 10000 pas de temps, sous l'influence d'une entrée sensorielle à évolution lente. La période principale est estimée par spectre de Fourier sur une fenêtre de 500 pas de temps (cette période étant représentative de la période propre de chaque neurone actif). a) Activité moyenne. b) Période du mode principal (en échelle logarithmique) (Beslon & Daucé, 2002)

Une autre étude, présente dans mon manuscrit de thèse (Daucé, 2000), développe l'analyse spatio-temporelle des patrons d'activité de type cycle limite ou chaos léger. L'existence d'une chaîne d'activation spontanée (*synfire chain* (Abeles, 1991)) est mise en évidence par analyse de la matrice de covariance des signaux individuels. Différents modes (patrons d'activation spatio-temporels) se développent spontanément lorsque le réseau est soumis à des entrées statiques différentes (Daucé & Quoy, 2000). Enfin, l'analyse de la réponse du réseau à une stimulation périodique indique une "bande passante" de quelques unités temporelles, le réseau étant capable de "capturer" (et donc d'amplifier) des signaux dont la période varie entre 2 et 10-12 unités de temps (autrement dit à sélectionner un mode dynamique périodique interne compatible avec la période imposée) (Daucé & Quoy, 2000) – voir figure 6.4. Le comportement spontané des réseaux récurrents aléatoires est par ailleurs à la base du mécanisme de perception par feedback positif présentés dans la section 7.2.2, page 91.

L'intérêt du comportement spontané des réseaux récurrents à connectivité aléatoires a été largement confirmé depuis. Une étude plus systématique de la réponse linéaire dans les réseaux récurrents (Cessac & Samuelides, 2007) met en évidence l'existence de modes et de résonances extrêmement variés lorsque le signal est appliqué sur un seul nœud du réseau (micro-stimulation). Les modèles de "réservoir computing" proposés par l'équipe de Wolfgang Maas (Maass et al., 2002) reposent sur le même principe : un réseau de neurones aléatoire récurrent soumis à un signal spatio-temporel est capable de conserver une mémoire de quelques unités de temps utilisables pour produire des classificateurs sensibles

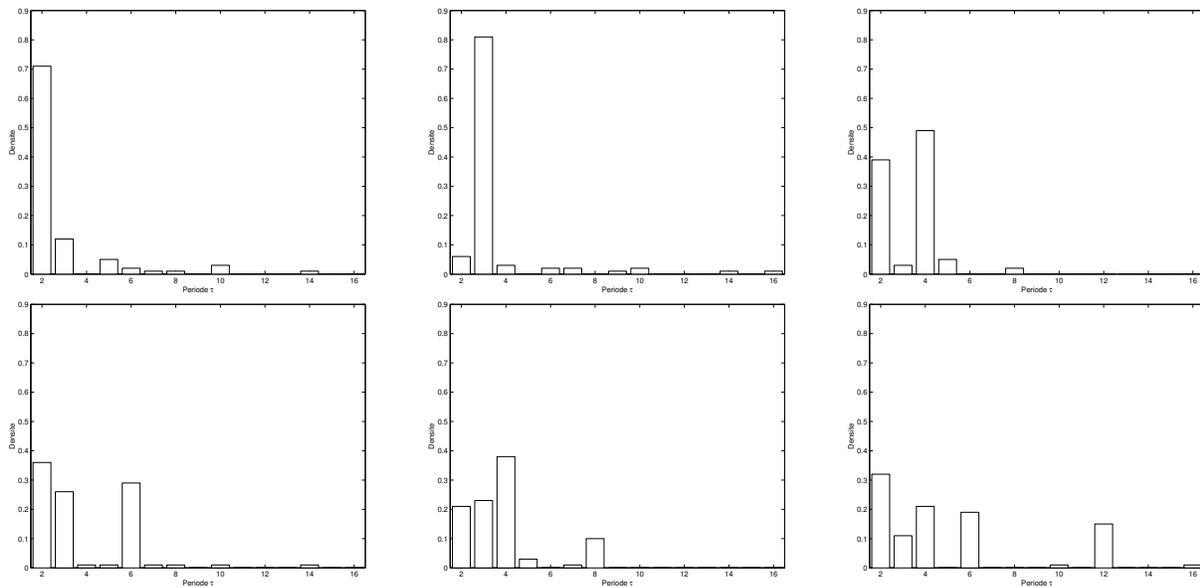


FIGURE 6.4 – **Capture de période pour différentes valeurs de  $\tau$  (période du signal).** On a représenté sur ces histogrammes la répartition des périodes, mesurée sur 100 couples (réseau, séquence d'entrée) pour  $\tau = 2$  (haut, gauche),  $\tau = 3$  (haut, milieu),  $\tau = 4$  (haut, droite),  $\tau = 6$  (bas, gauche),  $\tau = 8$  (bas, milieu) et  $\tau = 12$  (bas, droite). Paramètres :  $N = 200$ ,  $g = 8$ ,  $\bar{I} = 0$ ,  $\mu_\theta = 0$ ,  $\sigma_\theta = 0$ ,  $\mu_J = 0$ ,  $\sigma_J = 1$ ,  $\sigma_I = 0.1$  (Daucé, 2001).

au contexte temporel.

Ce modèle de réseau à connexions aléatoire offre un exemple de substrat dont la caractéristique principale est la *capture dynamique*, reposant sur des transformations morphologiques des patrons d'activités (changements des distributions spatiales et temporelles des activités). Bien que construits sur un tirage aléatoire arbitraire, ils offrent donc un bon support pour le traitement (et la transformation) des signaux temporels. Faire reposer le traitement sur un tirage arbitraire rejoint d'ailleurs un certain nombre de techniques de plongement (ou projection) aléatoires utilisées en apprentissage automatique, tels le "compressive sensing" (Baraniuk, 2007) ou l'architecture "extreme learning" (Huang et al., 2006).

### 6.2.3 Modèles de champ neuronal

La question adressée par le modèle publié en 2004 (Daucé, 2004b) (voir annexe D) est celle de l'implémentation d'un modèle de champ neuronal (voir pages 29 et 75) sur un support composé d'unités discrètes. Sur les modèles jusqu'alors proposés, dans un cas l'activité focale était soit uniquement réactive (Hansel & Sompolinsky, 1996), dans l'autre cas, l'activité persistante reposait sur un mécanisme de bistabilité cellulaire (Camperi & Wang, 1998) ou encore sur une plasticité synaptique à court terme (Compte et al., 2000). Notre modèle est une des premières implémentations d'un champ neuronal sur support discret reposant sur des mécanismes de réseau uniquement.

L'article étend le modèle récurrent aléatoire (Daucé et al., 2001) à des connectivités plus complexes, incluant des délais variables et les connexions dépendantes de la distance. Un des apports de cet article est de montrer la stabilité de certains indicateurs (comme le niveau de réponse moyen) à travers les modèles. Par exemple, l'introduction de délais différents entre les populations excitatrices et les populations inhibitrices conduit à des changements qualitatifs (régime d'oscillations lentes synchronisées) mais non quantitatif (même réponse moyenne).

L'introduction d'une connectivité dépendante de la distance opère un changement plus radical, puisque l'hypothèse fondatrice du modèle (indépendance des tirages) n'est plus vérifiée, et donc l'hypothèse de chaos local (indépendance des activités) n'est plus valide. Les unités neuronales de chaque population reçoivent des coordonnées spatiales (ici prises uniformément sur un intervalle borné). Le tirage des poids synaptiques devient dépendant de la distance qui sépare deux unités. La valeur moyenne d'un poids est identique à celle du modèle précédent, mais les nœuds les plus proches ont en espérance une valeur plus forte, et les nœuds les plus éloignés une valeur plus faible (la distribution des poids synaptiques repose sur le produit entre la distribution uniforme initiale et un noyau gaussien conservatif). Le mécanisme du champ neuronal, qui se caractérise par une connectivité majoritairement excitatrice à courte distance, et inhibitrice à longue distance, est implémenté à l'aide de ces noyaux gaussiens tels que le rayon du noyau des liens excitateurs est plus faible que celui des liens inhibiteurs.

L'article montre alors que les propriétés des réseaux aléatoires récurrents et celles des champs neuronaux s'additionnent dans ce modèle, puisqu'on peut mettre en évidence sur ce substrat la présence simultanée d'une activité de fond chaotique, de grandes oscillations lentes et de foyers d'activité de type champ neuronal. Comme dans le cas du champ neu-

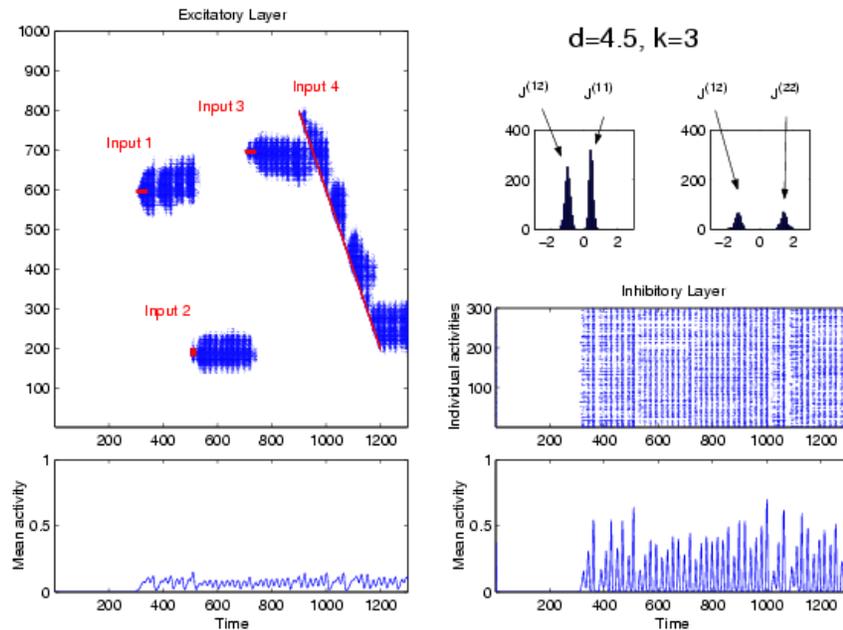


FIGURE 6.5 – Réponse à différentes entrées sur un réseau de neurones organisé topographiquement sur la couche excitatrice. Le réseau comporte 1000 neurones excitateurs et 300 neurones inhibiteurs. Les liens excitateurs sont globalement plus rapides que les liens inhibiteurs. Quatre entrées différentes (trace rouge) sont présentées au réseau. Le premier signal excite les neurones 591-600 entre les instants 301 et 340. Le second signal excite les neurones 181-200 entre les instants 501 et 520. Le troisième signal excite les neurones 691-700 entre les instants 701 et 740. Le dernier signal est une rampe qui traverse la couche excitatrice en activant 10 neurones à chaque instant. **En haut, à gauche** : patron d'activité global de la couche excitatrice. **En bas, à gauche** : activité moyenne de la couche excitatrice. **En haut, à droite** : Histogramme de la somme des poids entrants pour la population excitatrice (gauche) et la population inhibitrice (droite). **Au milieu, à droite** : patron d'activité global de la couche inhibitrice. **En bas, à droite** : activité moyenne de la couche inhibitrice. D'après (Daucé, 2004).

ronal, l'activité persistante repose sur l'activité excitatrice locale, avec un effet de lissage lié à l'hétérogénéité des connexions. Les oscillations lentes reposent comme précédemment l'existence de délais différenciés. Ce sont ces oscillations qui donnent au substrat la capacité de répondre à des changements de faible amplitude (augmentent donc la sensibilité du substrat au signal d'entrée). Il en résulte une implémentation d'un champ neuronal sur un substrat composé d'unités neuronales discrètes.

De manière plus générale, ce modèle nous fournit un premier exemple, dans le cadre des réseaux récurrents aléatoires, de l'effet de mémoire produit par une boucle de rétroaction positive.

Notre modèle propose une passerelle simple entre les modèles de champ neuronal continu et de réseaux de neurones à état et à temps discret. Il s'étend facilement à des réseaux de neurones intègre-et-tire (non publié), avec des propriétés peu différentes qualitativement de celles qu'on obtient avec des modèles à fréquence de décharge. Ce modèle

a été utilisé en tant que couche perceptive dans des architectures de contrôle (Daucé, 2004a; Daucé, 2007). D'autres modèles visant une plus grande fidélité aux processus biologiques ont étudié l'effet de délais variables et de patrons de connexion hétérogènes (voir par exemple (Roxin et al., 2005)), et plus généralement l'étude de l'interaction entre délais variables, topographie et dynamique reste un sujet d'actualité en neurosciences computationnelles (Voges & Perrinet, 2012; Lundqvist et al., 2012).

### 6.2.4 Extension aux populations de neurones impulsionnels

Le passage des neurones binaires aux neurones à potentiels d'action modifie la nature des régimes dynamiques produits par le réseau. Une différence importante avec les modèles à sortie continue est la définition précise de l'échelle temporelle sur les délais de transmission comme sur les constantes de temps et les périodes réfractaires. L'activité produite par ces réseaux apparaît plus "désordonnée" que celle des réseaux à sortie continue. Ainsi, aucune activité de type cycle limite n'est mise en évidence : seules les activités de type chaotiques sont observées lorsque le réseau est actif ; si les liens récurrents ne sont pas assez forts, l'activité s'éteint. Pas de paramètre de contrôle ici permettant de décrire une route vers le chaos par quasi-périodicité, qui était une des caractéristique des réseaux récurrents aléatoires à sortie continue.

Les architectures multi-populations ont été approfondies à partir de l'approche par "faisceaux d'axones" (voir section 6.1), indépendants de la définition spatiale. Les divers paramètres de contrôle en jeu, à savoir le taux d'amplification, le coefficient de variation, le délai moyen participent à définir des architectures de contrôle diversifiées définies principalement par le nombre et la fonction des populations en jeu. D'autres paramètres, comme la densité de connexion, le poids synaptique moyen etc. apparaissent conditionnés aux paramètres structurels principaux.

Mes contributions dans ce domaine portent sur l'étude des transferts d'échelle entre neurones fréquentiels et neurones à potentiels d'action, à la définition de patrons de connectivité séparant clairement neurones excitateurs et neurones inhibiteurs, et à la mise en place d'un formalisme architectural indépendant des modèles. Le principe qui, dans notre cas, unifie l'approche des neurones à sortie continue et des neurones à sortie discrète est présenté dans (Daucé, 2007) - voir annexe G. Il repose sur un modèle de neurone intègre-et-tire normalisé à résolution temporelle variable. Le passage du neurone discret "à fuite" au neurones binaires (McCulloch et Pitts), puis au neurone à sortie continue est obtenu par simple variation du pas d'intégration du schéma d'Euler de l'équation différentielle qui définit le neurone. Il en ressort une unification des modèles à travers les échelles temporelles où les neurones intègre-et-tire représentent une résolution de l'ordre de la milliseconde, les neurones binaires une résolution de l'ordre de la dizaine de millisecondes et les neurones à sortie continue une résolution de l'ordre de la centaine de millisecondes.

Les architectures à populations excitatrices et inhibitrices "équilibrées" présentent, dans les mêmes gammes de paramètres que précédemment, des comportements de population très similaires. Les régimes d'oscillations lentes synchronisées se retrouvent également dans ces modèles. Des transitions de phase sont observables entre différents régimes dynamiques sous l'effet de la variation de certains paramètres de contrôle, ou sous l'effet de la plasticité synaptique (Henry & Daucé, 2008a; Henry & Daucé, 2008b; Henry & Daucé,

2009) - voir annexe H.

Manipuler des architectures neuronales à différentes résolutions spatiales et temporelles permet d'estimer les éléments structurels communs aux différents modèles indépendamment du niveau de résolution (et, de façon complémentaire, d'identifier les éléments structurels spécifiques à un niveau de résolution spatial ou temporel donné, et/ou d'estimer les limites d'expressivité des modèles à basse résolution). Ainsi, nous avons pu confirmer le caractère *non spécifique* des régimes d'oscillations lentes, qui n'apparaissent pas conditionnés au type de cellule utilisée (à sortie discrète ou continue), mais reposent plutôt sur les paramètres et effets de réseau, comme les fréquences moyennes, le taux d'amplification et les délais de transmission entre les différentes populations. Les effets de réseau (de "masses" neuronales) apparaissent prédominants dans les réseaux récurrents aléatoires et les simulations à "haute résolution" viennent souvent confirmer des comportements déjà obtenus sur les modèles à basse résolution.

Nos études n'ont pas permis de mettre en évidence des régimes dynamiques nouveaux propres aux échelles temporelles fines et à l'activité discrète. Cependant, certains effets de mémoire à court terme, sur le principe du "réservoir" de dynamiques vu précédemment, ont montré que l'activité auto-entretenu permet la prise en compte de dépendances entre des événements séparés par plus de 100 ms (Henry et al., 2007) - voir annexe F. La présence de séquences d'activation spatio-temporelles reproductibles (dites "groupes polychrones"), telles que postulées par (Izhikevich, 2006) en tant qu'unités opérationnelles élémentaires, peut être mise en évidence dans nos réseaux aléatoires. Une étude que j'ai publiée récemment, utilisant des unités neuronales un peu plus détaillées (Daucé, 2014) - voir annexe M, montre que la réponse de population à des stimuli spatio-temporels suit une séquence reproductible (mais bruitée) qui peut être amplifiée par la plasticité. La sensibilité acquise à des patrons spatio-temporels spécifiques, telles qu'existant également dans les modèles fréquentiels, peut donc être observée à l'échelle de quelques millisecondes, mettant en jeu un faible nombre de potentiels d'action par neurone. La clé semble donc être la mise en évidence de mécanismes similaires à différentes échelles, les opérations à échelle temporelle courte étant elles-mêmes "portées" (ou embarquées) dans des opérations à échelle temporelle plus larges.

### 6.2.5 Modèles de la connectivité fonctionnelle large-échelle

L'apparition récente des techniques de tractographie par tenseurs de diffusion (Le Bihan et al., 2001) permet d'analyser la matière blanche du cerveau pour suivre la direction des faisceaux de fibres. Il est possible d'en déduire un patron de connectivité (le "connectome") reliant les principales régions du cortex. Le patron de connectivité issu de cette analyse forme un réseau, constitué par deux matrices, l'une fournissant la distance entre les régions connectées, l'autre fournissant le "poids" de cette connexion en fonction de la densité de fibres. Ce réseau est relativement stable à travers les sujets.

Le connectome offre une vue d'ensemble de l'architecture du cerveau. L'analyse du connectome (analyse "structurelle") vient en complément de l'analyse des patrons d'activité observés dans l'exercice d'une tâche (analyse fonctionnelle). Dans le cadre de cette analyse structurelle, plusieurs indicateurs ont été proposés pour distinguer les nœuds les plus centraux (les plus "importants") des nœuds périphériques. Il est possible de mettre

en évidence d'un "noyau central" (Hagmann et al., 2008) regroupant des nœuds médians servant de lien (ou de relais) entre la plupart des autres régions. Cet ensemble de nœuds présente des points communs avec le patron d'activité "par défaut" observé en imagerie fonctionnelle (l'activité observée lorsque le sujet n'est pas occupé à accomplir une tâche) (Raichle et al., 2001).

En partant de la structure du connectome, nous avons utilisé pour l'activation des nœuds différentes variantes du modèle "à réponse graduelle" de Hopfield (Hopfield, 1984), en changeant les caractéristiques du seuil d'activation afin de contrôler le niveau d'activation moyen (Golos, Jirsa & Daucé, 2015) - voir annexe P. Ces travaux ont permis de mettre en évidence un nombre très important d'attracteurs dans certaines gammes de paramètres. Ainsi, en condition de "basse température", le nombre d'attracteurs distincts peut atteindre plusieurs milliers (sur un modèle comportant 998 nœuds). L'analyse des ensembles d'attracteurs obtenus permet d'identifier une huitaine de "modes" différents (voir figure 6.6), chaque mode correspondant à l'activation d'une région anatomique particulière. Ces modes peuvent être comparés aux modules régionaux mis en évidence par une analyse structurelle (Hagmann et al., 2008) ainsi qu'avec les composants indépendants identifiés sur les signaux fMRI de l'état de repos (Damoiseaux et al., 2006).

L'ajout d'un terme de bruit dans l'équation de mise à jour permet de simuler une activité multistable *temporellement*, s'apparentant aux régimes dits de "dynamique itinérante" (Tsuda, 1992; Kaneko, 1990), et dont les caractéristiques sont plus proches des signaux observés à l'état de repos. Cette multistabilité temporelle conduit fréquemment le réseau vers des régimes d'activité plus simples (nœuds uniformément actifs ou inactifs) par dérive lente de l'activité moyenne. Seul un modèle à "feedback négatif global" présente une multistabilité temporelle à long terme, grâce au contrôle du niveau d'activité qui empêche d'atteindre les régimes triviaux.

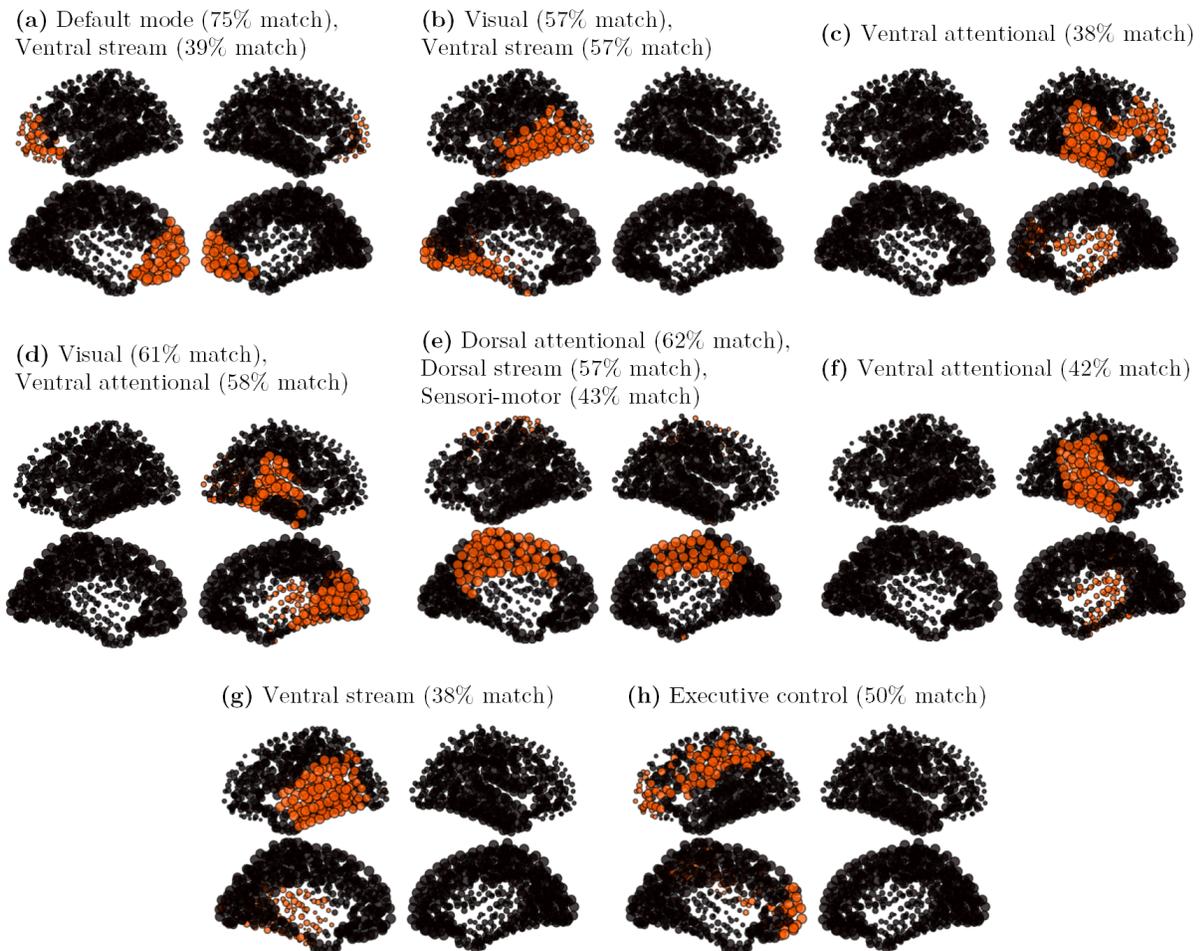


FIGURE 6.6 – Vecteurs caractéristiques des huit principaux clusters observés sur un jeu d'attracteurs issu du connectome humain (Hagmann et al., 2007). Le modèle de mise à jour est inspiré de (Hopfield, 1984). Un ensemble de 28,849 attracteurs distincts est extrait de 30,000 simulations. Les clusters sont obtenus à l'aide d'une mesure de similarité basée sur la distance de Hamming, avec un seuil de similarité fixé à  $k = 0.8$ . Les vecteurs caractéristiques sont définis comme les 10% des nœuds les plus actifs dans un cluster. **a** : vecteur caractéristique du cluster 1 (4,089 attracteurs) ; **b** : vecteur caractéristique du cluster 2 (3,048 attracteurs) ; **c** : vecteur caractéristique du cluster 3 (2,045 attracteurs) ; **d** : vecteur caractéristique du cluster 4 (1,573 attracteurs) ; **e** : vecteur caractéristique du cluster 5 (907 attracteurs) ; **f** : vecteur caractéristique du cluster 6 (892 attracteurs) ; **g** : vecteur caractéristique du cluster 7 (881 attracteurs) ; **h** : vecteur caractéristique du cluster 8 (844 attracteurs) (d'après Golos et al (2015)).

# 7

## Plasticité synaptique et apprentissage

Ce chapitre est consacré à l'étude des mécanismes d'apprentissage dans les réseaux de neurones récurrents. Nous regardons comment les règles de plasticité synaptique vont permettre à certaines stimulations sensorielles (désignées comme intéressantes) de “creuser” au sein du réseau un mode de réponse particulier (une “résonance”), cette résonance étant le support des interpolations perceptives futures. Nous regardons également en quoi ces architectures dites “à feedback positif”, contrairement au modèle classique du “predictive coding”, implémentent le comportement de “résistance au changement” (ou persévérance).

### 7.1 Modèles de la perception

Selon les modèles classiques, la perception sensorielle s'apparente à un mécanisme de *filtrage* des entrées perceptives. Le filtrage d'un signal correspond fonctionnellement au fait de pouvoir *prédire le présent*. Tout écart entre la prédiction et l'observation constitue un erreur de prédiction. La perception sensorielle est interprétée comme le résultat de l'intégration des erreurs perceptives afin de mettre à jour le modèle.

Le filtrage repose sur un *modèle* prédictif (ou modèle interne) généralement implémenté selon le principe de l'“observateur” (Kalman, 1960; Wolpert & Ghahramani, 2000), où :

- La sortie du modèle correspond à la prédiction actuelle,
- Le signal d'entrée correspond à l'observation effective.

La présence d'une erreur sensorielle fait entrer le système dans un conflit qui peut se résoudre de deux manières :

- Les modèles de la perception fondés sur la rétroaction négative (Rao & Ballard, 1999; Friston & Kiebel, 2009) font disparaître de la scène sensorielle les éléments prédits, de sorte que, selon cette théorie, ce qui est “perçu” (ou intégré) est l'erreur sensorielle, autrement dit la différence entre sensation et la prédiction.
- Les modèles de la perception fondés sur la rétroaction positive “complètent” le signal d'entrée avec le signal prédit :
  - Si l'écart est faible, le système tend à ignorer ce changement et persévère dans

sa prédiction.

- Si l'écart est important et/ou réitéré, le système met à jour (reconfigure) son activité interne (son "modèle") afin de mieux anticiper l'entrée.

La persévération présente de nombreuses analogies avec les mécanismes de mémoire à court terme (voir page 81). Si une partie du signal disparaît par exemple, l'activité interne est capable de compléter les morceaux manquants, selon un mécanisme auto-associatif qui s'apparente à celui des réseaux de Hopfield. Le réseau peut continuer par exemple à "croire" à l'intégrité des éléments de la scène visuelle (sensorielle) non perçus à cause de phénomènes d'occlusion.

D'un point de vue fonctionnel, chacun de ces modèles de la perception présente des avantages et des inconvénients :

- La rétroaction positive apporte de la robustesse, de la mémoire à court terme (complétion, interpolation des données manquantes, traitement des ambiguïtés perceptives) mais peut également produire une résistance excessive aux changements de la scène sensorielle. Le cas limite correspond au réseau qui ignore totalement son entrée sensorielle, persévérant indéfiniment sur une illusion perceptive produite par l'activité interne ("hallucination").
- La rétroaction négative apporte une parcimonie (perception de la différence d'activité), qui permet d'implémenter un principe de "matching pursuit" pour un traitement rapide du flux visuel (Perrinet, 2010). Il présente cependant une sensibilité au bruit de mesure, et nécessite dans certains cas la mise en place d'une remise à zéro périodique pour "réinitialiser" le modèle.

## 7.2 Contributions personnelles

### 7.2.1 Route inverse

L'étude du mécanisme de reconnaissance par "réduction" de la dynamique constitue le cœur de ma thèse, sous la direction de Bernard Doyon, chercheur à L'INSERM au sein de l'équipe dirigée par Manuel Samuelidès à l'ONERA de Toulouse, entre 1996 et 1999. Certains travaux réalisés pendant la thèse de Mathias Quoy (1991-1994) (Quoy, 1994) avaient déjà montré l'effet de réduction produit par la plasticité Hebbienne dans les réseaux aléatoire récurrents à dynamique chaotique intrinsèque.

L'article qui présente ces idées de la manière la plus complète est celui de Neural Networks (Daucé et al., 1998) (voir annexe A page 143). Il porte sur l'étude de la réaction (spontanée ou acquise) d'un réseau de neurones récurrent aléatoire à des patrons (stimuli) également aléatoires. Il fait la synthèse de plusieurs résultats sur les réseaux récurrents aléatoires, à la fois sur les comportements de grande taille et sur les effets de taille finie, et présente pour la première fois les effets de la plasticité sur ce type de réseau. Il reprend pour ce qui me concerne des travaux réalisés pendant ma première année de thèse. Les réseaux étudiés étant normalisés, leur comportement est réglé par deux paramètres : le gain de la fonction de réponse neuronale (assimilable à l'inverse d'une température) et la variabilité du seuil d'activation. Le signal présenté en entrée du réseau est un patron aléatoire statique qui vient se surimposer au seuil d'activation (rendant donc les neurones plus ou moins

excitables). Le réseau soumis à un tel signal converge vers un nouvel attracteur (autrement dit se reconfigure).

Le changement observé suite à la présentation d'un stimulus statique ne correspond pas à de la multistabilité au sens classique (un système dynamique multistable pouvant converger vers différents attracteurs pour des conditions initiales différentes). Ici le stimulus fait partie des "paramètres" du réseau. Ainsi, chaque stimulus distinct définit système dynamique distinct convergeant vers un attracteur distinct. Le fait de modifier le stimulus au cours d'une même simulation revient à modifier le système dynamique "en cours de route". La réalisation d'une opération (sous la forme d'une réorganisation de la dynamique suite à l'imposition d'un motif statique) repose donc sur un mécanisme différent de celui proposé par Hopfield :

- Il s'agit essentiellement d'une relaxation sous contrainte, chaque stimulus étant l'expression d'une contrainte différente.

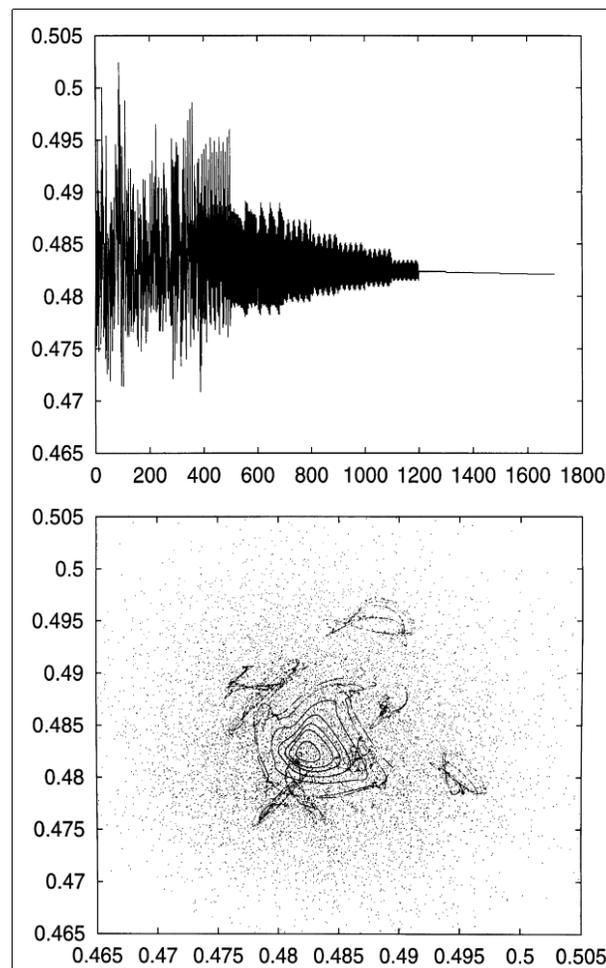


FIGURE 7.1 – (Haut) Activité moyenne  $m_{\text{net}}(t)$  en dynamique contrainte en fonction du temps  $t$ , au cours d'une séance d'apprentissage. La règle de plasticité est appliquée tous les 100 pas de temps. Dans cet exemple, la dynamique se réduit sur un point fixe après 12 cycles d'apprentissage ( $t = 1200$ ). (Bas) Même activité, représentée dans le plan de phase  $[m_{\text{net}}(t), m_{\text{net}}(t + 1)]$ . On notera la décroissance du diamètre du cycle limite à chaque application de la règle de plasticité (extrait de Daucé et al., 1998).

— L'ensemble des réponses aux contraintes statiques possibles constitue le *répertoire dynamique* du réseau.

Une règle de plasticité synaptique est ajoutée, correspondant à une dynamique lente sur les couplages, qui modifie le graphe selon une règle inspirée de Hebb (Hebb, 1949) :

$$J_{ij}(t) - J_{ij}(t-1) \propto (s_i^{\text{out}}(t) - 0.5) \times (s_j^{\text{in}}(t-1) - 0.5)$$

(à savoir que le lien entre deux neurones se renforce s'il existe une corrélation entre l'activité pré-synaptique et l'activité post-synaptique). L'article étudie les changements dans le répertoire dynamique lorsque le réseau est soumis à ces changements synaptiques dans le contexte d'une contrainte statique spécifique. L'article met en évidence des changements de comportements liés à cette dynamique sur les poids, conduisant systématiquement à une simplification de l'activité intrinsèque (par exemple passage d'une activité chaotique à une activité de type cycle limite). On parle de "réduction" de la dynamique, la plasticité produisant une "route inverse" de la complexité vers le point fixe (voir figure 7.1).

Une fois la session d'apprentissage terminée, la présentation du stimulus appris a pour effet d'induire cette dynamique moins complexe (obtenue en fin de session). Cette activité est qualitativement différente, plus simple que l'activité induite par les autres stimuli. La plasticité permet donc d'obtenir une réaction spécifique à une entrée particulière (ici une forme arbitraire générée par un tirage aléatoire).

La plasticité Hebbienne a donc pour effet de créer des comportements "acquis" spécifiques. Le mécanisme de Hebb (renforcement des liens entre les neurones les plus actifs) tend à produire une classe de réactions différente face aux motifs appris. Les motifs en question sont "reconnus". Le reste des motifs, ne présentant pas ces caractéristiques spatiales spécifiques, induisent une réaction "par défaut" qualitativement différente. Il y a donc, pour des motifs quantitativement similaires (issus d'un tirage aléatoire de même loi), des réponses qualitativement différentes. Ce mécanisme de séparation entre le connu et l'inconnu (appariement conditionnel) rejoint l'hypothèse formulée par Freeman (Skarda & Freeman, 1987) à propos de la perception des odeurs : activité simplifiée pour des odeurs connues, activité de fond chaotique pour les odeurs non familières.

Chaque réponse à un motif se traduit par une activité spatio-temporellement distincte, l'apprentissage ayant pour effet de capturer (ou plutôt accentuer) un *mode dynamique* particulier. Le réseau est vu comme un "réservoir" de comportements dynamiques spatio-temporellement différents. Le comportement spatio-temporel associé au motif change sous l'effet de la plasticité, en conservant certaines de ses caractéristiques comme la période intrinsèque, qui se manifeste de façon "pure" sous la forme d'un cycle limite. Ce comportement temporellement périodique finit par disparaître si l'apprentissage est poussé jusqu'au point fixe. La règle de plasticité Hebbienne, dans sa forme simple étudiée ici, n'est pas convergente et l'apprentissage est stoppé de manière arbitraire à l'apparition du cycle limite pour les besoins de l'étude.

La principale difficulté rencontrée pour ces réseaux est la contamination de la réaction induite à des patrons non appris. Si des motifs voisins du motif initialement appris induisent également une réduction de la dynamique, la répétition de l'apprentissage sur plusieurs motifs différents conduit invariablement à un effondrement généralisé de la complexité, étendant la réduction de dynamique à l'ensemble des motifs. La capacité d'un

réseau est définie comme le nombre de motifs pouvant être appris sans provoquer cet effondrement. Cette capacité s'avère assez faible en pratique. Dans notre cas, comme dans le cas du réseau de Hopfield (Hopfield, 1982), elle semble dépendre linéairement de la taille du réseau.

Des changements marginaux sur la règle de plasticité, en limitant drastiquement le nombre de liens éligibles, permettent d'augmenter significativement la capacité des réseaux (Daucé & Doyon, 1998b; Daucé & Doyon, 1998a), avec une "capacité de charge" ("*loading capacity*") empirique de l'ordre de 8% (le nombre de motifs pouvant être appris avant l'effondrement catastrophique est de l'ordre de 8 % du nombre de neurones, ce qui reste relativement faible).

L'article de Neural Networks, en démontrant pour la première fois l'effet de la plasticité Hebbienne sur la dynamique intrinsèque des réseaux récurrents aléatoires, a reçu une bonne audience, dans la mesure où il est un des premiers à présenter un cadre mathématique pour l'étude des réseaux de neurones récurrents perturbés par un signal extérieur. Le même principe est utilisé, dans une certaine mesure, dans les modèles de contrôle proposés par Jun Tani dans les années 90 (Tani, Ito & Sugita, 2004), ainsi que dans les modèles de réservoir computing proposés par Maass et Markram (Maass, Natschläger & Markram, 2002; Buonomano & Maass, 2009).

### 7.2.2 Architectures multi-couches

Les travaux présentés dans (Daucé, 2000; Daucé et al., 2002) (voir annexe C) présentent une architecture neuronale permettant d'implémenter un modèle de la perception des signaux spatio-temporels. Ils reposent sur une architecture de réseaux récurrents aléatoires à plusieurs couches (deux ou trois). Le formalisme multi-couches est le même que celui utilisé dans les réseaux équilibrés (voir section 6.1). Ici les couches se distinguent non par la nature de leurs neurones mais par le rôle qu'elles prennent dans le processus d'apprentissage :

- une (ou deux) couche(s) primaire(s) ayant un rôle passif (sans dynamique intrinsèque)
- et une couche secondaire (ou associative) possédant une dynamique intrinsèque, analogue aux réseaux récurrents aléatoires simples présentés dans la section 7.2.1.

Le modèle se présente comme une architecture générique pour l'apprentissage de motifs spatio-temporels (voir figure 7.2). La contrainte extérieure se manifeste ici sous la forme d'un signal spatio-temporel périodique, et la relaxation s'apparente à celle d'un oscillateur forcé. Dans cette étude, comme dans l'étude précédente, l'accent est mis sur la plasticité et l'apprentissage, où le comportement "différent" est un comportement acquis.

La règle Hebbienne utilisée est une règle d'ordre 1 (fondée sur les différences d'activité), qui modifie les couplages synaptiques selon une estimation locale de la covariance entre les neurones pré et post-synaptiques, en tenant compte du délai de transmission (Sejnowski, 1977).

$$J_{ij}(t) - J_{ij}(t-1) \propto (s_i^{\text{out}}(t) - m_i^{\text{out}}(t)) \times (s_j^{\text{in}}(t-1) - m_j^{\text{in}}(t-1))$$

avec :

$$m_i(t) = \beta s_i(t) + (1 - \beta)m_i(t-1), \text{ avec } \beta \in ]0, 1[$$

Cette règle, contrairement à la précédente, tend à amplifier les variations d'activité. Elle tend asymptotiquement vers une activité de type cycle limite très robuste à large amplitude individuelle (et non vers un point fixe). Comme dans le cas précédent, du fait de la non-convergence asymptotique, la durée d'apprentissage est fixée par l'utilisateur/concepteur en vue d'optimiser un certain comportement (ici la résonance perceptive).

Dans la configuration initiale, la couche primaire reçoit passivement le signal spatio-temporel (séquence périodique de motifs). Les liens feed-forward aléatoires transmettent ce signal à la couche secondaire (dite couche associative). L'activité de la couche associative est un mélange entre cette projection aléatoire du signal d'entrée et l'activité intrinsèque chaotique (entretenu par les liens latéraux aléatoires de la couche secondaire). Les liens de feedback (secondaire vers primaire) sont initialement nuls.

La plasticité synaptique a lieu sur deux classes de liens : latéraux et feedback. Outre la réduction de la complexité de l'activité, la plasticité des liens de feedback crée une boucle d'amplification positive entre la couche associative et la couche sensorielle, qui transforme la corrélation statistique en lien de causalité (voir figure 7.3).

- L'existence de correspondances statistiques entre l'activité interne et le signal d'entrée produit en effet une amplification du signal d'entrée via cette activité de feedback.
- Cette projection induite de l'activité interne vers la couche sensorielle agit comme une prédiction *dans le présent* (voir section 7.1), autrement dit fait reposer la connaissance de l'entrée sensorielle actuelle sur des informations observées dans le passé.

Le changement par rapport au modèle à une couche réside dans l'utilisation effective des changements qualitatifs induits par la plasticité sous la forme du signal de feedback. La plasticité permet de *constituer des résonances* spécifiques à certaines configurations spatio-temporelles apprises. Ces configurations connues, en activant une résonance, révèlent leur caractère prédictible par le réseau.

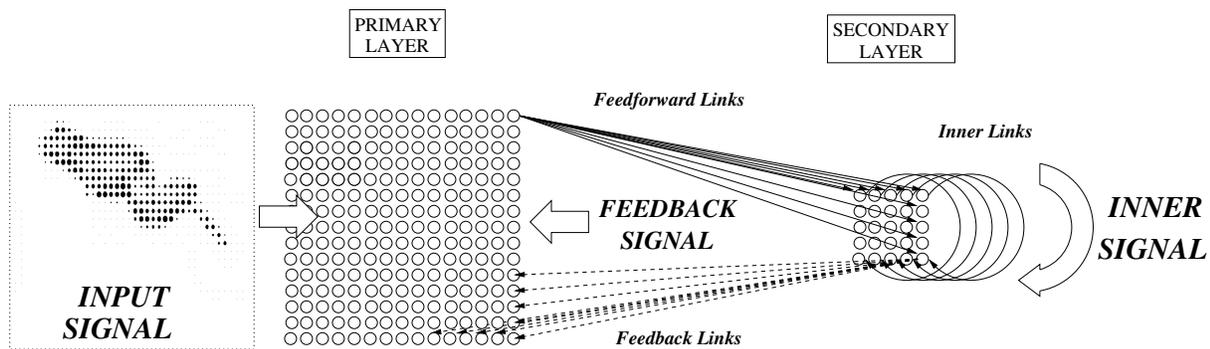


FIGURE 7.2 – Architecture du modèle ReST. Notre modèle est composé de deux couches, dont la taille n'est pas nécessairement égale. Seuls quelques liens sont représentés. La couche primaire est soumise à un signal d'entrée spatio-temporel. La couche secondaire n'est soumise à aucun signal. Les liens sont monodirectionnels. L'activité de la couche secondaire (signal interne) est chaotique. L'activité de la couche primaire dépend à la fois du signal d'entrée et du signal de rétroaction en provenance de la couche secondaire (extrait de Daucé et al, 2002).

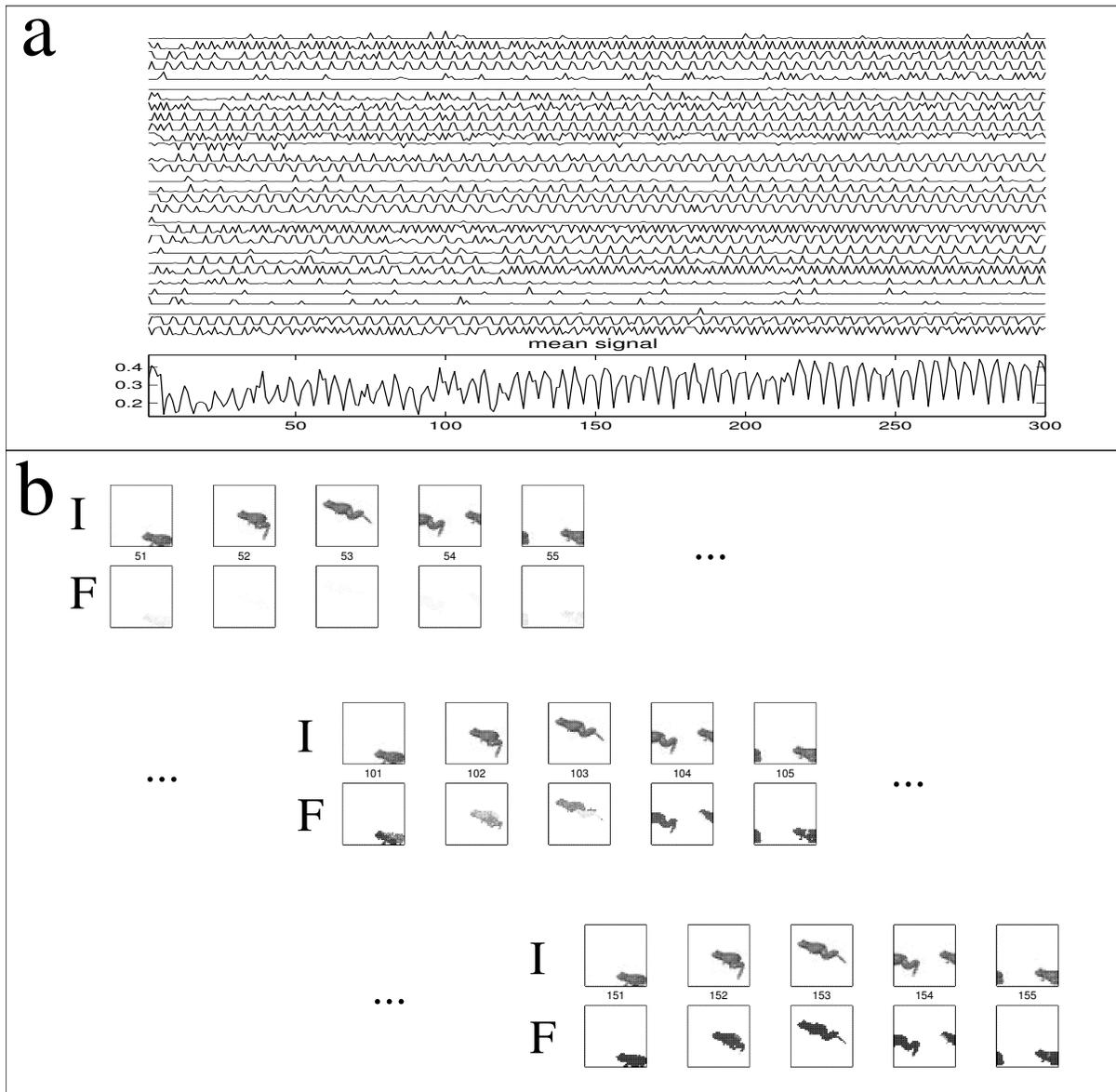


FIGURE 7.3 – **Dynamique d'apprentissage, entre  $t = 1$  et  $t = 300$ , avec  $N^{(1)} = 1600$  et  $N^{(2)} = 200$ .** Le système est continuellement stimulé par un signal spatio-temporel périodique de période 6 (représentant un saut de grenouille) **a.** Activité neuronale sur la couche secondaire. 30 signaux individuels (sur 200) sont représentés. La trace inférieure donne l'activité globale moyenne. **b.** Evolution au cours du temps du signal d'entrée  $I(t)$  et du signal de retour de la couche secondaire  $F(t)$ , pour  $t \in 51, \dots, 55$ ,  $t \in 101, \dots, 105$  et  $t \in 151, \dots, 155$ . A chaque pas de temps, les 1600 valeurs projetées sur la couche d'entrée sont présentées en niveaux de gris sous forme de vignettes  $40 \times 40$ , avec le blanc correspondant à 0 et le noir à 1 (extrait de Daucé et al, 2002).

Plusieurs articles de conférences apportent des éclairages complémentaires sur certains aspects du modèle. L'article de la conférence ICANN 2001 (Daucé, 2001) présente une version préliminaire du modèle perceptif à deux couches avec une règle d'apprentissage légèrement plus simple que celle de l'article de Biological Cybernetics. L'article de SAB (2000) (Daucé & Quoy, 2000) comporte une partie montrant l'effet de l'apprentissage sur des signaux mixtes composés d'une partie statique et d'une partie périodique. L'association apprise entre motifs statiques et motifs dynamiques se manifeste alors par un niveau de corrélation élevé entre la dynamique induite par les motifs mixtes et la dynamique induite par les motifs statiques seuls (les activités étant par contre qualitativement différentes).

### 7.2.3 Plasticité dans les populations de neurones impulsionnels

En complément de l'analyse des effets de population précédemment mentionnés, j'ai étudié, dans le cadre de l'ACI "Temps et Cerveau", les conditions de l'émergence d'oscillations synchronisées *induites* dans des réseaux équilibrés constitués d'unités binaires (autrement dit neurones impulsionnels sans mémoire). Les principaux résultats concernant l'application de la plasticité sur ces réseaux sont détaillés dans l'article EPJ-ST de 2007 (Daucé, 2007) (voir annexe G). Entre autres,

- l'article présente une nouvelle règle de plasticité Hebbienne, la règle TD (qui présente des analogies avec la règle des différences temporelles de Sutton et Barto (Sutton & Barto, 1998)), qui est une règle du premier ordre (voir page 49) stricte fondée sur la conjugaison de l'activité entrante et la *différence d'activité* sortante :

$$\dot{J}_{ij} = F((s_i^{\text{out}}(t) - s_i^{\text{out}}(t-1)) \times s_j^{\text{in}}(t), J_{ij})$$

- L'étude présente en premier lieu une région paramétrique dans laquelle on peut observer des transitions entre un régime de faible activité et un régime chaotique synchronisé, en fonction du coefficient de variation (une moindre variabilité des poids produisant ici un régime plus chaotique).
- En partant d'un de ces régime à activité faible non-synchronisée, la plasticité a pour effet de produire une transition vers un régime à activité synchronisée. La passage d'une activité non-synchrone à une activité synchronisée est spécifique aux réseaux équilibrés (la synchronisation à grande échelle reposant sur les effets de balancier entre population excitatrice et inhibitrice).
- Comme dans les autres études, il est possible d'appliquer la plasticité lorsque le réseau est soumis à un motif statique. Dans ce cas, la plasticité inscrit dans le réseau des comportements spécifiques (ici de grandes oscillations synchronisées) associées aux motifs appris.

Le travail sur des modèles plus réalistes biologiquement a été réalisé, en co-direction avec Gilles Montagne, sur le projet de thèse de Frédéric Henry, étudiant du master de sciences cognitives à Lyon (voir section 1.2).

- Frédéric Henry a travaillé sur un simulateur de réseaux de neurones impulsionnels :
- simulant des délais de transmission variables entre les différents neurones

- implémentant différentes règles de plasticité synaptique, dont la STDP (Spike-Timing Dependent Plasticity).

**STDP** La plasticité dépendante des instants de tir (*Spike-timing dependent plasticity* – STDP) est une règle de plasticité dépendant de l'ordre de tir des neurones pré-synaptique et post-synaptique. Cette règle est facilitatrice pour les synapses où l'activité pré-synaptique précède causalement l'activité post-synaptique, et, à l'inverse, suppressive pour les synapses où l'activité post-synaptique précède (de manière anti-causale) l'activité pré-synaptique. Cette règle tend donc à favoriser les conjugaisons d'activité causalement liées, mais surtout, de manière symétrique, d'empêcher les activités non causalement liées.

La mise en évidence de cette règle sur les synapses naturelles (Bi & Poo, 1998) a favorisé le développement de nombreux modèles computationnels :

- Dans des architectures feed-forward, la STDP réduit la latence de réponse des neurones aux patrons spatio-temporels appris, permettant d'implémenter un codage par rang (Sen Song, 2000; Guyonneau et al., 2005).
- Le cas des réseaux récurrents est plus ambigu :
  - dans certaines implémentations, la STDP semble favoriser l'émergence de séquences d'activation stables (Izhikevich et al., 2004; Izhikevich, 2006).
  - D'autres modèles, utilisant une implémentation différente, indiquent une absence de structuration de l'activité, avec des synapses instables sur toute la durée de la simulation (Morrison et al., 2007).

Une formulation possible de la règle est la suivante :

$$\dot{J}_{ij} = F((s_i^{\text{out}} \times m_j^{\text{in}}) - (s_j^{\text{in}} \times m_i^{\text{out}}), J_{ij})$$

où  $s_j^{\text{in}}$  est l'activité pré-synaptique,  $s_i^{\text{out}}$  est l'activité post-synaptique,  $m_j^{\text{in}}$  une trace exponentielle de l'activité pré-synaptique, et  $m_i^{\text{out}}$  une trace exponentielle de l'activité post-synaptique. Cette règle peut formellement être reliée aux règles de plasticité d'ordre 1 (voir page 49).

En utilisant la technique de transfert d'échelle temporelle (augmentation du pas de discrétisation d'Euler), il est par ailleurs facile d'établir la correspondance entre la STDP et la règle TD décrite précédemment (Daucé, 2007), comme cela avait déjà été noté par (Rao & Sejnowski, 2001).

**STDP dans les réseaux récurrents aléatoires** L'étude présentée dans (Henry & Daucé, 2008b) (voir annexe H) est motivée par de nombreuses observations sur le rôle des oscillations et de la synchronisation dans le fonctionnement du système nerveux. Les réseaux aléatoires équilibrés offrent, comme on l'a vu (section 6.1), la possibilité de développer des régimes dynamiques plus variés, permettant en particulier la production d'oscillations synchronisées de grande amplitude dans certaines configurations. Il est intéressant dans ce cadre d'étudier comment la plasticité interagit avec la dynamique de population pour produire des comportements nouveaux, en particulier des oscillations synchronisées.

Les modèles de neurones impulsionnels présentés dans (Henry & Daucé, 2008b; Daucé, 2014) montrent, à une échelle temporelle fine, une capacité à produire une bouffée d'activité de population (sur une soixantaine de millisecondes), suivie par un retour à l'état de

base. Les dynamiques contraintes relaxent vers des régime présentant le même degré de chaos que la dynamique spontanée. Seule la répartition des niveaux d'activité varie. Les réponses induites par les motifs ne sont pas qualitativement différentes.

Plusieurs indicateurs de complexité ont été mis en place pour mesurer la nature des dynamiques spontanée et induite. L'analyse du spectre de Fourier ne permet pas d'établir de fréquence caractéristique. L'autocorrélogramme du signal moyen montre une décroissance exponentielle caractéristique d'un bruit filtré. Nous utilisons également un indicateur du nombre de degrés de liberté (reposant sur l'Analyse en Composantes Principales (Wright et al., 2001)) qui montre une dimension intrinsèque du signal de l'ordre de 35 pour un réseau de 100 neurones, ce qui indique une très faible corrélation entre l'activité des différents neurones. Le réseau développe donc un régime de chaos fort.

Pour séparer les réponses, on applique la STDP sous l'influence d'un motif particulier. Selon les gammes de paramètres, deux effets peuvent être observés. Dans certains cas (Henry & Daucé, 2008a), un régime synchronisé est obtenu avec de grandes oscillations et une période de l'ordre de 50 à 100 ms (10-15 Hz), lorsque la dynamique est soutenue par la stimulation (et s'éteint en l'absence de stimulation). Dans ce cas, la dynamique finale s'éteint puis renaît périodiquement sous l'effet de la stimulation externe. Dans d'autres cas, le régime est très proche du régime périodique asynchrone obtenu en dynamique spontanée (Henry & Daucé, 2008b), avec un comportement "quasi synchrone" observable lorsque le réseau est soumis à des motifs binaires.

L'application de la STDP conduit la dynamique vers un régime plus simple. La dimension intrinsèque décroît continument jusqu'à atteindre une valeur faible (entre 2 et 3), qui caractérise l'entrée dans un régime périodique (Henry & Daucé, 2008a; Henry & Daucé, 2008b). Cette activité périodique est maintenue de manière asymptotique, ce qui confirme l'appartenance de la STDP aux règles de plasticité d'ordre 1. La STDP a également un effet régulateur sur le niveau d'activité moyen. Pour différents taux d'activité individuels initiaux (entre 50 et 400 Hz), la STDP conduit le système vers un régime où l'activité individuelle est de l'ordre de 200 Hz (voir figure 7.4). Le caractère régulateur (homéo-

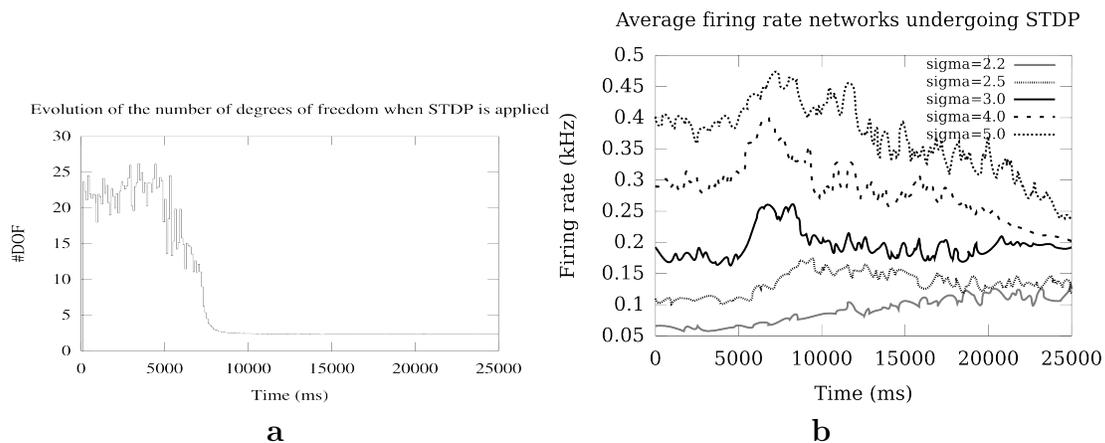


FIGURE 7.4 – **a.** Estimation du nombre de degrés de libertés, par pas de 100 ms, sur 25 secondes (25000 ms) de dynamique. La STDP est appliquée après 5 secondes. **b.** Evolution du taux de décharge moyen sur la même durée. Les différentes traces correspondent à différentes valeurs du couplage moyen (extrait de Henry & Daucé, 2008b).

statique) de la STDP sur le taux de décharge moyen est une caractéristique connue déjà notée par (Kempster et al., 1999).

L'activité finale est caractérisée par une périodicité intrinsèque marquée, de l'ordre de 20 ms (50 Hz). Cette période correspond à environ deux fois le délai moyen (Henry & Daucé, 2008b), avec une même séquence d'activation qui se répète en boucle. L'effet de simplification et de stabilisation d'un régime périodique dans les réseaux développant une activité intrinsèque a été moins souvent noté dans la littérature. Il semble ici être le résultat de la capture de régularités présentes dans la dynamique initiale, qui se distingue donc d'un processus aléatoire stationnaire. La différence entre périodicité globale (50 Hz) et activité individuelle (200 Hz) s'explique par le fait que les neurones tirent en moyenne quatre fois par cycle.

L'activité périodique finale est une activité très robuste. Si l'on interrompt l'apprentissage suffisamment tôt, le comportement du réseau permet de séparer l'activité induite par le (ou les) motif(s) appris de celle induite par les autres motifs (Henry & Daucé, 2008a). En optimisant la durée d'apprentissage, une étude (non publiée) montre une capacité d'encodage similaire à celle des réseaux continus : le nombre maximal de motifs pouvant induire une réponse spécifique est de l'ordre de 5% du nombre de neurones dans le réseau.

D'un point de vue computationnel, il n'y a qualitativement pas de différence entre le nouveau modèle et l'ancien (Daucé et al., 1998). Le comportement induit par la STDP ne semble donc pas reposer sur l'ordre ou l'instant de tir des potentiels d'action, mais plutôt sur leur fréquence de décharge.

**STDP dans les réseaux équilibrés** L'utilisation d'un patron de connectivité plus réaliste a été proposé dans une étude datée de 2009 (Henry & Daucé, 2009), avec séparation fonctionnelle des neurones excitateurs et inhibiteurs, et utilisation de neurones à conductance (Kepler et al., 1992). Cette étude met en évidence un comportement très analogue à celui observé sur le modèle binaire (Daucé, 2007), confirmant la parenté fonctionnelle entre la règle TD et la STDP (voir aussi (Rao & Sejnowski, 2001)). La STDP est appliquée uniquement sur les connexions entre neurones excitateurs. Une convergence vers un régime d'oscillations synchronisées, avec une période globale de l'ordre de 25 Hz (et une fréquence moyenne individuelle de l'ordre de 200 Hz), est mise en évidence. Il s'agit d'une synchronisation observable au niveau de la population, les activités individuelles restant relativement irrégulières et hétérogènes. Une analyse de la plasticité en fonction des délais met en évidence la potentiation des délais longs ( $> 8$  ms) et une dépression des délais intermédiaires (4-8 ms) qui favorise l'entretien du comportement synchronisé (non publié).

**STDP et action de population** Une dernière étude, publiée récemment dans les proceedings d'ESANN (Daucé, 2014) (voir annexe M), étudie l'effet de la STDP dans un réseau équilibré sous influence d'un signal spatio-temporel. Deux cas sont considérés : signal périodique (se répétant toutes les 30 ms) ou signal non-périodique. L'activité spontanée sous influence de ces deux types de signaux est qualitativement la même initialement. La plasticité est appliquée dans les deux situations. Lorsque le stimulus est aperiodique, aucun changement qualitatif n'est observé. Lorsque le stimulus est périodique, l'activité devient plus régulière, avec de grandes oscillations (période 16 Hz) observables au niveau

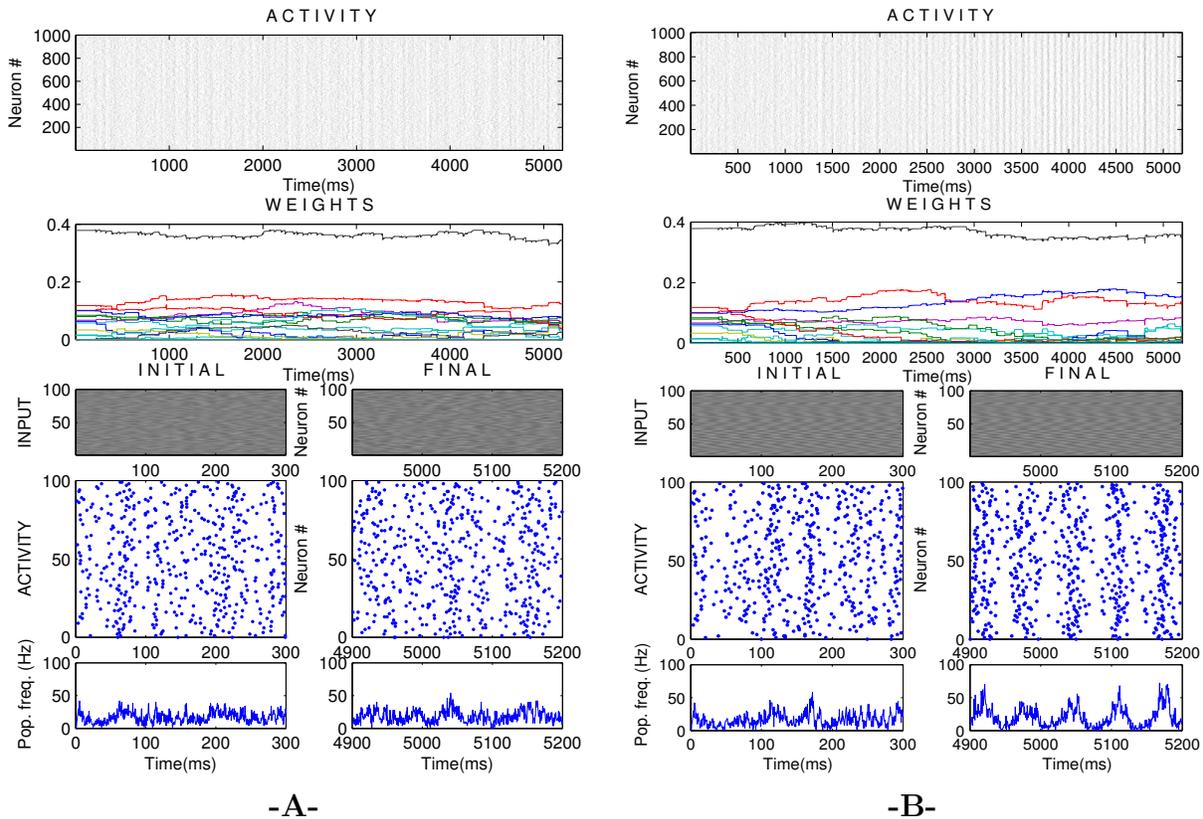


FIGURE 7.5 – **Activité de la population excitatrice durant la STDP.** **A.** Entrée non-périodique. **B.** Entrée périodique. De haut en bas : la première figure donne l’activité de tous les neurones de la couche excitatrice, sur 5200 ms ; la seconde donne l’évolution, sur la même durée, de quelques poids synaptiques ; la troisième donne l’aspect du signal d’entrée, en niveau de gris, pour  $t \in [0, 300]$  ms et  $t \in [4900, 5200]$  ms ; la quatrième donne l’activité de 100 neurones, sur les mêmes intervalles ; la cinquième donne le taux de décharge moyen en Hz, sur les mêmes intervalles (extrait de Daucé, 2014).

de la population (voir figure 7.5). Le comportement est assez proche du comportement de grandes oscillations lentes obtenu sous influence statique. On observe néanmoins une synchronisation des oscillations du réseau avec la période du stimulus, autrement dit il y a un accrochage de phase de sorte que la période interne (60 ms) est exactement le double de la période du signal (30 ms). Ce comportement d’accrochage est assez comparable à celui observé sur les réseaux plus simples. Le réseau devient “réactif” aux motifs périodiques appris, développant une “action de population” qui est un analogue global des potentiels d’action neuronaux.

Une sensibilité acquise à des patrons spatio-temporels spécifiques, telles qu’observée dans les modèles fréquentiels (voir paragraphes précédents), peut donc être obtenue à l’échelle de quelques millisecondes, mettant en jeu un faible nombre de potentiels d’action par neurone. Le même mécanisme peut donc être identifiés à différentes échelles, les opérations à échelle temporelle courte étant elles-même possiblement “portées” (ou embarquées) dans des opérations à échelle temporelle plus larges.

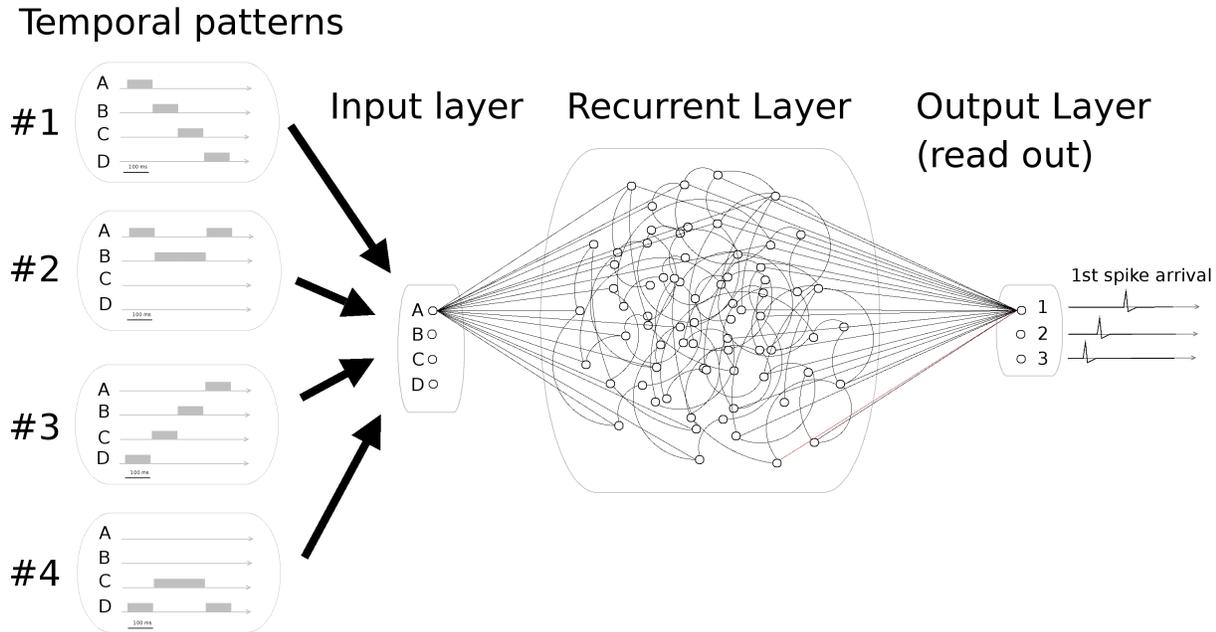


FIGURE 7.6 – **Dispositif expérimental.**  $N = 4$  signaux spatio-temporels sont présentés au réseau afin d’être classés en  $K = 3$  catégories. Le réseau est composé de 3 populations. La couche d’entrée est composée de 4 neurones (étiquetés A, B, C et D). Les connexions d’entrée suivent une loi Gaussienne de moyenne 0 et d’écart-type 0,04. La couche cachée contient 100 neurones complètement connectés. Les poids synaptiques de la couche récurrente suivent une loi Gaussienne de moyenne 0 et d’écart-type 0,02. La couche de sortie est composée de 3 neurones, avec des connexions latérales inhibitrices (non représentées). Les poids de sortie suivent une loi Gaussienne de moyenne 0,09 et d’écart-type 0,01 (extrait de Henry et al., 2007).

### 7.2.4 Discrimination de séquences spatio-temporelles

Les protocoles d’apprentissage présentés jusqu’à présent avaient pour but de former une réponse induite. La forme de la réponse n’était pas fixée à priori et dépendait des caractéristiques du substrat. Il n’y avait pas de “read-out”, mais une simple résonance entre le signal d’entrée et la dynamique développée en interne.

Dans une étude antérieure (Henry et al., 2007) (voir annexe F), nous avons également considéré une architecture à entrées-sorties complète :

- Le réseau possède une couche d’entrée qui transmet passivement le signal, une couche récurrente aléatoire, et une couche de sortie constituée dans ce cas de quelques neurones dont la sortie est lue.
- La forme de la réponse induite est imposée par l’expérimentateur, de manière supervisée ou semi-supervisée.

La tâche consiste à discriminer quatre motifs spatio-temporels. Ces motifs possèdent des éléments en communs, de telles sorte qu’il est impossible de séparer les signaux sur la base de l’entrée instantanée. Chaque neurone de sortie représente une catégorie. La réponse est fixée par le neurone qui émet le premier potentiel d’action suite à la présentation du motif (voir figure 7.6).

Le protocole d’apprentissage consiste à présenter les séquences périodiquement, lire les réponses produites, informer le réseau du caractère correct ou incorrect de sa réponse et

appliquer une règle de plasticité visant à augmenter le taux de réponses correctes, selon un principe qui se rapproche de l'apprentissage par renforcement.

L'apprentissage utilise les effets symétriques des règles Hebbienne et anti-Hebbienne sur la dynamique des réseaux récurrents aléatoires, montrés dans (Daucé et al., 2005). Par exemple, l'anti-STDP (règle miroir de la STDP) a pour effet d'augmenter la complexité de la dynamique intrinsèque, soit un effet contraire à la STDP. L'application successive de la STDP puis de l'anti-STDP sur un réseau récurrent aléatoire a pour effet de réduire la complexité puis de l'augmenter à nouveau.

L'étude est également une implémentation originale du codage par rang, utilisant un principe concurrent d'avance vs. retard de la réponse. Si la réponse est correcte, la STDP est appliquée sur la couche de sortie et sur la couche récurrente, ce qui tend à anticiper l'instant de réponse du neurone ayant tiré. Si la réponse est incorrecte, c'est l'anti-STDP qui est appliquée, tendant à retarder la réponse. L'article (Henry et al., 2007) utilise des signaux temporels très lents, où la séquence est composée de 4 motifs durant chacun 100 ms, de sorte qu'une séquence complète dure 400 ms. Dans ce cas, le temps nécessaire pour atteindre des taux de réussite raisonnables s'exprime en heures (environ 100000 essais). Les résultats présentés dans (Daucé & Henry, 2006) utilisent des séquences plus courtes (40 ms). Dans ce cas, environ 500 essais suffisent pour obtenir une discrimination parfaite.

La différence de taux d'apprentissage et de taux de réussite s'explique principalement par la différence d'échelle temporelle des séquences à apprendre. La mémoire d'une unité neuronale est de l'ordre de sa constante de membrane, soit 10 à 20 ms. Néanmoins, à la lumière des résultats du chapitre précédent, la plasticité semble capable d'induire suffisamment de régularité dans la dynamique intrinsèque pour conserver une mémoire plus importante au sein de la population. La difficulté à discriminer les séquences de 400 ms semble néanmoins fixer la limite supérieure de cette capacité (100 ms étant en pratique suffisant pour discriminer deux séquences).

# 8

## Architectures de contrôle

Ce chapitre aborde la question plus globale de la formation des comportements moteurs. Nous commençons par rappeler certains principes et problèmes d'ordre général, comme la différence d'extensivité et d'expressivité du contrôleur et de son environnement. Nous présentons ensuite un certain nombre de résultats ayant en commun le principe du contrôle en "boucle fermée". Nous abordons ainsi successivement la question de la stabilisation et de la déstabilisation des patrons d'interaction vus comme des transitions entre activité autonome et activité commandée, et la question de l'activité autonome (chaos, bruits "structurés") pour l'implémentation des mécanismes d'exploration dans le cadre de l'apprentissage par renforcement, sur divers substrats neuromimétiques. De manière plus appliquée, nous détaillons les principes d'une architecture de contrôle destinée à compléter des déficiences motrices à partir de l'analyse de l'EEG de surface (interfaces cerveau-machines non-invasives).

### 8.1 Contrôleur et environnement

L'approche dite "située" en robotique mobile et en intelligence artificielle (Brooks, 1991; Varela et al., 1991) prône la prise en compte de la contrainte du corps comme élément structurant du développement de compétences cognitives plus élaborées. Cette contrainte, qui s'exprime par des contingences sensori-motrices, et la prise en compte des disparités d'échelle, augmente néanmoins la complexité des tâches d'apprentissage et de leur implémentation par plasticité synaptique.

Le cadre que nous considérons ici est celui de la commande motrice. Nous considérons un programme doté d'effecteurs moteurs capables de déplacer et faire pivoter des masses articulées dans l'espace physique. L'espace physique est décrit, d'une part, par ces effecteurs directement contrôlables, ainsi que d'autres éléments physiques non directement contrôlables.

- Le programme est le *contrôleur*.
- L'espace physique est l'*environnement*.

L'interaction d'un contrôleur avec son environnement passe par des transducteurs, qui sont les canaux de communication entre les deux milieux : appareils sensoriels et proprioceptifs chez les animaux, et contractions musculaires permettant de déplacer des masses articulées ; senseurs et actionneurs pour les dispositifs de contrôle artificiels : capteurs qui

traduisent des signaux physiques en signaux électriques, moteurs, joints et pistons pour les déplacements de masses.

Le cadre que nous avons regardé jusqu'à présent est l'apprentissage en "boucle ouverte", dans lequel les actions physiques sur le milieu sont découplées des sensations. La prise en compte plus réaliste des deux domaines matériels que sont le circuit logiciel et le milieu physique environnant conduit à déplacer le cadre conceptuel de l'apprentissage vers l'apprentissage en "boucle fermée". Il se définit formellement par la présence de corrélations entre les actions ou les réponses produites et les manifestations sensorielles suivantes.

### Production motrice

Au cerveau traiteur de signal, tel qu'il est classiquement étudié en neurosciences computationnelles, s'oppose de manière symétrique le cerveau *émetteur* de signal. L'activité du cerveau de l'animal éveillé consiste en grande partie à produire des ordres moteurs qui se traduisent directement au sein de l'environnement par des déplacements de masses, des rotations d'articulations, des déformations etc. résumés par l'équation (3.2), page 36, où l'environnement est le canal de transmission et de traitement de l'information motrice émise par le cerveau. Cette description purement formelle de l'activité motrice comme signal est bien sûre peu habituelle pour les activités motrices élémentaires (orientation locomotion, etc.) mais se comprend facilement si on considère l'activité motrice consistant à produire des sons articulés chez l'homme ou plus généralement tout type de communication animale.

Le fait de considérer le cerveau comme simple traiteur de signal, ou symétriquement comme simple émetteur de signal, n'est bien sûr pas satisfaisant, le seul niveau de description adéquat étant celui de la boucle de rétroaction formé par le contrôleur neuronal et son environnement, autrement dit le *couplage* entre,

- d'une part, l'appareillage moteur et son environnement (eq. (3.2), page 36)
- d'autre part l'appareillage sensoriel et le contrôleur (eq. (3.3), page 36)

La prise en compte de la boucle de rétroaction illustre le caractère profondément imbriqué des activités motrices et perceptives, rendant difficile l'identification d'une causalité (d'une "source") autre que la relation agent-environnement elle-même<sup>1</sup>.

### Circuits neuronaux de la commande motrice

La recherche de principes généraux à l'apprentissage moteur en biologie se heurte donc à un certain nombre de difficultés liées au caractère extrêmement complexe et imbriqués des circuits biologiques, et au caractère multi-tâches du cerveau.

Le cadre général défini jusqu'à présent permet néanmoins de fixer quelques lignes directrices.

1. Nous mettons en avant le rôle de l'activité endogène dans la production motrice, et particulièrement l'exploration motrice dans le cadre de l'apprentissage.
2. Nous cherchons à identifier des mécanismes d'appariement sensori-moteurs compatibles avec le paradigme de l'apprentissage Hebbien et de la formation d'assemblées

---

1. Ce paradoxe de l'origine a été noté de longue date par l'école cybernétique (Dupuy & Varela, 1992).

neuronales. Plus généralement, dans le cadre des systèmes dynamiques couplés (voir les équations (3.5)-(3.8), page 38), nous regardons le principe de l'appariement conditionnel vu page 23.

3. Nous cherchons également à établir un lien plus général entre apprentissage moteur et reconnaissance de formes. Nous développons des modèles séquentiels dans lesquels la règle de plasticité s'apparente à un apprentissage “par coeur” de la fonction de réponse du contrôleur  $\phi_{\text{out}}^{-1}(\mathbf{x}^*)$  (voir page 43).

### Rôle de l'activité endogène dans la production motrice

Les réseaux de neurones récurrents aléatoires offrent, comme nous l'avons vu, un cadre propice à l'étude des couplages entre un contrôleur neuronal et un environnement physique. Dans le cadre de la production des ordres moteurs, l'activité récurrente endogène, telle qu'elle se développe par exemple dans les réseaux récurrents aléatoires, apparaît propre à susciter des commandes relativement indépendantes du contexte sensoriel.

Ici, notre hypothèse de travail est simplement que *l'activité endogène est le mécanisme principal permettant à un agent d'enrichir son répertoire de comportements*. L'activité endogène spontanée, telle qu'observée dans les réseaux récurrents aléatoires, constitue un “réservoir” de comportements moteurs dont la réalisation peut être facilitée ou supprimée via des mécanismes de renforcement, mais aussi via des aspects contingents comme l'historique des couplages sensori-moteurs entre l'agent et son environnement, en particulier, dans le cadre de tâches sous-spécifiées (ou non-convexes) comme les déplacements de masses articulées.

### Apprentissage par renforcement

Le lien entre ces mécanismes d'apprentissage Hebbiens et les méthodes acteur-critique classiques n'est pas direct. Dans les méthodes classiques, le choix de l'action (“actor”) est découplé de l'évaluation des conséquences du choix de l'action (“critic”). Chaque situation sensorielle se voit attribuer une valeur en relation avec les bénéfices futurs attendus. Dans le cas que nous considérons (réseaux de neurones aléatoires), la critique n'est pas implémentée. Le réseau se contente d'explorer l'espace des réponses associées à différents contextes. Ce cadre se rapproche formellement d'une version simplifiée de l'apprentissage par renforcement appelée “Direct policy gradient” (Williams, 1992; Bartlett & Baxter, 1999), où la réponse est construite par ajustement direct des paramètres de la politique (la règle qui détermine le choix entre les différentes réponses possibles). L'amélioration de la politique au cours du temps passe par un principe de descente de gradient guidé par les récompenses (c'est-à-dire qui vise à augmenter l'espérance de la récompense au cours du temps). La traduction de ces principes sous la forme d'une règle de plasticité conduit à une règle qui s'apparente à la règle de Hebb (Florian, 2006).

Dans le cadre des réseaux de neurones récurrents aléatoires :

- D'une part, les propriétés de criticalité et le caractère chaotique de la dynamique contrainte suggèrent une capacité à explorer différents répertoires comportementaux au cours d'un historique d'interactions avec un environnement.

- D'autre part, le caractère simplificateur (réducteur) de la règle de Hebb sur la dynamique interne est favorable à la stabilisation de couplages sensori-moteurs dans l'espace de la tâche.

Le principe de l'apprentissage par renforcement suggère également de prendre en compte (et produire les modifications synaptiques) liées à la totalité de la trajectoire dans le cas de récompenses ponctuelles (problème du “temporal credit assignment” (Sutton & Barto, 1998)). L'utilisation de récompenses ponctuelles nécessite ainsi de mettre en place un mécanisme de *trace synaptique*, c'est-à-dire une mémoire locale des co-activations récentes pour appliquer la modification synaptique, dans un sens Hebbien ou anti-Hebbien selon les cas.

**Dilemme exploration exploitation.** Le dilemme exploration/exploitation s'exprime, sous sa forme la plus simple, comme un problème d'échantillonnage aléatoire de l'espace des choix (problème dit du “bandit manchot”). Chaque tirage apporte une information qu'il convient d'exploiter au mieux, en établissant un compromis entre recherche d'information (réduction d'incertitude) et exploitation effective de l'information (pour maximiser le gain). Cette forme d'apprentissage séquentiel établit des modèles qui visent à réduire la différence entre les gains obtenus par l'algorithme et les gains attendus en toute connaissance de cause. L'apprentissage séquentiel statistique est le socle d'une famille de méthodes d'apprentissage fondée sur l'échantillonnage de l'environnement, comme le Q-learning ou le TD-learning dans le cadre de l'apprentissage par renforcement classique. Cette approche a renouvelé le domaine, en proposant par exemple des politiques non-stochastiques (algorithme UCB), ou en étendant le modèle au cas “adversarial” (théorie des jeux) (Auer et al., 2002).

**Temporal credit assignment** Cette importance de la commande aléatoire dans l'acquisition de connaissance a été formalisé dans le cadre de *l'apprentissage par renforcement* (Sutton & Barto, 1998) Cette approche repose sur la notion de signal de renforcement, qui est un “méta signal” dont le rôle est de transformer le manière dont le logiciel traite le signal, selon des modalités variées comme la consommation énergétique, l'exposition aux signaux nociceptifs ou bénéfiques, ou encore la réduction d'incertitude. Si le milieu contient des éléments nocifs ou aversifs par exemple, l'amélioration peut passer par une augmentation de la sensibilité logicielle à une classe de signaux associés à (ou précurseurs de) la sensation aversive, et à une facilitation des comportements permettant de mieux éviter les situations où le signal aversif est présent. La plupart des modèles de renforcement reposent sur des tables de correspondance (“Look-up table”), adaptées aux environnements et aux choix d'action discrets, mais souffrent de problèmes d'explosion combinatoire dans des environnements continus et/ou complexes.

Une séance d'apprentissage consiste à effectuer une série de “choix” qui orientent le cours futur de la séance. La tâche n'est plus définie par rapport à la justesse des réponses instantanées, mais par rapport à la justesse de ce parcours dans l'espace des trajectoires possibles.

Les signaux de renforcement (méta-signaux) servent à “sculpter” le programme au cours des interactions entre le logiciel et son environnement. Le programme qui résulte de ces interactions varie selon l'expressivité du réseau sous-jacent (la structure du circuit

logique), le caractère déterministe ou stochastique des politiques d'exploration et enfin le caractère convexe ou non de la tâche. Le choix entre les réponses peut être départagé selon l'espérance de récompense associée à chacune d'elles, mais aussi selon le degré d'incertitude associé, certaines réponses mal connues (ou certaines sources et/ou chemins mal modélisés) pouvant potentiellement apporter des récompenses supérieures. C'est le fameux dilemme "exploration/exploitation".

### Plasticité induite

Dans le cadre du contrôle moteur, l'apprentissage vise à produire des régularités dans cette dynamique afin de stabiliser certaines réponses motrices. Le but de la plasticité est de favoriser les réponses liées à un renforcement positif (apprentissage "Hebbien"), et de défavoriser les réponses liées à un renforcement négatif (apprentissage "anti-Hebbien"). On parle de plasticité "induite".

La plasticité induite est un mécanisme de plasticité déclenché par un signal extérieur. Il peut s'agir, dans un cadre biologique, d'un neurotransmetteur ou d'une entrée synaptique spécifique, i.e. :

$$\dot{J}_{ij} = F(\mathbf{S}_i(t), \mathbf{S}_j(t), \mathbf{y}(t), J_{ij}) \quad (8.1)$$

où  $\mathbf{y}(t)$  est un signal extérieur qui vient "moduler" la plasticité.

On notera que l'application alternée d'une plasticité Hebbienne et anti-Hebbienne ouvre la voie à l'implémentation de mécanismes d'apprentissage par renforcement, ainsi qu'à la modélisation des mécanismes de plasticité liés à la récompense et à la punition dans le cerveau (Barto, 1995; Schultz et al., 1997; Gurney et al., 2001).

### Apprendre et oublier dans les environnements non stationnaires

Les modèles probabiliste de l'apprentissage reposent souvent sur une hypothèse de stationnarité qui suppose que les données d'entrée sont issues d'un processus stochastique dont les caractéristiques ne varient pas au cours du temps. Cette hypothèse n'est pas vérifiée dans le cas du contrôle en boucle fermée, où non seulement le programme mais également l'environnement voient leurs caractéristiques évoluer au cours du temps.

La non-stationnarité est une caractéristique d'environnements eux-mêmes soumis à des transformations, soit par l'action du contrôleur, soit par des actions extérieures (sans rapport avec le contrôleur). L'interaction entre un logiciel et un humain produit par exemple un historique d'interactions, correspondant généralement à des étapes de "calibration" (côté logiciel) et à une "prise en main" de l'interface (côté l'humain).

Ainsi les régularités observées durant les premières phases de l'apprentissage ne sont plus nécessairement présentes à une étape ultérieure de l'apprentissage : certaines trajectoires sont "abandonnées". Dans une perspective de parcimonie (ou de ressource limitée), il est avantageux d'oublier certains faits passés pour mieux appréhender les faits nouveaux, autrement dit d'oublier pour mieux apprendre. Cette prise en compte du "vieillessement" de l'environnement conduit à accorder plus de crédit à des observations récentes qu'à des observations anciennes (Kivinen et al., 2004). L'apprentissage "en ligne" (Le Cun & Bottou, 2004) apporte un cadre d'analyse naturel au traitement de la non stationnarité,

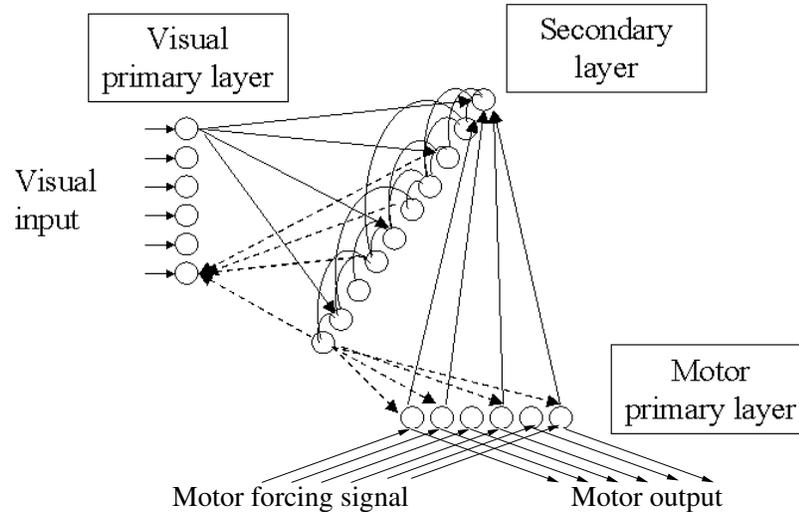


FIGURE 8.1 – **Architecture de contrôle.** La couche primaire est séparée en deux couches perceptives : couche visuelle primaire et couche sensori-motrice primaire. L'intégration des informations visuelles et sensori-motrices est réalisée par la couche récurrente (extrait de Dauce et al., 2002).

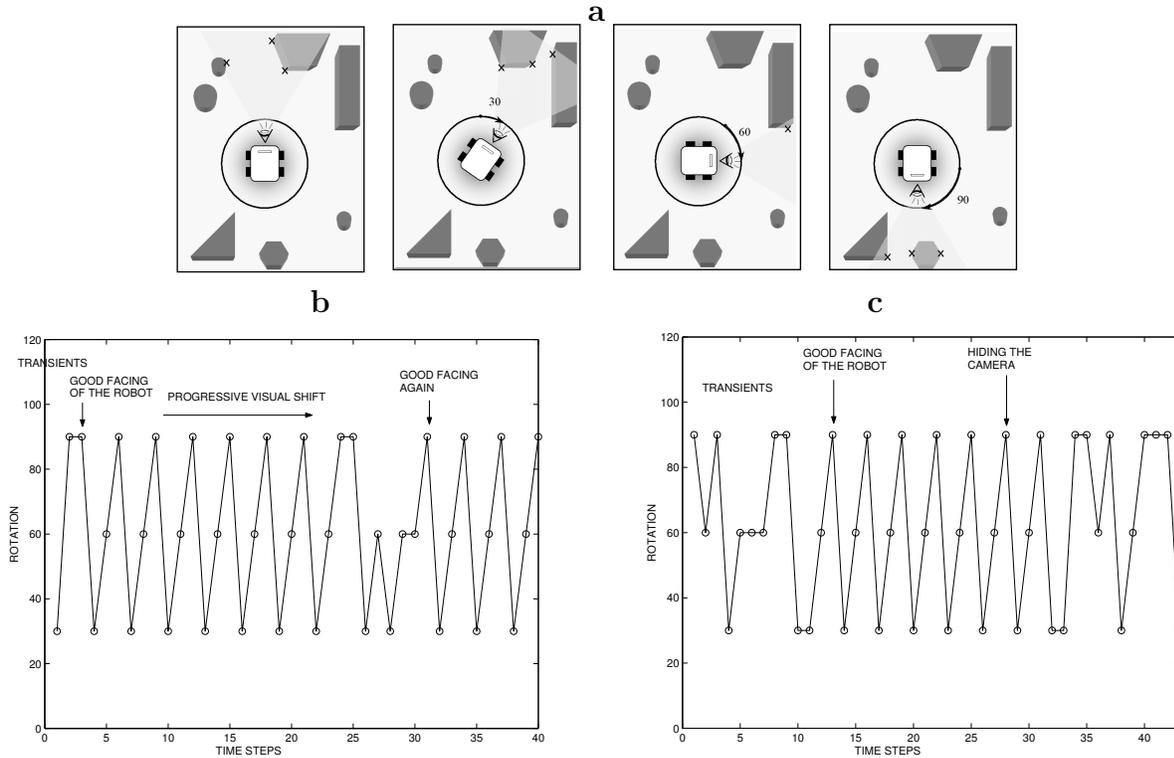
l'ajustement permanent du contrôleur lui permettant en effet de suivre les changements environnementaux en mettant à jour son modèle.

## 8.2 Contributions personnelles

### 8.2.1 Alternance comportementale en robotique autonome

Dans le cadre d'un court séjour de fin de thèse (juillet 1999), j'ai participé à une étude sur l'apprentissage moteur sur une plate-forme robotique réelle au sein de l'équipe de neurocybernetique dirigée par Philippe Gaussier au laboratoire ETIS (Cergy-Pontoise). L'expérience a été conduite sur la plate-forme de contrôle robotique PerAc (Gaussier & Zrehen, 1995), et publiée dans (Quoy & Dauce, 2000; Dauce & Quoy, 2000; Quoy et al., 2001; Dauce et al., 2002) (voir annexe C).

L'architecture de contrôle est constituée de trois couches : deux couches primaires (une couche visuelle et une couche motrice) et une couche associative (voir figure 8.1). L'apprentissage reposait sur l'imposition sur la couche motrice d'une commande périodique constituée d'une séquence de trois déplacements angulaires distincts répétés indéfiniment. A l'issue de l'apprentissage, les liens de feedback positif de la couche associative vers les couches primaires permettent au robot de produire une commande motrice (périodique) conforme à la séquence sensori-motrice observée pendant l'apprentissage (autrement dit de prédire son entrée visuelle en même temps que sa commande motrice). Les entrées sensorielles et motrices sont contrôlées afin que la séquence visuelle soit bien prévisible pendant l'apprentissage. Pendant le test, on laisse de petits décalages angulaires liés aux frottements se produire, introduisant un décalage progressif de la scène visuelle, et une perte de correspondance entre la prédiction et l'observation. On observe alors, au niveau comportemental, une alternance claire entre un comportement périodique, correspondant



**FIGURE 8.2 – Mouvements du robot après apprentissage.** **a** : Positions successives du robot durant une séquence motrice ( $+30^\circ, +60^\circ, +90^\circ$ ). La robot est placé dans un environnement ouvert (ce n'est pas une simulation). L'association entre un jeu d'amers visuels (points de contraste élevé, notés par un 'x' sur la figure) et leur position angulaire constitue l'entrée visuelle. Après trois mouvements, le robot a fait un demi-tour (la répétition de ces trois mouvements ramène le robot sur sa position initiale). **b** : Exemple de recalage moteur suite à un décalage visuel. Les deux premiers pas de temps correspondent à la dynamique de relaxation. Les mouvements suivants correspondent à la synchronisation de la séquence périodique apprise ( $+30^\circ, +60^\circ, +90^\circ$ ) sur la séquence visuelle apprise. En raison de frottements au niveau du sol, le mouvement effectif, qui ne correspond pas exactement à la commande, conduit à une dissociation progressive entre la scène visuelle et la séquence de mouvements. Un changement de comportement est observé à  $t = 25$ , correspondant au décrochage entre le flux visuel et les mouvements associés. Finalement, après de nouveaux transitoires, le robot trouve à nouveau une bonne correspondance entre le mouvement et la vision et réitère sa séquence périodique. **c** : Après que le robot a atteint son comportement périodique, l'entrée visuelle est masquée. L'absence d'information visuelle conduit rapidement le robot vers un comportement "chaotique" (d'après Daucé et al., 2002).

à la tâche apprise, et un comportement irrégulier sans correspondance avec la tâche (voir figure 8.2). Les rotations irrégulières induisent une succession de signaux visuels qui, par hasard, se mettent à correspondre à la séquence apprise, et font basculer l'ensemble dans le régime périodique où les entrées visuelles redeviennent prévisibles.

Cette expérience robotique offre un exemple clair de couplage entre une activité mécanique (des déplacements) et une activité logicielle (un réseau de neurones récurrent). Une activité neuronale périodique couplée avec une activité mécanique périodique et/ou une activité neuronale irrégulière couplée avec une activité mécanique irrégulière. La commande ne repose pas sur la correction d'erreur, comme dans le contrôle traditionnel, mais sur

- l'appariement de séquences spatio-temporelles (pattern matching)

- et sur la persistance de cet appariement pour des versions dégradées du signal appris.

L'erreur (la non-correspondance) ne produit pas une correction mais une transition vers une activité irrégulière qui s'apparente à une exploration aléatoire de l'espace des couplages sensori-moteurs jusqu'à atteindre un nouvel attracteur comportemental.

Une interprétation possible de ces résultats du point de vue de l'approche dynamique de la cognition artificielle a été proposée dans le livre que j'ai co-édité avec Agnès Guillot en 2002 (Guillot & Daucé, 2002), selon le principe du *passage d'échelle* (voir discussion page 59).

### 8.2.2 Contrôle adaptatif en boucle fermée

L'apprentissage en boucle fermée a été tenté sur des architectures analogues à (Daucé et al., 2002) (couches primaires vs. couche associative récurrente) dans des environnements variés à états continus (robotiques simulée, pendule inverse, ...), dans lesquels la commande motrice est issue de l'activité de la (ou des) couche(s) associative(s).

- Les changements comportementaux reposent sur l'activité de la couche associative, soumise à une influence interne (l'activité intrinsèque) et une influence externe (signal en provenance de la couche sensorielle). En l'absence de consigne, la commande motrice doit en effet être le produit de l'activité interne.
- Le caractère irrégulier de l'activité initiale doit traduire l'absence de direction privilégiée de la réponse motrice. Le caractère organisateur des règles Hebbiennes était censée induire un comportement plus prédictible au fil du temps. La prise en compte du signal de renforcement dans la règle de plasticité était censée augmenter l'adéquation des comportements à la contrainte exprimée par le signal.

Cette tâche s'est révélée plus difficile à implémenter que la précédente :

- Dans le cas précédent, la séquence à apprendre (le répertoire comportemental) était fixée par le concepteur. Ici, le répertoire comportemental est supposé émerger des interactions avec l'environnement.
- Une autre facilité dans le cas précédent était que le concepteur stoppait la séance d'apprentissage lorsqu'il estimait que le réseau avait suffisamment appris. Dans le cas de l'apprentissage par renforcement, la plasticité est appliquée de manière continue, ce qui peut conduire à des problèmes de surapprentissage (et de résistance au changement de la dynamique interne qui rend le réseau insensible à son environnement).
- De plus, un environnement continu classique (contrairement à un environnement discret) présente peu d'occasions d'interactions périodiques avec l'environnement. Dans un contexte d'environnement continu, la dynamique externe est lente par rapport à la dynamique neuronale.

Dans ce cadre,

- Les règles Hebbiennes d'ordre 0 (fondées sur la conjonction d'activité) se sont révélées plus efficaces pour la résolution des tâches, bien que moins propices aux effets de synchronisation et de résonance.

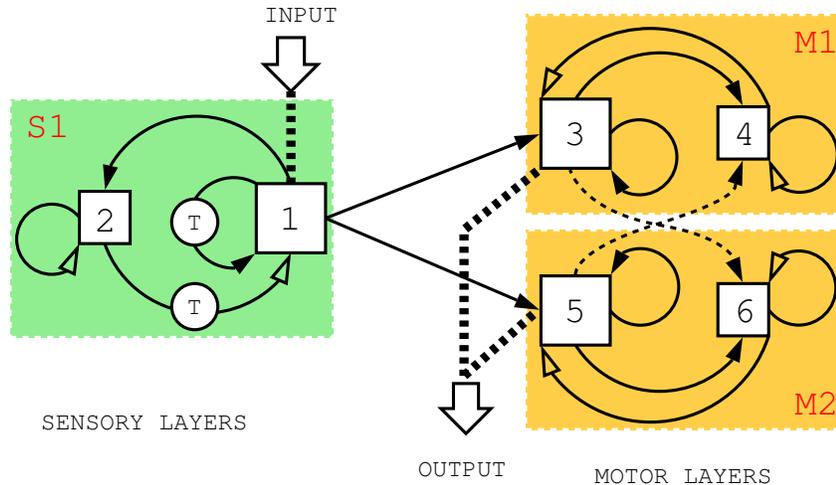


FIGURE 8.3 – **Architecture du réseau perception-action.** Le réseau est composé de 6 populations. Les flèches pleines représentent les connexions excitatrices. Les flèches évidées représentent les connexions inhibitrices. Un module fonctionnel est composé de deux populations fortement interconnectées, l’une excitatrice, l’autre inhibitrice. Le module sensoriel **S1** est composé des populations 1 et 2, le module moteur **M1** des populations 3 et 4, et le module moteur **M2** des populations 5 et 6. Le module S1 présente un patron de connexions topologiquement organisé (liens mentionnés avec un “T”). Les 3 modules sont interconnectés par des liens excitateurs (les neurones inhibiteurs agissant seulement localement). Le module S1 envoie un signal (excitateur) vers les modules M1 et M2. Les modules M1 et M2 peuvent s’inhiber mutuellement en excitant les neurones inhibiteurs de leur voisin (lignes pointillées). Ces interactions latérales sont initialisées à zéro (d’après Daucé, 2004a).

- Les problèmes de surapprentissage ont été résolus par l’utilisation d’une règle qui bloque la plasticité en cas d’influence exclusive d’une population afférente. La plasticité ne se développe que sur les neurones soumis à plusieurs influences, interne ou externe. Le terme limitatif bloque la plasticité lorsqu’une classe de liens afférents atteint à elle seule le seuil d’activation du neurone.

### Modèle de 2004

Le premier article présentant une application des problèmes de renforcement dans le cadre de réseaux récurrents aléatoires a été publié dans les proceedings de la conférence SAB’04 (Simulation of Adaptive Behaviors) (Daucé, 2004a) (voir annexe E). L’architecture neuronale repose sur 3 modules : un module sensoriel et deux modules moteurs. Chaque module est constitué de deux populations de neurones : une population excitatrice et une population inhibitrice (voir figure 8.3). L’environnement est un pendule inversé. La tâche consiste à transmettre au pendule l’effet de deux forces opposées produites par le niveau d’activité des deux couches motrices. Le module perceptif reçoit la position angulaire courante en codage topographique (rétine 1D). La couche sensorielle est organisée topographiquement comme dans (Daucé, 2004b). Le régime dynamique, analogue à (Daucé, 2004b), est une activité de type neural field avec oscillations synchronisées. Contrairement à (Daucé et al., 2002), il n’y a pas de couche associative, et pas de liens de feedback de la couche motrice vers la couche sensorielle. L’apprentissage repose sur l’application d’une règle Hebbienne fondée sur la conjonction d’activité simple (d’ordre 0). La plasticité est

appliquée sur les liens excitateurs uniquement. Deux chemins de renforcement sont définis : un chemin de renforcement positif : liens excitateurs feed-forward et liens excitateurs vers la population inhibitrice contralatérale, et un chemin de renforcement négatif : liens excitateurs vers la population inhibitrice ipsilatérale. Le signal de renforcement est calculé à partir de la position angulaire et la vitesse angulaire courante du pendule. Les différents essais sont initialisés avec des conditions initiales variées. Après plusieurs essais (de l'ordre de 300), le système parvient à contrôler le pendule dans le plupart des cas.

L'intérêt de ce modèle est qu'il repose sur un principe de ciblage de la plasticité sur certaines catégories de liens qui est cohérent avec le principe biologique de libération de neurotransmetteurs ciblant des synapses différentes dans le cas de signaux de renforcement positifs et négatifs (Daw & Doya, 2006).

### Modèle de 2007

Un second modèle, présenté dans (Daucé, 2007) (voir annexe G), implémente cette fois-ci le principe d'une application alternée de règles Hebbiennes (resp. anti-Hebbiennes) pour les récompenses positives (resp. négatives), conformément aux idées exprimées dans (Daucé et al., 2005), sur les liens excitateurs uniquement. L'architecture comprend cette fois-ci deux modules : un module perceptif et un module moteur, constitués chacun d'une population excitatrice et d'une population inhibitrice. Une particularité de ce modèle est l'influence inhibitrice du module moteur vers le module perceptif qui a pour effet d'inhiber l'entrée sensorielle lorsque la couche motrice est active. Des résultats comparables au cas précédent sont obtenus, avec un signal de renforcement plus simple reposant uniquement sur la vitesse angulaire.

Dans ces deux modèles, l'exploration des réponses possibles repose sur la dynamique intrinsèque qui s'avère suffisamment variée pour explorer différents types de réponses et stabiliser celles qui conviennent le mieux. La dynamique chaotique joue donc ici principalement le rôle d'un bruit d'exploration. Les comportements obtenus sont principalement réactifs, il n'y a pas de mise en évidence de résonance et/ou de persévérance entre certains couplages sensori-moteurs et la dynamique interne.

### 8.2.3 Codage topographique et renforcement

Deux études portant sur des mécanismes d'apprentissage par renforcement dans des architectures neuronales ont été publiées dans le cadre de l'étude du codage topographique dans l'ANR MAPS (voir section 1.2, page 11).

#### Modèle de 2009

La première étude, publiée dans (Daucé, 2009) (voir annexe I), est une extension des architectures présentées dans la section précédente. La tâche est l'apprentissage d'un comportement de poursuite oculaire dans un environnement en forme de tore (de "donut"). L'environnement contient une cible visuelle effectuant un mouvement circulaire simple. Deux populations de neurones à impulsions sont utilisées. Le signal de renforcement repose

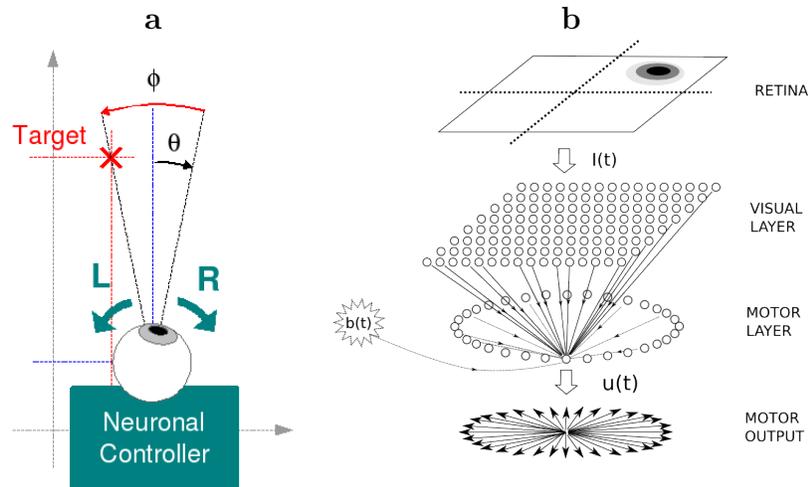


FIGURE 8.4 – **a : Environnement** : Un globe oculaire doté d’un mouvement de rotation poursuit une cible ponctuelle (croix) :  $\theta$  est la position courante de l’œil,  $\phi$  est la direction subjective de la cible (relativement à la direction de la pupille). **b : Contrôleur** : la couche visuelle est constituée de 256 neurones envoyant des axones excitateurs vers une couche motrice constituée de 32 neurones. Les neurones prémoteurs s’inhibent mutuellement. Un bruit blanc filtré  $b(t)$  est ajouté sur chaque neurone prémoteur.  $I(t)$  est l’entrée visuelle, et  $u(t)$  est la commande (d’après Daucé, 2009).

sur la distance entre l’orientation visuelle courante et la position de la cible. La couche primaire est une rétine avec une déformation “fovéale”, sur la quelle la position de la cible est codée sous la forme d’une “bulle” d’activité. La deuxième couche est une couche motrice primaire constituée de 32 neurones, chaque neurone codant pour une direction différente (voir figure 8.4). La commande finale est la composition de ces commandes individuelles. La principale différence avec le dernier modèle de la section précédente est l’absence d’activité intrinsèque dans la couche secondaire. Celle-ci est remplacée par un bruit exploratoire conforme à la théorie du Policy gradient (Williams, 1992).

La règle de Policy Gradient (voir section 8.1) a été implémentée dans un cadre neuronal réaliste (neurones impulsionnels), à partir du modèle SRM (Spike Response Model) de Gerstner (Gerstner & Kistler, 2002). Le modèle SRM possède en particulier un terme appelé le potentiel de membrane “libre” (ou encore “entrée synaptique”) utile pour calculer précisément la probabilité de tir du neurone. La règle d’apprentissage repose sur le produit entre l’activité pré-synaptique et un terme post-synaptique constitué de la différence entre potentiel d’action effectif et probabilité de tir estimée. Cet écart entre la prédiction et le patron de tir effectif est utilisé, en positif ou en négatif selon la récompense, pour modifier la synapse. Cette règle présente de bonnes propriétés de convergence, et peut donc être appliquée de manière continue sans risque de faire diverger l’activité du réseau.

L’interprétation de cette règle, dans le cadre des réseaux récurrents aléatoires, consiste à définir des activités “concurrentes”, et à leur assigner des rôles différents (bruit ou signal) selon le faisceau d’axones considéré. Dans le schéma le plus simple (une couche primaire et une couche associative récurrente aléatoire), le signal d’origine sensorielle est interprété comme un terme de bruit par les synapses des axones récurrents, et, inversement le signal intrinsèque est interprété comme du bruit par les synapses des axones afférents. La plasticité repose ainsi sur l’écart entre la prédiction d’activité liée à l’influence d’un

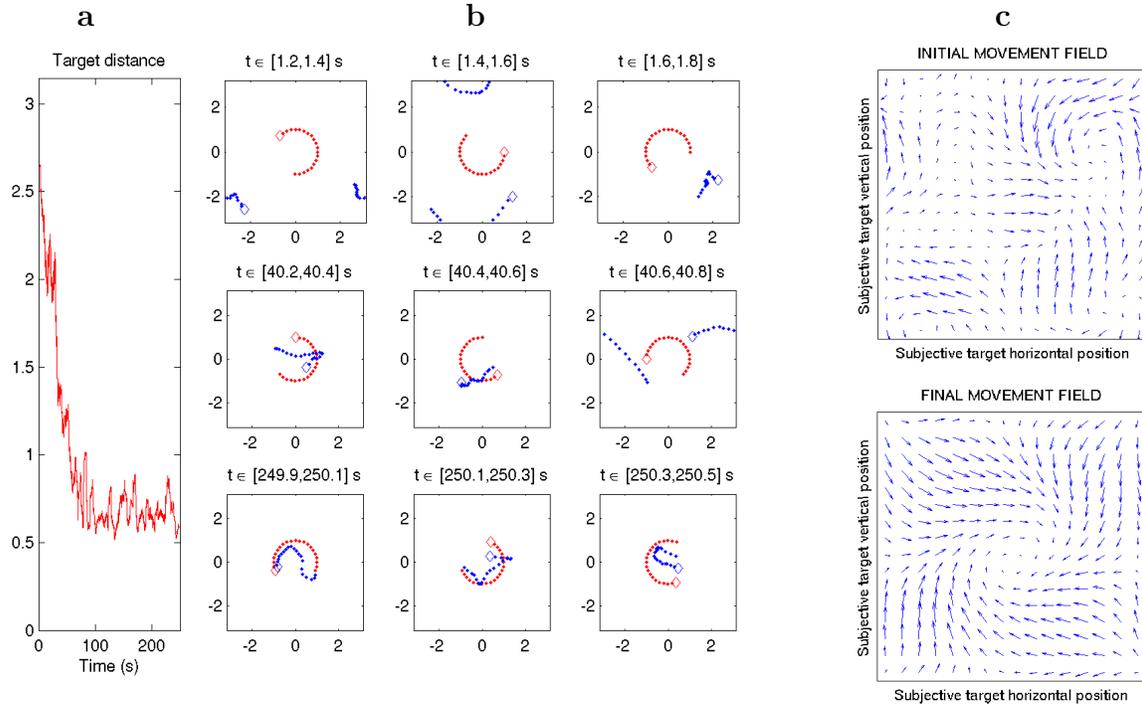


FIGURE 8.5 – **a** : Évolution de la distance moyenne ( $\|\psi - \theta\|$ ) sur des intervalles d'1 s, durant les 250 s du processus d'apprentissage. **b** : extrait des trajectoires du contrôleur et de la cible, sur des intervalles de 200 ms, par pas de 10 ms. Les points rouges donnent la position angulaire de la cible (le diamant donnant la position finale sur l'intervalle considéré). Les points bleus donnent la position angulaire de l'œil. **c** : (haut) carte des réponses motrices initiales, (bas) carte des réponses motrices finales (d'après Daucé, 2009).

faisceau d'axones particulier et l'activité effective observée. Cette règle cherche dans ce cas à atteindre un point d'équilibre entre les influences interne et externe pour établir la réponse motrice. Dans un cadre de robotique simulée (non publié), cette règle a également permis de faire émerger un comportement d'évitement d'obstacles à partir de la perception d'un méta-signal aversif aux collisions.

La plasticité porte ici sur les liens afférents qui vont progressivement faire émerger des champs récepteurs sur les neurones de la couche motrice. Les neurones se spécialisent progressivement en devenant plus sensibles aux activités situées dans le champ visuel opposé à leur mouvement, implémentant un mécanisme de correction d'erreur. Le caractère régulier (et donc prévisible) du mouvement de la cible induit de plus un composant rotationnel, visible sur le champ de mouvement (voir figure 8.5).

La différence principale avec des modèles similaires (Florian, 2006) est l'utilisation d'un bruit "lent" dont le diagramme d'autocorrélation se rapproche de celui de l'erreur motrice (la distance à la cible). Ce bruit lent permet de réduire considérablement les durées d'apprentissage, en testant les réponses (les écarts à l'activité induite par la couche primaire) sur une échelle temporelle similaire à l'erreur motrice.

L'activité intrinsèque des modèles précédents est donc remplacée par un bruit structuré dont les caractéristiques sont plus faciles à maîtriser, au prix de la perte de certaines caractéristiques comme la réduction de complexité (la réduction du bruit) à l'issue de

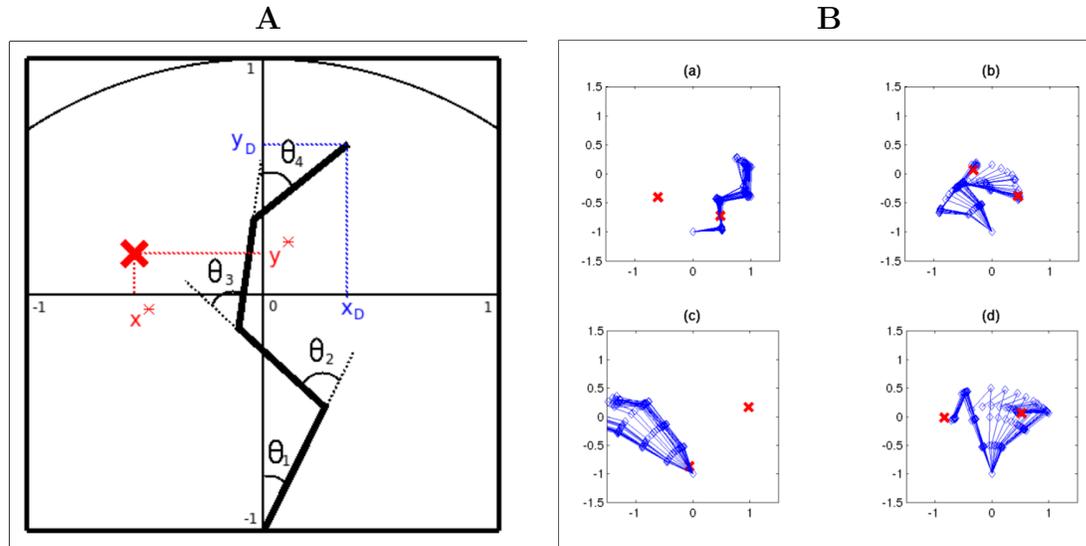


FIGURE 8.6 – **A** : Apprentissage du contrôle d'un bras articulé. Le bras est constitué de quatre segments rigides, le premier segment étant attaché en  $(0, -1)$ . La position du bras est décrite par les quatre positions angulaires des articulations :  $\theta_1, \dots, \theta_4$ . La position de l'extrémité est donnée par  $(x_D, y_D)$ . La position de la cible est donnée par  $(x^*, y^*)$  **B** : Mouvements produits après un déplacement instantané de la cible (représentée par une croix rouge) (a) en début d'expérience (b) après apprentissage (c) après blocage de la troisième articulation (d) après récupération (d'après Daucé & Dutech, 2010).

l'apprentissage.

### Modèle de 2010

Le deuxième modèle, présenté dans (Daucé & Dutech, 2010) (voir annexe J), est l'équivalent linéaire de l'architecture précédente. Ce modèle simplifié est appliqué à des tâches de contrôle plus complexes. Il fournit également une méthodologie d'apprentissage par renforcement relativement originale, bien adaptée aux environnements continus. Le traitement d'états continus passe, comme déjà proposé par (Doya, 2000), par l'utilisation de fonctions noyau qui réalisent un codage topographique des grandeurs extensives perçues : position (et vitesse) de la cible et de l'effecteur final dans l'espace visuel rétinocentré, et données proprioceptives (positions angulaires des joints des effecteurs).

Deux tâches sont étudiées

- Saccades vers des cibles en mouvement
- Contrôle d'un bras articulé

Nous regardons ici comment un bruit générateur interne va permettre de faire des choix en stabilisant certaines réponses dans le cadre de processus d'optimisation non-convexes (plusieurs contrôleurs/solutions possibles), conduisant à des solutions différentes (et donc des environnements perceptifs différents) selon le hasard des initialisation et des choix aléatoires produits par le modèle.

**Saccades vers des cibles en mouvement** Une première tâche, en boucle ouverte, est l'apprentissage de saccades vers des cibles en mouvement, avec une rétine fovéale.

- L'entrée sensorielle est constituée de deux couches topographiques primaires :
  - une couche perçoit la position
  - une couche perçoit la vitesse
- et la commande apprise est la combinaison linéaire de deux commandes (deux champs de mouvement) distincts, conformément à l'hypothèse de "double drive" proposée par (Keller et al., 1996) dans le cadre de l'étude de la saccade oculaire.

Nous montrons que les différents termes de bruit, sur les nœuds de sortie, développent une variabilité suffisante pour permettre à la règle de plasticité de dissocier les deux composantes de la commande.

**Contrôle d'un bras articulé** La seconde tâche est l'apprentissage du contrôle d'un bras articulé à nombreux degrés de liberté (tâche dite de cinématique inverse) (figure 8.6).

- La scène visuelle (position de la cible) est codée sur une carte bidimensionnelle.
- Les données proprioceptives sont également apportées sous forme de quatre cartes unidimensionnelles traduisant la position angulaire des différentes articulations.
- La couche de sortie est constituée de quatre unités motrices dont la fréquence de décharge définit la vitesse angulaire de l'articulation (commande en vitesse).

La cible effectue périodiquement des "sauts" aléatoires dans l'espace visuel. La tâche consiste à guider l'extrémité du bras vers la cible (autrement dit trouver la combinaison de positions angulaires qui fait coïncider la position effective de la cible avec la position effective de l'extrémité du bras). Il s'agit d'une tâche d'optimisation non convexe (de nombreuses solutions existent) dans le cadre d'un dispositif de contrôle en boucle fermée.

- L'apprentissage repose sur une règle de policy gradient adaptée au cas d'une commande linéaire perturbée par un bruit gaussien, comme considéré ici.
- Le signal de récompense repose sur la distance de l'effecteur final à la cible, avec un terme de régularisation pénalisant la vitesse.

Le contrôleur apprend à une commande correspondant, dans l'espace de la tâche, à la différence entre la position de la cible et la position de l'effecteur final déduite des données proprioceptives, par combinaison linéaire du champ récepteur visuel et des champs proprioceptifs angulaires. L'utilisation d'un codage topographique permet de "lisser" la solution, des positions visuellement proches conduisant à des réponses motrices proches.

Cette étude permet d'aborder la question de la prise en compte de différents référentiels dans l'élaboration d'une commande motrice. Certaines données sensorielles sont définies dans un certain référentiel (référentiel de la rétine pour l'œil par exemple) tandis que les commandes opèrent dans un référentiel différent (référentiel du tronc dans le cas de la saccade). Dans notre cas, la commande motrice est calculée à partir du référentiel constitué par la position courante du bras articulé. Cette étude apporte une intuition sur le principe d'un apprentissage "par cœur" des règles de translation permettant de prendre en compte la position du corps dans l'élaboration de la commande motrice dans le cadre du codage topographique (voir aussi (Pouget & Sejnowski, 1997)).

### 8.2.4 Modèles biologiques de l'orientation

Cette section présente une série de résultats relatifs à des projets plus directement orientés vers la modélisation des mécanismes perceptifs et moteurs tels qu'ils apparaissent dans les études neurophysiologiques et comportementales. L'orientation est une des modalités fondamentales de l'activité motrice, consistant à déplacer les organes récepteurs dans une direction qui optimise leur exposition à des signaux d'intérêt. Nous étudions deux modalités de ce comportement d'orientation :

- d'une part l'orientation visuelle, et plus particulièrement la saccade oculaire, en tant que modèle de boucle sensori-motrice élémentaire,
- et de l'autre l'orientation temporelle, c'est-à-dire la capacité du sujet à concentrer ses ressources attentionnelles à un instant choisi dans le temps.

#### Modèles de l'orientation spatiale

L'orientation visuelle est un comportement moteur très spécifique, qui s'apparente plus au monde de la perception qu'au monde de la motricité, reposant sur des circuits de contrôle distincts des circuits de commande de l'appareil moteur. Ce système, situé dans le tectum (partie supérieure du tronc cérébral) au niveau des couches profondes du colliculus supérieur, développe une activité topographiquement organisée responsable de la commande des mouvements rapides des globes oculaires.

La saccade oculaire est un mouvement présent chez l'ensemble des vertébrés consistant à orienter la partie centrale de la rétine (la fovea) vers un point d'intérêt de l'espace visuel. Le comportement d'orientation saccadique, par opposition aux comportements locomoteurs ou alimentaires, est constitué de mouvements brefs (discrets), permettant d'explorer la scène visuelle. La saccade oculaire est le prototype des mouvements dits "balistiques", partant d'une posture initiale et s'achevant sur une posture finale (par opposition à des mouvements d'ajustement ou de contrôle guidé par un feedback comme la poursuite oculaire par exemple).

Etant donnée une cible visuelle présente en périphérie de la rétine, la tâche consiste à positionner le corps (tronc, tête et globe oculaire) de telle sorte que cette cible soit placée au centre de la rétine. Le mouvement consiste donc, étant donnée la posture corps-tête-œil initiale, à produire un déplacement corps-tête-œil conduisant à la posture cible finale. La séquence qui, à partir des données sensorielles et proprioceptives, génère une activité phasique (une bouffée d'activité) des motoneurons conduisant au déplacement désiré est appelée la transformation visuo-motrice (dite transformation "spatio-temporelle" (Moschovakis et al., 1998)). Elle implique la traduction d'une activité exprimant des coordonnées spatiales vers une activité temporellement bornée exprimant une vitesse motrice transitoire. Une version simplifiée de la tâche est la saccade à tête fixe consistant à orienter le globe oculaire seul.

Même si on se limite aux saccades à tête fixe, le circuit de la commande saccadique demeure assez complexe et non totalement élucidé. Si l'on exclut les étapes de traitement visuel conduisant à l'identification d'un point d'intérêt en périphérie de la rétine, la structure au carrefour de la transformation visuo-motrice est une carte rétinotopique située sur la partie supérieure du tectum, appelée le colliculus supérieur. Le déclenchement de la saccade implique de nombreuses catégories de neurones du tronc cérébral, incluant des

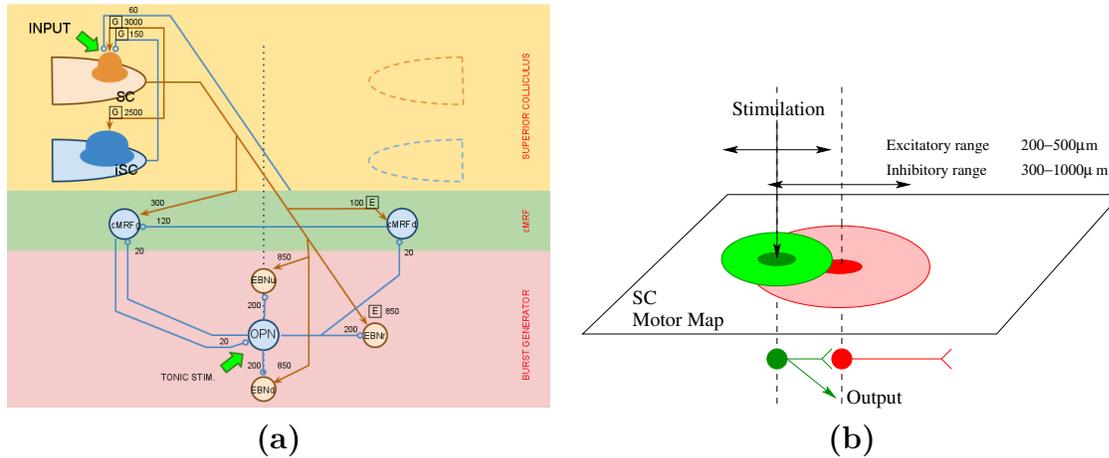


FIGURE 8.7 – (a) Schéma général du contrôle de la saccade, constitué d’une carte topographique bidimensionnelle (Colliculus Supérieur - SC) et du chemin vers les neurones pré-moteurs (générateurs de bouffées - EBN), incluant la région centrale de la formation réticulaire mésencéphalique (cMRF) et les neurones omni-pause (OPN). Le “G” dénote des connexions topographiquement organisées, en fonction de la distance, selon un noyau Gaussien, et le “E” dénote des poids exponentiel décroissant selon la *position* du neurone sur la carte colliculaire. (b) Inhibition disynaptique sur la carte colliculaire. Seuls les neurones excitateurs envoient un axone vers les neurones générateurs de bouffées (d’après Mouraud et al., 2011).

inhibiteurs toniques, des neurones prémoteurs excitateurs et inhibiteurs variés appartenant à la formation réticulée, une dérivation via le cervelet et enfin des motoneurones oculaires.

Le colliculus est une carte topographique bilatérale (champ visuel droit, champ visuel gauche) dont la stimulation électrique permet de déclencher des saccades. L’amplitude des saccades dépend de la position de la stimulation sur la carte, et non de la fréquence ou de l’intensité de la stimulation. La commande motrice étant ballistique (sans feedback visuel ou proprioceptif), les premiers travaux de modélisation (Robinson, 1975) suggèrent l’existence d’un intégrateur dans le circuit de la commande déterminant la durée de la saccade en fonction de la distance déjà parcourue (intégration de la commande de vitesse au cours du temps). Pour tenir compte de certaines observations (en particulier la résistance à des variations de fréquences conduisant à des saccades plus lentes mais toujours précises), l’hypothèse d’un chemin double contrôlant à la fois la vitesse et la durée de la saccade a été émise par (Groh, 2001).

Le travail réalisé en collaboration avec Alain Guillaume et Anthony Mouraud a consisté à étudier et implémenter sur un simulateur de réseaux de neurones réaliste cette hypothèse de chemin double vers les neurones prémoteurs. Le logiciel utilisé est le simulateur de neurones impulsifs “DAMNED” développé par Anthony Mouraud et Didier Puzenat (Mouraud & Puzenat, 2009). Notre modèle étudie la transformation sensori-motrice à l’aide d’une architecture réaliste impliquant plusieurs milliers de neurones. La carte colliculaire est organisée de manière topographique selon un principe de champ neuronal (connexions excitatrices à courte distance, connexions inhibitrices à longue distance).

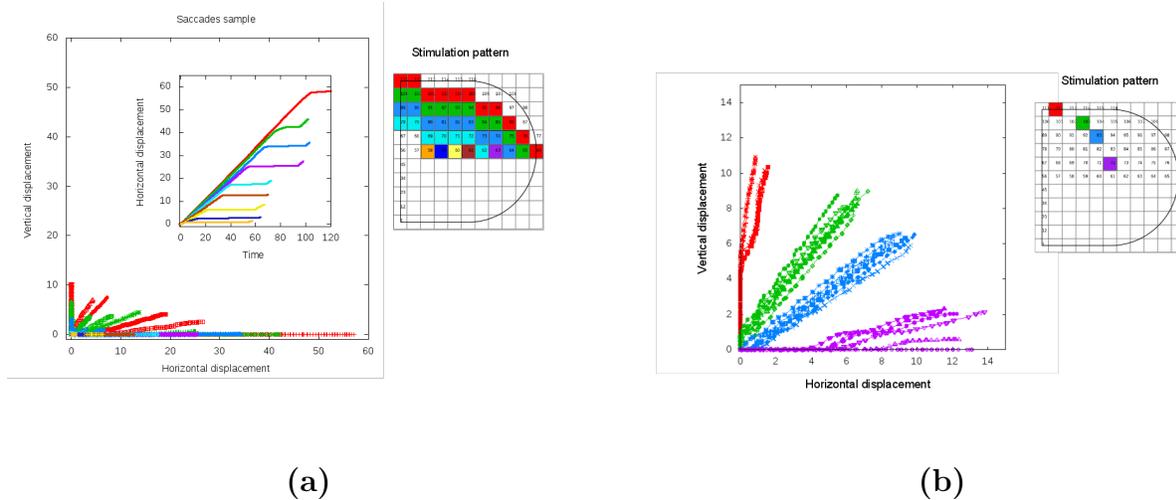


FIGURE 8.8 – (a) Carte des déplacements oculaires en fonction du site de stimulation indiqué en couleur sur la figure de droite (inclusion : décours temporel des saccades horizontales seulement). (b) Déplacement de l’œil pour différentes fréquences et intensités de la stimulation, pour quatre différents sites de stimulation (d’après Mouraud et al., 2011).

Les projections directes des neurones colliculaires vers les neurones pré-moteurs obéissent à un gradient exponentiel, l’amplitude de la saccade augmentant non linéairement avec l’excentricité.

L’étude et la reproduction des mécanismes neuronaux d’orientation visuelle, tels que tentés dans (Daucé et al., 2011) (voir annexe L), montrent l’existence d’une activité autonome contrôlée dans le temps et dans l’espace sur la durée de la saccade (100 ms environ), c’est-à-dire sur un intervalle temporel précis, “programmé”. Notre étude, en implémentant de façon très détaillée les populations de neurones en jeu, a permis de souligner l’importance sans doute sous-évaluée des neurones d’une structure de la région réticulaire (centro-medial reticular formation - cMRF) dans le contrôle de la saccade.

Notre modèle repose sur l’hypothèse d’un contrôle effectué selon trois modalités : contrôle du déclenchement, contrôle de la vélocité et contrôle de la durée. Si la vélocité repose bien sur les projection directe du colliculus vers les neurones pré-moteurs, nous faisons l’hypothèse que le contrôle du déclenchement et le contrôle de la durée reposent sur une structure de la formation réticulée appelée cMRF (centro-medial reticular formation) dont le rôle dans le contrôle de la saccade est connu (Waitzman et al., 1996).

Un premier modèle, présenté à l’édition 2011 de la conférence CNS (Daucé et al., 2011), attribue un rôle double aux cellules du cMRF. Les cellules contralatérales règlent le déclenchement de la saccade tandis que les cellules ipsilatérales participent à sa terminaison via une intégration sous-liminaire du signal émis par les neurones pré-moteurs (voir figures 8.7 et 8.8). Ce modèle utilise des neurones dont la constante de membrane lente permet de faire l’économie d’un mécanisme d’intégration plus élaboré. La brève bouffée émise par les neurones ipsilatéraux lorsque le potentiel atteint le seuil interrompt simultanément l’activité des neurones prémoteurs et celle du colliculus, mettant fin à la saccade.

Ce premier modèle, de fonctionnement relativement simple, ne prend pas en compte

certaines observations qui reportent une prise en compte par le système saccadique de l'activité effective des neurones pré-moteurs (Barton et al., 2003). Un second modèle, incluant un lien de feedback des neurones pré-moteurs vers les neurones du cMRF a été implémenté. Ce modèle, qui permettait de prendre en compte l'interruption temporaire de l'activité des neurones pré-moteurs présentait néanmoins des tracés de saccades moins précis, et n'a pas fait l'objet de publication.

## Modèles de l'orientation temporelle

La problématique de l'orientation temporelle est celle des mécanismes neuronaux supports de l'orientation de l'attention vers des instants précis, "situés" sur un axe temporel. L'objectif général du projet "NOAT" (voir section 1.2) était de comparer deux formes distinctes d'orientation de l'attention temporelle : l'attention implicite (stimulus inclus dans un contexte temporellement structuré) et attention explicite (indices endogènes donnant une information de durée), et de proposer de les caractériser fonctionnellement et neuralemment. Dans le cadre des études en psychologie expérimentale, la focalisation de l'attention peut être mesurée en comparant des temps de réponse entre les essais sur lesquels les cibles apparaissent au moment escompté et ceux sur lesquels elles apparaissent plus tôt ou plus tard. Les études d'imagerie identifient quant à elles le réseau de l'orientation temporelle dans les cortex latéralisés gauches pré-moteur inférieur et pariétal (Coull et al., 2004).

Du côté de la modélisation, les modèles de l'attention temporelle offrent un défi intéressant dans la mesure où la trame temporelle, telle qu'elle est utilisée en physique newtonienne, n'a aucune contrepartie sensorielle. Seule existe dans notre environnement sensible la succession entre les événements, et éventuellement le caractère causal de ces successions. La difficulté des tâches d'orientation temporelle (estimation de l'instant de survenue d'un événement) se traduit par des écarts de jugement temporel importants d'un essai à l'autre, augmentant linéairement avec la durée demandée (ce qui suggère une difficulté proportionnelle à la durée), cette propriété étant observable à la fois chez l'homme et chez l'animal (Gibbon, 1977).

D'un point de vue computationnel, la trame temporelle se présente comme une variable explicative cachée, que le sujet doit reconstruire pour interpréter son environnement. Les modèles computationnels proposés dans la littérature supposent l'existence d'une activité intrinsèque, agissant comme le tic-tac d'une horloge, cette activité étant ensuite intégrée pour former une estimation interne du temps passé (Gold & Shadlen, 2001; Simen et al., 2011).

Nous avons regardé en particulier le cas de la période préparatrice variable ("variable foreperiod"), où les stimuli apparaissent à des intervalles variant d'un essai à l'autre. Dans ce cas, les temps de réponse indiquent une meilleure réactivité pour les intervalles les plus longs. Ainsi les ressources attentionnelles ne sont pas équitablement réparties. Conformément à un principe de révision des croyances, le passage du temps permet en effet d'éliminer les hypothèses portant sur les intervalles courts, donnant plus de vraisemblance (et donc des temps de réaction plus courts) aux hypothèses les plus tardives. L'idée est donc que la réactivité au stimulus (readiness) est proportionnelle à cette probabilité estimée d'apparition.

Le modèle que nous avons proposé, en collaboration avec Gaurav Malhotra, qui a

été recruté en contrat post-doctoral entre 2010 et 2011, utilise l'hypothèse d'une représentation probabiliste de l'environnement, autrement dit une reconstruction à partir des données du modèle génératif de ces données, et une prédiction (sous la forme d'une vraisemblance) de l'instant probable d'un événement.

Le processus implémenté (Malhotra & Daucé, 2011) (voir annexe K) utilise un principe d'accumulation d'évidence concurrente (voir les équations (3.12) et (3.13), page 39). Les différentes hypothèses ont initialement le même a-priori (chaque hypothèse étant assignée à un intervalle temporel distinct). Le déroulement temporel (l'intégration du signal endogène) permet de mettre à jour le modèle génératif au fur et à mesure que certaines hypothèses perdent en plausibilité. Il y a donc une mise à jour dynamique du modèle correspondant à un changement de modèle génératif (une "révision des croyances" : 4 hypothèses, puis 3 hypothèses, puis 2...). De manière intéressante, la meilleure correspondance avec les données psychométriques ne correspond pas au modèle exact des données. Nous avons été obligés de biaiser les probabilités a priori ("responsibilités") en donnant un poids plus important à la source la plus proche de l'estimation temporelle courante, sur le principe d'un biais "optimiste". Une autre hypothèse serait que le poids accordé à l'instant courant serait proportionnel au sentiment d'achèvement, c'est-à-dire à la proportion d'intervalle temporel subjectif parcouru.

Par ailleurs, la similarité formelle de ce modèle avec l'approche probabiliste du contrôle moteur (Haruno, Wolpert & Kawato, 2001) suggère un lien possible entre l'estimation temporelle et la simulation motrice (Jeannerod, 2001).

### 8.2.5 Interfaces cerveau-machine

Une interface cerveau-machine ("Brain Computer Interface" - BCI) est un dispositif qui mesure l'activité électrique à la surface du cerveau (en général de manière non-invasive) afin de produire une commande. Les applications possibles sont en premier lieu la supplémentation motrice pour des handicapés moteurs sévères, mais également les jeux, le monitoring, la réhabilitation motrice, etc. Les dispositifs BCI sont de bons exemples de systèmes "embarqués", avec des possibilités d'interaction en boucle fermée entre le logiciel et l'humain : la réponse produite par le logiciel (le feedback) a une conséquence directement mesurable sur l'activité électrique émise par le sujet, et vice versa. Dans ce cadre, le sujet humain est la source des signaux (il joue le rôle de l'"environnement" du logiciel BCI).

Dans un cadre adaptatif, l'interface logicielle doit donc apprendre, pour chaque utilisateur, à discriminer les tâches différentes activités électriques produites. De la part du sujet, la maîtrise de l'interface comprend également une phase d'apprentissage pouvant induire des modifications dans son activité électro-encéphalographique, dans le sens d'un signal plus net, mais aussi dans la répartition spatiale et les temps de réponse de l'activité électrique. Par ailleurs, des dégradations du signal peuvent aussi être observées en cas de lassitude, de distraction, de somnolence...

Dans le cadre du projet CO-ADAPT (voir section 1.2), ma contribution a principalement porté sur une interface de type clavier virtuel appelée "P300 speller". L'onde "P300" est une activité positive à dominante occipitale qui apparaît environ 300 ms après un stimulus "surprenant", ou différent des autres, dans une série de stimuli qui lui est présentée. Le P300-speller est constitué d'une grille de symboles. Des groupes de symboles

sont illuminés de façon brève et aléatoire à fréquence élevée. Le sujet doit simplement fixer le symbole qu'il souhaite épeler. Une réponse distincte de l'activité de fond est attendue chaque fois que la lettre souhaitée est illuminée. En recoupant les groupes de symboles qui ont induit une réponse P300, le symbole attendu est écrit à l'écran et une nouvelle série de flash démarre.

**Apprentissage direct de la politique** La qualité de la réponse repose sur la justesse du classifieur qui doit discriminer les signaux P300 des signaux non-P300. La première étude sur l'apprentissage par renforcement du P300-speller reprend le schéma de Kivinen (Kivinen et al., 2004) (prise en compte de l'oubli) lorsque le classifieur est soumis à des signaux de renforcement et non des signaux d'erreur classiques, selon le principe du policy-gradient (apprentissage direct de la politique).

Nous avons également regardé comment le choix d'un biais sur la récompense permet de réduire la variance de l'estimateur. Nous avons enfin regardé la tolérance de l'algorithme au bruit, soit en dégradant le signal traité, soit en dégradant le signal de renforcement. Le modèle proposé utilise un terme de régularisation combiné à un paramètre d'apprentissage permettant de définir une constante de temps caractéristique correspondant à la durée pendant laquelle les données du passé sont prises en compte avant d'être oubliées (Daucé et al., 2015) (voir annexe O).

**Accumulation d'évidence** Considérons un classifieur probabiliste qui, à chaque illumination d'un nouveau groupe de lettres (toutes les 200 ms), donne un nombre entre 0 et 1, interprété comme la probabilité que ce groupe de lettre ait produit un P300. A l'aide de ce nombre, il est possible de mettre à jour les probabilités attachées à chaque symbole de la grille. Ainsi, à chaque nouveau flash, la probabilité s'affine jusqu'à ce qu'un symbole soit déclaré "gagnant" lorsque la probabilité qui lui est attachée dépasse un certain seuil. Il est possible de construire un classifieur générique à l'aide des données issues d'un groupe de sujet et de l'appliquer sans calibration sur un nouveau sujet, avec des taux de réussites proches de l'état de l'art des classifieurs spécifiques aux sujets. Le principe d'accumulation permet également d'optimiser le nombre de flashes. Enfin, en ajoutant une hypothèse Gaussienne, il est possible de définir un algorithme adaptatif (gradient logistique) qui met à jour en permanence le classifieur afin de mieux correspondre aux données du sujet et améliorer le taux de réussite et la vitesse de la réponse. La validation du modèle adaptatif Gaussien sur des données P300 a été présenté à la conférence ESANN 2014 (Daucé & Thomas, 2014) (voir annexe N). Le schéma d'accumulation d'évidence fait partie de la mise à jour du logiciel OpenVibes (Renard et al., 2010) développé entre autres à l'INRIA, et a fait l'objet de plusieurs publications dans des conférences (Thomas et al., 2013; Thomas et al., 2014).

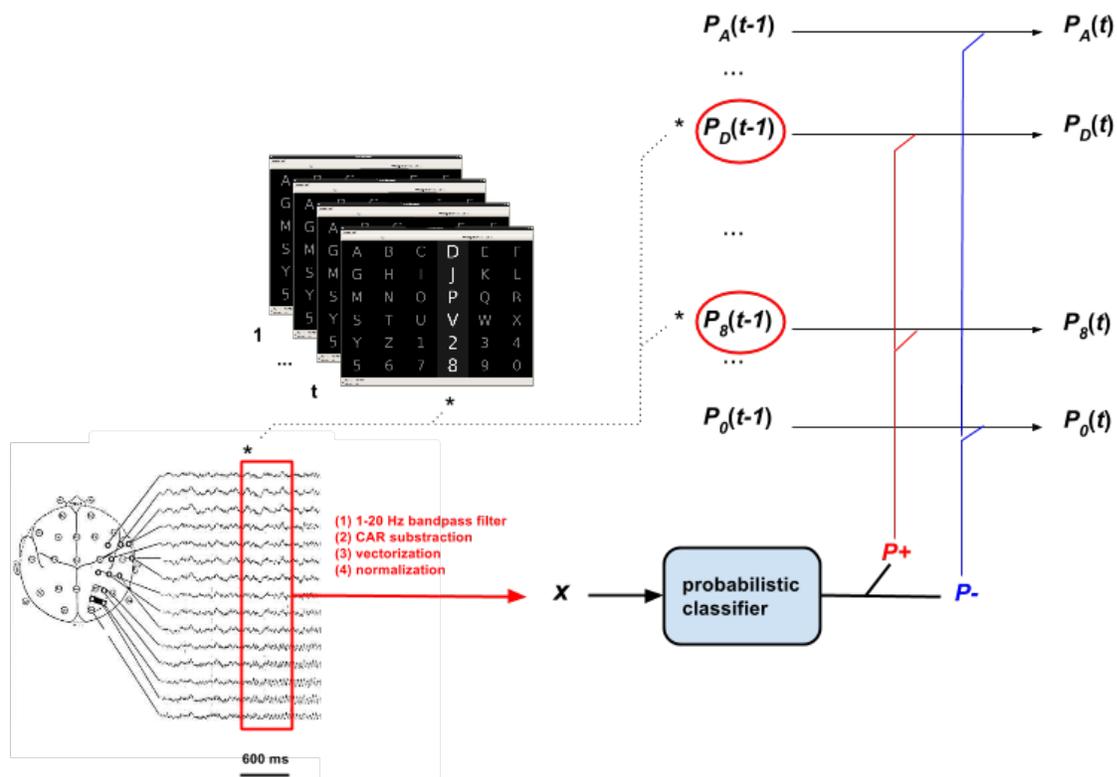


FIGURE 8.9 – Accumulation d’évidence dans le cadre du clavier virtuel “P300”. A chaque illumination d’un groupe de lettres est associée une probabilité  $P^+$  d’émission d’un P300, qui permet de mettre à jour un vecteur de probabilités associé aux différents symboles du clavier virtuel (d’après Daucé & Thomas, 2014b).



# Bibliography

- Abbott, L. & Nelson, S. B. (2000). Synaptic plasticity : taming the beast. *Nature neuroscience supplement*, 3, 1178–1183.
- Abeles, M. (1991). *Corticotronics : Neural circuits of the Cerebelar Cotex*. Cambridge : Cambridge University Press.
- Abeles, M., Bergman, H., Margalit, E. & Vaadia, E. (1993). Spatiotemporal firing patterns in the frontal cortex of behaving monkeys. *J. neurophys.*, 70(4), 1629–1638.
- Amari, S. (1972). Characteristics of random nets of analog neuron-like elements. *IEEE Trans. Syst. Man. Cyb.*, SMC-2(5), 643–657.
- Amari, S. (1977). Dynamics of pattern formation in lateral-inhibition type neural fields. *Biological Cybernetics*, 27, 77–87.
- Amit, D. J. & Brunel, N. (1997). Global spontaneous activity and learned local delay activity in cortical conditions. *Cerebral Cortex*, 7(2), 237–252.
- Amit, D. J., Gutfreund, H. & Sompolinsky, H. (1985). Spin-glass models of neural networks. *Physical Review A*, 32(2), 1007.
- Amit, D. J. & Treves, A. (1989). Associative memory neural network with low temporal spiking rates. *Proceedings of the National Academy of Sciences of the United States of America*, 86, 7871–7875.
- Andry, P., Gaussier, P., Moga, S., Banquet, J.-P. & Nadel, J. (2001). Learning and communication via imitation : An autonomous robot perspective. *Systems, Man and Cybernetics, Part A : Systems and Humans, IEEE Transactions on*, 31(5), 431–442.
- Arnold, V. (1978). *Chapitre supplémentaire sur la théorie des équations différentielles ordinaires*. Moscou : Mir.
- Auer, P., Cesa-Bianchi, N. & Fischer, P. (2002). Finite-time analysis of the multiarmed bandit problem. *Machine learning*, 47(2-3), 235–256.
- Averbeck, B. B., Latham, P. E. & Pouget, A. (2006). Neural correlations, population coding and computation. *Nature Reviews Neuroscience*, 7(5), 358–366.
- Babloyantz, A. (1985). Evidence of chaotic dynamics of brain activity during the sleep cycle. *Phys. Lett. A*, 111(3), 152–156.
- Baraniuk, R. (2007). Compressive sensing. *IEEE signal processing magazine*, 24(4).
- Bartlett, P. & Baxter, J. (1999). Hebbian synaptic modifications in spiking neurons that learn. Technical report, School of Information Science and Engineering, Australian National University.

- Barto, A. G. (1995). *Adaptive Critics and the Basal Ganglia* (pp. 215–232). MIT Press.
- Barton, E. J., Nelson, J. S., Gandhi, N. J. & Sparks, D. L. (2003). Effects of partial lidocaine inactivation of the paramedian pontine reticular formation on saccades of macaques. *J Neurophysiol*, *90*(1), 372–386.
- Baum, L. E. & Petrie, T. (1966). Statistical inference for probabilistic functions of finite state markov chains. *The annals of mathematical statistics* (pp. 1554–1563).
- Beggs, J. M. & Plenz, D. (2004). Neuronal avalanches are diverse and precise activity patterns that are stable for many hours in cortical slice cultures. *The Journal of neuroscience*, *24*(22), 5216–5229.
- Bellman, R. (1956). Dynamic programming and lagrange multipliers. *Proceedings of the National Academy of Sciences of the United States of America*, *42*(10), 767.
- Berger, P., Pommeau, Y. & Vidal, C. (1988). L'ordre dans le chaos. *Hermann, Paris*.
- Beslon, G. & Daucé, E. (2002). Modularité et Apprentissage dans les Réseaux de Neurones Récurrents. In : Guillot A., Daucé E. (Coord.) Approche Dynamique de la Cognition Artificielle, Paris (France) : Eds. Hermès-Lavoisier, pp. 61-80.
- Bi, G.-Q. & Poo, M.-M. (1998). Synaptic modifications in cultured hippocampal neurons : Dependence on spike timing, synaptic strength, and postsynaptic cell type. *The Journal of Neuroscience*, *18*(24), 10464–10472.
- Bi, G.-q. & Poo, M.-m. (2001). Synaptic modification of correlated activity : Hebb's postulate revisited. *Annual Review of Neuroscience*, *24*, 139–166.
- Bicho, E. & Schönner, G. (1997). The dynamics approach to autonomous robotics demonstrated on a low-level vehicle platform. *Robotics and Autonomous System*, *21*, 23–35.
- Bienenstock, E. L., Cooper, L. N. & Munro, P. W. (1982). Theory for the development of neuron selectivity : orientation specificity and binocular interaction in visual cortex. *The Journal of Neuroscience*, *2*(1), 32–48.
- Billings, G. & van Rossum, M. C. (2009). Memory retention and spike-timing-dependent plasticity. *Journal of neurophysiology*, *101*(6), 2775–2788.
- Binzegger, T., Douglas, R. J. & Martin, K. A. (2004). A quantitative map of the circuit of cat primary visual cortex. *The Journal of Neuroscience*, *24*(39), 8441–8453.
- Bishop, C. M. (2006). *Pattern recognition and machine learning*, Volume 1. springer New York.
- Bliss, T. V. P. & Lomo, T. (1973). Long-lasting potentiation of synaptic transmission in the dendate area of the anaesthetized rabbit following stimulation of the perforant path. *J. Physiol.*, *232*, 331–356.
- Boerlin, M., Machens, C. K. & Denève, S. (2013). Predictive coding of dynamical variables in balanced spiking networks. *PLoS Comput Biol*, *9*(11), e1003258.
- Bonhoeffer, T., Grinvald, A. et al. (1991). Iso-orientation domains in cat visual cortex are arranged in pinwheel-like patterns. *Nature*, *353*(6343), 429–431.
- Brembs, B. (2003). Operant conditioning in invertebrates. *Current Opinion in Neurobiology*, *13*, 710–717.

- Bremer, F. (1949). Considérations sur l'origine et la nature des "ondes" cérébrales. *Electroencephalography and clinical neurophysiology*, 1(1), 177–193.
- Brette, R., Rudolph, M., Carnevale, N. T., Hines, M. L., Beeman, D., Bower, J. M., Diesmann, M., Morrison, A., Goodman, P., Friederick, H. J., Zirpe, M., Natschläger, T., Pecevski, D., Ermentrout, G. B., Djurfeldt, M., Lansner, A., Rochel, O., Vieville, T., Muller, E., Davidson, A., El Boustani, S. & Destexhe, A. (2007). Simulation of networks of spiking neurons : A review of tools and strategies. *J. Comp. Neur.*, 23(3), 349–398.
- Brooks, R. (1991). Intelligence without representation. *Artificial Intelligence*, 47, 139–159.
- Brunel, N. (2000a). Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *Journal of Computational Neuroscience*, 8, 183–208.
- Brunel, N. (2000b). Persistent activity and the single-cell frequency-current curve in a cortical network model. *Network : Computation in Neural Systems*, 11(4), 261–280.
- Brunel, N. & Wang, X.-J. (2001). Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. *Journal of Computational Neuroscience*, 11, 63–85.
- Buesing, L., Bill, J., Nessler, B. & Maass, W. (2011). Neural dynamics as sampling : a model for stochastic computation in recurrent networks of spiking neurons. *PLoS Comput. Biol.*, 7(11), e1002211.
- Buonomano, D. V. & Maass, W. (2009). State-dependent computations : spatiotemporal processing in cortical networks. *Nature Reviews Neuroscience*, 10(2), 113–125.
- Busch, N. A., Dubois, J. & VanRullen, R. (2009). The phase of ongoing eeg oscillations predicts visual perception. *The Journal of Neuroscience*, 29(24), 7869–7876.
- Buzsáki, G. & Moser, E. I. (2013). Memory, navigation and theta rhythm in the hippocampal-entorhinal system. *Nature neuroscience*, 16(2), 130–138.
- Cabana, T. & Touboul, J. (2013). Large deviations, dynamics and phase transitions in large stochastic and disordered neural networks. *Journal of Statistical Physics*, 153(2), 211–269.
- Camperi, M. & Wang, X.-J. (1998). A model of visuospatial working memory in prefrontal cortex : Recurrent network and cellular bistability. *J. of Computational Neuroscience*, 5, 383–405.
- Carpenter, G. A. & Grossberg, S. (1987). A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer vision, graphics and image processing*, 37, 54–115.
- Cessac, B. (1994). *Propriétés statistiques des dynamiques de réseaux neuromimétiques*. PhD thesis, Université Paul Sabatier.
- Cessac, B. (1995). Increase in complexity in random neural networks. *Journal de Physique I*, 5, 409–432.
- Cessac, B., Daucé, E., Perrinet, L. & M., S. (2007). Topics in dynamical neural networks from large scale neural networks to motor control and vision – introduction. *European Physical Journal - Special Topics*, 142(1), 1–5.

- Cessac, B., Doyon, B., Quoy, M. & Samuelides, M. (1994). Mean-field equations, bifurcation map and route to chaos in discrete time neural networks. *Physica D*, *74*, 24–44.
- Cessac, B. & Samuelides, M. (2007). From neuron to neural networks dynamics. *The European Physical Journal Special Topics*, *142(1)*, 7–88.
- Compte, A. (2006). Computational and in vitro studies of persistent activity : Edging towards cellular and synaptic mechanisms of working memory. *Neuroscience*, *139(1)*, 135–151.
- Compte, A., Brunel, N., Goldman-Rakic, P. S. & Wang, X.-J. (2000). Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cerebral Cortex*, *10*, 910–923.
- Coull, J. T., Vidal, F., Nazarian, B. & Macar, F. (2004). Functional anatomy of the attentional modulation of time estimation. *Science*, *303(5663)*, 1506–1508.
- Crammer, K., Dekel, O., Keshet, J., Shalev-Shwartz, S. & Singer, Y. (2006). Online passive-aggressive algorithms. *The Journal of Machine Learning Research*, *7*, 551–585.
- Cui, Y., Paille, V., Xu, H., Genet, S., Delord, B., Fino, E., Berry, H. & Venance, L. (2015). Endocannabinoids mediate bidirectional striatal spike-timing-dependent plasticity. *The Journal of physiology*.
- Damoiseaux, J., Rombouts, S., Barkhof, F., Scheltens, P., Stam, C., Smith, S. M. & Beckmann, C. (2006). Consistent resting-state networks across healthy subjects. *Proceedings of the national academy of sciences*, *103(37)*, 13848–13853.
- Daucé, E. (1995). *Dynamique chaotique et apprentissage au sein de réseaux neuromimétiques*. ONERA-CERT. rapport de DEA.
- Daucé, E. (2000). *Adaptation dynamique et apprentissage dans les réseaux de neurones récurrents aléatoires*. PhD thesis, ENSAE.
- Daucé, E. (2001). Learning from chaos : a model of dynamical perception. In Dorffner, G., Bischof, H. & Hornik, K. (Eds.), *ICANN 2001, LNCS 2130* (pp. 1129–1134). Berlin Heidelberg : Springer-Verlag.
- Daucé, E. (2002). *Approche dynamique de la cognition artificielle*, chapter Systèmes dynamiques pour les sciences cognitives, (pp. 33–44). Lavoisier.
- Daucé, E. (2004a). Hebbian reinforcement learning in a modular dynamic network. In *Proceedings of the Eighth International Conference on Simulation of Adaptive Behavior (SAB'04)* (pp. 305–314).
- Daucé, E. (2004b). Short term memory in recurrent networks of spiking neurons. *Natural Computing*, *2(3)*, 135–157.
- Daucé, E. (2007). Learning an control with large dynamic neural networks. *European Physical Journal - Special Topics*, *142(1)*, 123–161.
- Daucé, E. (2009). A model of neuronal specialization using hebbian policy-gradient with "slow" noise. In *Artificial Neural Networks–ICANN 2009* (pp. 218–228). Springer.

- Daucé, E. (2014). Toward stdp-based population action in large networks of spiking neurons. In Verleysen, M. (Ed.), *proceedings of the 22<sup>nd</sup> European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN 2014)* (pp. 29–34). Bruges, Belgium : i6doc.com.
- Daucé, E. & Doyon, B. (1998a). Novelty learning in a discrete-time chaotic network. In Niklasson, L., Boden, M. & Ziemke, T. (Eds.), *ICANN 98*, Volume 2 (pp. 1051–1056). Springer.
- Daucé, E. & Doyon, B. (1998b). Un processus d'apprentissage hebbien dans des réseaux dynamiques à temps discret. In *actes des Neuvièmes journées Neurosciences et Sciences de l'Ingénieur (NSI)*. Munster, France.
- Daucé, E. & Dutech, A. (2010). Online learning with noise : A kernel-based policy-gradient approach. In Neurocomp (Ed.), *Proc. of the fifth french plenary conference on computational neurosciences (Neurocomp'10)* (pp. 252–257). Lyon, France.
- Daucé, E. & Henry, F. (2006). Hebbian learning in large recurrent neural networks. In *Première conférence française de neurosciences computationnelles (Neurocomp'06)* (pp. 202–205). Pont-à-Mousson, France.
- Daucé, E., Mouraud, A. & Guillaume, A. (2011). Simple spatio-temporal transformation with sub-threshold integration in the saccadic system. In *Abstracts from the Twentieth Annual Computational Neuroscience Meeting : CNS\*2011*, Volume 12 (Suppl 1) (p. P133). BMC Neuroscience.
- Daucé, E., Moynot, O., Pinaud, O. & Samuelides, M. (2001). Mean-field theory and synchronization in random recurrent neural networks. *Neural Processing Letters*, 14, 115–126.
- Daucé, E., Moynot, O., Pinaud, O., Samuelides, M. & Doyon, B. (1999). Mean field equations reveal synchronization in a 2-populations neural network model. In Verleysen, M. (Ed.), *ESANN 99* (pp. 7–12). D-Facto.
- Daucé, E. & Proix, T. (2013). P300-speller adaptivity to change with a backspace key. In *proc. of TOBI workshop IV : Practical Brain-Computer Interfaces for End-Users : Progresses and Challenges* (pp. 105–106). Sion, Switzerland.
- Daucé, E., Proix, T. & Ralaivola, L. (2013). Fast online adaptivity with policy gradient : example of the bci "p300" speller. In Verleysen, M. (Ed.), *proc. of the 21th European Symposium on Artificial Neural Networks, computational intelligence and machine learning (ESANN 2013)* (pp. 197–202). Bruges, Belgium.
- Daucé, E., Proix, T. & Ralaivola, L. (2015). Reward-based online learning in non-stationary environments : adapting a p300-speller with a "backspace" key. In *Proceedings of the 2015 International Joint Conference on Neural Networks (IJCNN 2015)*. Killarney, Ireland.
- Daucé, E. & Quoy, M. (2000). Random recurrent neural networks for autonomous systems design. In *Proceedings of the Sixth International Conference on Simulation of Adaptive Behavior (SAB'2000)* (pp. 31–40). Meyer, J.-A.
- Daucé, E., Quoy, M., Cessac, B., Doyon, B. & Samuelides, M. (1998). Self-organization and dynamics reduction in recurrent networks : stimulus presentation and learning. *Neural Networks*, 11, 521–533.

- Daucé, E., Quoy, M. & Doyon, B. (2002). Resonant spatiotemporal learning in large random recurrent networks. *Biological Cybernetics*, *87*, 185–198.
- Daucé, E. & Ralaivola, L. (2012). Approche adaptative pour les interfaces cerveau-machine. In *Actes des XIXèmes rencontres de la société francophone de classification*. Marseille, France.
- Daucé, E., Soula, H. & Beslon, G. (2005). Learning methods for dynamic neural networks. In *Proceedings of 2005 International Symposium on Nonlinear Theory and its Applications (NOLTA2005) Bruges, Belgium 2005* (pp. 598–601).
- Daucé, E. & Thomas, E. (2014). Evidence build-up facilitates on-line adaptivity in dynamic environments : example of the bci p300-speller. In Verleysen, M. (Ed.), *proceedings of the 22<sup>nd</sup> European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN 2014)* (pp. 377–382). Bruges, Belgium : i6doc.com.
- Daw, N. D. & Doya, K. (2006). The computational neurobiology of learning and reward. *Current Opinion in Neurobiology*, *16*, 199–204.
- De Zeeuw, C. I., Hoogenraad, C. C., Koekkoek, S., Ruigrok, T. J., Galjart, N. & Simpson, J. I. (1998). Microcircuitry and function of the inferior olive. *Trends in neurosciences*, *21(9)*, 391–400.
- Dean, A. (1981). The variability of discharge of simple cells in the cat striate cortex. *Experimental Brain Research*, *44(4)*, 437–440.
- Dean, P., Porrill, J., Ekerot, C.-F. & Jörntell, H. (2010). The cerebellar microcircuit as an adaptive filter : experimental and computational evidence. *Nature Reviews Neuroscience*, *11(1)*, 30–43.
- Dehaene, S. & Changeux, J.-P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron*, *70(2)*, 200–227.
- Dempster, A. P., Laird, N. M. & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)* (pp. 1–38).
- Destexhe, A. (1997). Conductance-based integrate-and-fire models. *Neural Computation*, *9(3)*, 503–514.
- Destexhe, A., Rudolph, M. & Paré, D. (2003). The high-conductance state of neocortical neurons in vivo. *Nature reviews neuroscience*, *4(9)*, 739–751.
- Doucet, A., De Freitas, N. & Gordon, N. (2001). *An introduction to sequential Monte Carlo methods*. Springer.
- Doya, K. (2000). Reinforcement learning in continuous time and space. *Neural Computation*, *(12)*, 219–245.
- Doyon, B., Cessac, B., Quoy, M. & Samuelides, M. (1993). Control of the transition to chaos in neural networks with random connectivity. *International Journal of Bifurcation and Chaos*, *3(02)*, 279–291.
- Dupuy, J.-P. & Varela, F. J. (1992). Understanding origins : an introduction. In *Understanding Origins* (pp. 1–25). Springer.

- Eliasmith, C. & Anderson, C. H. (2004). *Neural engineering : Computation, representation, and dynamics in neurobiological systems*. MIT press.
- Faugeras, O., Frégnac, Y. & Samuelides, M. (2007). A future for systems and computational neuroscience in france? *Journal of Physiology-Paris*, *101(1)*, 1–8.
- Faugeras, O., Touboul, J. & Cessac, B. (2009). A constructive mean-field analysis of multi-population neural networks with random synaptic weights and stochastic inputs. *Frontiers in computational neuroscience*, *3*.
- Fiser, J., Berkes, P., Orbán, G. & Lengyel, M. (2010). Statistically optimal perception and learning : from behavior to neural representations. *Trends in cognitive sciences*, *14(3)*, 119–130.
- Flash, T. & Hochner, B. (2005). Motor primitives in vertebrates and invertebrates. *Current opinion in neurobiology*, *15(6)*, 660–666.
- Florian, R. V. (2006). Reinforcement learning through modulation of spike-timing dependent synaptic plasticity. *Neural Computation*.
- Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C. & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the United States of America*, *102(27)*, 9673–9678.
- Friston, K. (2010). The free-energy principle : a unified brain theory? *Nature Reviews Neuroscience*, *11(2)*, 127–138.
- Friston, K. & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B : Biological Sciences*, *364(1521)*, 1211–1221.
- Funahashi, S., Bruce, C. J. & Goldman-Rakic, P. S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J. Neurophysiol.*, *61*, 331–349.
- Furber, S. B., Lester, D. R., Plana, L., Garside, J. D., Painkras, E., Temple, S., Brown, A. D. et al. (2013). Overview of the spinnaker system architecture. *Computers, IEEE Transactions on*, *62(12)*, 2454–2467.
- Gaussier, P. & Zrehen, S. (1995). Perac : A neural architecture to control artificial animals. *Robotics and Autonomous Systems*, *16(2)*, 291–320.
- Georgopoulos, A. P., Schwartz, A. B. & Kettner, R. E. (1986). Neuronal population coding of movement direction. *Science*, *233(4771)*, 1416–1419.
- Gerstner, W. & Kistler, W. (2002). *Spiking Neuron Models. Single Neurons, Populations, Plasticity*. Cambridge University Press.
- Gibbon, J. (1977). Scalar expectancy theory and weber's law in animal timing. *Psychological Review*, *84(3)*, 279.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston : Houghton-Mifflin.
- Gold, J. & Shadlen, M. N. (2001). Neural computations that underlie decisions about sensori stimuli. *Trends in Cognitive Sciences*, *5(1)*, 10–16.

- Golos, M., Jirsa, V. & Daucé, E. (2013). Multistability in large-scale models of brain activity. In *Abstracts from the Twenty Second Annual Computational Neuroscience Meeting : CNS\*2013*, Volume 14 (Suppl 1) (p. P84). BMC Neuroscience.
- Golos, M., Jirsa, V. & Daucé, E. (2015). Multistability in large scale models of brain activity. *PLoS Comput Biol*, *11*(12), e1004644.
- Graupner, M. & Brunel, N. (2012). 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), 3991–3996.
- Gray, C., König, P., Engel, A. & Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflect global stimulus properties. *Nature*, *7*, 334–338.
- Graziano, M. S., Taylor, C. S. & Moore, T. (2002). Complex movements evoked by microstimulation of precentral cortex. *Neuron*, *34*(5), 841–851.
- Greene, P. H. (1962). On looking for neural networks and "cell assemblies" that underlie behavior : I. a mathematical model. *The Bulletin of mathematical biophysics*, *24*(3), 247–275.
- Grillner, S., Wallén, P., Saitoh, K., Kozlov, A. & Robertson, B. (2008). Neural bases of goal-directed locomotion in vertebrates—an overview. *Brain research reviews*, *57*(1), 2–12.
- Groh, J. M. (2001). Converting neural signals from place codes to rate codes. *Biol Cybern*, *85*, 159–165.
- Grossberg, S. (1970). Neural pattern discrimination. *Journal of Theoretical Biology*, *27*(2), 291–337.
- Guillot, A. & Daucé, E. (2002). *Approche dynamique de la cognition artificielle*. Lavoisier.
- Gurney, K., Prescott, T. J. & Redgrave, P. (2001). A computational model of action selection in the basal ganglia. i. a new functional anatomy. *Biological cybernetics*, *84*(6), 401–410.
- Guyonneau, R., VanRullen, R. & Thorpe, S. (2005). Neurons tune to the earliest spikes through stdp. *Neural Computation*, *17*, 559–879.
- Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C. J., Wedeen, V. J. & Sporns, O. (2008). Mapping the structural core of human cerebral cortex. *PLoS biology*, *6*(7), e159.
- Haimovici, A., Tagliazucchi, E., Balenzuela, P. & Chialvo, D. R. (2013). Brain organization into resting state networks emerges at criticality on a model of the human connectome. *Physical review letters*, *110*(17), 178101.
- Haken, H. (1983). *Advanced synergetics*. Springer.
- Haken, H., Kelso, J. S. & Bunz, H. (1985). A theoretical model of phase transitions in human hand movements. *Biological cybernetics*, *51*(5), 347–356.
- Hansel, D. & Sompolinsky, H. (1996). Chaos and synchrony in a model of a hypercolumn in visual cortex. *J. Comp. Neurosc.*, *3*, 7–34.

- Harris, E. W., Ganong, A. H. & Cotman, C. W. (1984). Long-term potentiation in the hippocampus involves activation of n-methyl-d-aspartate receptors. *Brain research*, *323(1)*, 132–137.
- Haruno, M., Wolpert, D. & Kawato, M. (2001). Mosaic model for sensorimotor learning and control. *Neural computation*, *13(10)*, 2201–2220.
- Haykin, S. (1999). *Neural Networks : a comprehensive foundation*. Prentice-Hall.
- Hebb, D. (1949). *The Organization of behavior*. New York : Wiley.
- Henry, F. & Daucé, E. (2006). Temporal pattern identification using spike-timing dependent plasticity. In *proc. of the 15th annual Computational Neurosciences meeting (CNS\*2006)*. Edinburgh, U.K.
- Henry, F. & Daucé, E. (2008a). Emergence of stimulus-specific synchronous response through stdp in recurrent neural networks. In *16th european symposium on artificial neural networks, advances in computational intelligence and learning (ESANN 2008)*. Bruges, Belgium.
- Henry, F. & Daucé, E. (2008b). Spike-timing dependent plasticity and regime transitions in random recurrent neural networks. In *Proceedings of the second french conference on computational neurosciences Neurocomp08*. Marseille, France.
- Henry, F. & Daucé, E. (2009). Formation of synchronous activity through stdp in recurrent neural networks with heterogeneous delays. In Johnson, D. (Ed.), *Proc. of the eighteens annual computational neuroscience meeting (CNS\*2009)*. Berlin, Germany.
- Henry, F., Daucé, E. & Soula, H. (2007). Temporal pattern identification using spike-timing dependent plasticity. *Neurocomputing*, *70*, 2009–2016.
- Hinton, G. E., Osindero, S. & Teh, Y.-W. (2006). A fast learning algorithm for deep belief nets. *Neural computation*, *18(7)*, 1527–1554.
- Hodgkin, A. L. & Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology*, *117(4)*, 500–544.
- Hopfield, J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proc. Nat. Acad. Sci.*, *79*, 2554–2558.
- Hopfield, J. J. (1984). Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the National Academy of Sciences of the United States of America*, *81*, 3088–3092.
- Huang, G.-B., Zhu, Q.-Y. & Siew, C.-K. (2006). Extreme learning machine : theory and applications. *Neurocomputing*, *70(1)*, 489–501.
- Huang, Y. & Rao, R. P. (2014). Neurons as monte carlo samplers : Bayesian ?inference and learning in spiking networks. In Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence & K. Weinberger (Eds.), *Advances in Neural Information Processing Systems 27* (pp. 1943–1951). Curran Associates, Inc.
- Hubel, D. H. & Wiesel, T. N. (1962). Receptive fields, binocular interactions and functional architecture in the cat’s visual cortex. *Journal of Physiology*, *160*, 106–154.

- Ito, M. (1982). Cerebellar control of the vestibulo-ocular reflex—around the flocculus hypothesis. *Annual review of neuroscience*, 5(1), 275–297.
- Izhikevich, E. M. (2006). Polychronization : Computation with spikes. *Neural Computation*, 18, 245–282.
- Izhikevich, E. M., Gally, J. A. & Edelman, G. M. (2004). Spike-timing dynamics of neuronal groups. *Cerebral Cortex*, 14, 933–944.
- Jeannerod, M. (2001). Neural simulation of action : a unifying mechanism for motor cognition. *Neuroimage*, 14(1), S103–S109.
- Joel, D., Niv, Y. & Ruppin, E. (2002). Actor–critic models of the basal ganglia : New anatomical and computational perspectives. *Neural networks*, 15(4), 535–547.
- Kakade, S. M., Shalev-Shwartz, S. & Tewari, A. (2008). Efficient bandit algorithms for online multiclass prediction. In *Proceedings of the 25th international conference on Machine learning* (pp. 440–447).
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Fluids Engineering*, 82(1), 35–45.
- Kaneko, K. (1990). Clustering, coding, switching, hierarchical ordering and control in a network of chaotic elements. *Physica D*, (41), 137–172.
- Keller, E., Gandhi, N. & Weir, P. (1996). Discharge of superior collicular neurons during saccades made to moving targets. *Journal of neurophysiology*, 76(5), 3573–3577.
- Kempler, R., Gerstner, W. & Van Hemmen, J. L. (1999). Hebbian learning and spiking neurons. *Physical Review*, 59(4), 4498–4514.
- Kepler, T. B., Abbott, L. & Marder, E. (1992). Reduction of conductance-based neuron models. *Biological cybernetics*, 66(5), 381–387.
- Khoei, M. A., Masson, G. S. & Perrinet, L. U. (2013). Motion-based prediction explains the role of tracking in motion extrapolation. *Journal of Physiology-Paris*, 107(5), 409–420.
- Kindermans, P.-J., Verstraeten, D. & Schrauwen, B. (2012). A bayesian model for exploiting application constraints to enable unsupervised training of a p300-based bci. *PloS one*, 7(4), e33758.
- Kivinen, J., Smola, A. J. & Williamson, R. C. (2004). Online learning with kernels. *Signal Processing, IEEE Transactions on*, 52(8), 2165–2176.
- Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43, 59–69.
- König, P. & Engel, A. K. (1995). Correlated firing in sensory-motor systems. *Current Opinion in Neurobiology*, 5(4), 511–519.
- König, P. & Schillen, T. (1991). Stimulus dependent assembly formation of oscillatory responses : I. synchronization. *Neural Computation*, 3, 155–166.
- Kushner, H. J. (1964). On the dynamical equations of conditional probability density functions, with applications to optimal stochastic control theory. *Journal of Mathematical Analysis and Applications*, 8(2), 332–344.

- Lansner, A. (2009). Associative memory models : from the cell-assembly theory to biophysically detailed cortex simulations. *Trends in neurosciences*, 32(3), 178–186.
- Lapicque, L. (1907). Recherches quantitatives sur l'excitation électrique des nerfs traitée comme une polarisation. *J. Physiol. Pathol. Gen*, 9(1), 620–635.
- Le Bihan, D., Mangin, J.-F., Poupon, C., Clark, C. A., Pappata, S., Molko, N. & Chabriat, H. (2001). Diffusion tensor imaging : concepts and applications. *Journal of magnetic resonance imaging*, 13(4), 534–546.
- Le Cun, L. B. Y. & Bottou, L. (2004). Large scale online learning. *Advances in neural information processing systems*, 16, 217.
- LeCun, Y., Bengio, Y. & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444.
- Lee, D. D. & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755), 788–791.
- Leon, P. S., Knock, S. A., Woodman, M. M., Domide, L., Mersmann, J., McIntosh, A. R. & Jirsa, V. (2013). The virtual brain : a simulator of primate brain network dynamics. *Frontiers in neuroinformatics*, 7.
- Lindblad, T. & Kinser, J. M. (1998). *Image processing using pulse-coupled neural networks*. Springer.
- Lundqvist, M., Herman, P. & Lansner, A. (2012). Variability of spike firing during theta-coupled replay of memories in a simulated attractor network. *Brain research*, 1434, 152–161.
- Maass, W., Natschläger, T. & Markram, H. (2002). Real-time computing without stable states : A new framework for neural computation based on perturbations. *Neural computation*, 14(11), 2531–2560.
- MacLeod, K. & Laurent, G. (1996). Distinct mechanisms for synchronization and temporal patterning of odor-encoding cell assemblies. *Science*, 274, 976–979.
- MacLennan, B. J. (1999). Field computation in natural and artificial intelligence. *Information Sciences*, 119(1), 73–89.
- Malhotra, G. & Daucé, E. (2011). How and where does the brain predict the when : a bayesian approach to modeling temporal expectation. *BMC Neuroscience*, 12(Suppl 1), P53.
- Von der Malsburg, C. (1986). A neural cocktail-party processor. *Biol. Cybern.*, 54, 29–40.
- Markram, H. (2006). The blue brain project. *Nature Reviews Neuroscience*, 7(2), 153–160.
- Markram, H. (2012). The human brain project. *Scientific American*, 306(6), 50–55.
- Marr, D. (1982). *Vision : a computational investigation into the human representation and processing of visual information*. San Francisco : WH Freeman and Company.
- McClelland, J. L., McNaughton, B. L. & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex : insights from the successes and failures of connectionist models of learning and memory. *Psychological review*, 102(3), 419.
- McCulloch, W. S. & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.*, 5, 115–133.

- Morrison, A., Aertsen, A. & Diesmann, M. (2007). Spike-timing-dependent plasticity in balanced random networks. *Neural Computation*, *19*, 1437–1467.
- Moschovakis, A. K., Kitama, T., Dalezios, Y., Petit, J., Brandi, A. M. & Grantyn, A. A. (1998). An anatomical substrate for the spatiotemporal transformation. *J. Neurosci.*, *18*(23), 10219–10229.
- Mouraud, A. & Puzenat, D. (2009). Simulation of large spiking neural networks on distributed architectures, the "damned" simulator. In *Engineering Applications of Neural Networks* (pp. 359–370). Springer.
- Moynot, O. & Samuelides, M. (2002). Large deviations and mean-field theory for asymmetric random recurrent neural networks. *PTRF*, *123*(1), 41–75.
- Mussa-Ivaldi, F. A. & Solla, S. A. (2004). Neural primitives for motion control. *Oceanic Engineering, IEEE Journal of*, *29*(3), 640–650.
- O'Donnell, C., Nolan, M. F. & van Rossum, M. C. (2011). Dendritic spine dynamics regulate the long-term stability of synaptic plasticity. *The Journal of neuroscience*, *31*(45), 16142–16156.
- Oja, E. (1982). Simplified neuron model as a principal component analyzer. *Journal of mathematical biology*, *15*(3), 267–273.
- O'Keefe, J. & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford University Press.
- Olshausen, B. A. & Field, D. J. (1997). Sparse coding with an overcomplete basis set : A strategy employed by v1? *Vision research*, *37*(23), 3311–3325.
- O'Regan, K. & Noë, G. (2001). Sensorimotor theory of visual experience. *Behav. and Brain Sciences*.
- Paillard, J. & Beaubaton, D. (1978). De la coordination visuo-motrice à l'organisation de la saisie manuelle. In *Du contrôle moteur à l'organisation du geste* (pp. 225–260). Masson Paris.
- Palva, J. M., Palva, S. & Kaila, K. (2005). Phase synchrony among neuronal oscillations in the human cortex. *The Journal of Neuroscience*, *25*(15), 3962–3972.
- Perrinet, L. U. (2010). Role of homeostasis in learning sparse representations. *Neural computation*, *22*(7), 1812–1836.
- Perrinet, L. U., Adams, R. A. & Friston, K. J. (2014). Active inference, eye movements and oculomotor delays. *Biological cybernetics*, *108*(6), 777–801.
- Petermann, T., Thiagarajan, T. C., Lebedev, M. A., Nicolelis, M. A., Chialvo, D. R. & Plenz, D. (2009). Spontaneous cortical activity in awake monkeys composed of neuronal avalanches. *Proceedings of the National Academy of Sciences*, *106*(37), 15921–15926.
- Peters, J. & Schaal, S. (2008). Reinforcement learning of motor skills with policy gradients. *Neural Networks*, *21*, 682–697.
- Petitot, J., Varela, F., Pachoud, B. & Roy, J.-M. (2002). Naturaliser la phénoménologie. *Essais sur la phénoménologie contemporaine et les sciences cognitives*. Paris : CNRS Editions.

- Piaget, J. (1973). *Biologie et connaissance*. Gallimard idées.
- Potthast, R. et al. (2013). Universal neural field computation. *arXiv preprint arXiv :1312.3550*.
- Pouget, A., Deneve, S. & Duhamel, J.-R. (2002). A computational perspective on the neural basis of multisensory spatial representations. *Nature Reviews Neuroscience*, 3(9), 741–747.
- Pouget, A. & Sejnowski, T. (1997). Spatial transformations in the parietal cortex using basis functions. *Cognitive Neuroscience, Journal of*, 9(2), 222–237.
- Quiroga, R. Q., Reddy, L., Kreiman, G., Koch, C. & Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature*, 435(7045), 1102–1107.
- Quoy, M. (1994). *Apprentissage dans les réseaux neuromimétiques à dynamique chaotique*. PhD thesis, ENSAE.
- Quoy, M., Banquet, J.-P. & Daucé, E. (2001). Learning and control with chaos : From biology to robotics. *Behavioral and Brain Sciences*, 24(05), 824–825.
- Quoy, M. & Daucé, E. (2000). Visual and motor learning using a chaotic recurrent neural network : application to the control of a mobile robot. In *NC 2000*. accepted for oral presentation.
- Rachez, A., Proix, T., Maby, E., Mattout, J. & Daucé, E. (2011). Direct policygradient for online learning in bci. *International Journal of Bioelectromagnetism*, 13(1), 52–53.
- Raichle, M. E., MacLeod, a. M., Snyder, a. Z., Powers, W. J., Gusnard, D. a. & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, 98(2), 676–82.
- Rall, W. (1959). Branching dendritic trees and motoneuron membrane resistivity. *Experimental neurology*, 1(5), 491–527.
- Rao, R. & Sejnowski, T. (2001). Spike-timing-dependent hebbian plasticity as temporal difference learning. *Neural Computation*, 13, 2221–2237.
- Rao, R. P. & Ballard, D. H. (1999). Predictive coding in the visual cortex : a functional interpretation of some extra-classical receptive-field effects. *Nature neuroscience*, 2(1), 79–87.
- Renard, Y., Lotte, F., Gibert, G., Congedo, M., Maby, E., Delannoy, V., Bertrand, O. & Lécuyer, A. (2010). Openvibe : an open-source software platform to design, test, and use brain-computer interfaces in real and virtual environments. *Presence : teleoperators and virtual environments*, 19(1), 35–53.
- Revel, A. & Andry, P. (2009). Emergence of structured interactions : From a theoretical model to pragmatic robotics. *Neural networks*, 22(2), 116–125.
- Robinson, D. A. (1975). *Oculomotor control signals*, chapter Basic mechanisms of ocular motility and their clinical implications, (pp. 337–374). Oxford, UK : Pergamon.
- Rodriguez, E., George, N., Lachaux, J.-P., Martinerie, J., Renault, B. & Varela, F. (1999). Perception’s shadow : long-distance synchronization of human brain activity. *Nature*, 397(6718), 430–433.

- Rolfs, M., Jonikaitis, D., Deubel, H. & Cavanagh, P. (2011). Predictive remapping of attention across eye movements. *Nature neuroscience*, *14*(2), 252–256.
- Rosenblatt, F. (1958). The perceptron : A probabilistic model for information storage and organization in the brain. *Psychological Review*, *65*(6), 386–408.
- Roxin, A., Brunel, N. & Hansel, D. (2005). Role of delays in shaping spatiotemporal dynamics of neuronal activity in large networks. *Physical review letters*, *94*(23), 238103.
- Rubinov, M., Sporns, O., Thivierge, J.-P. & Breakspear, M. (2011). Neurobiologically realistic determinants of self-organized criticality in networks of spiking neurons. *PLoS Comput Biol*, *7*(6), e1002038.
- Rumelhart, D., Hinton, G. & Williams, R. (1986). *Parallel Distributed Processing : Explorations in the microstructure of cognition*, Volume I, chapter Learning internal representations by error propagation, (pp. 318–362). MIT Press.
- Samuelides, M., Cessac, B., Beslon, G., Daucé, E., Perrinet, L., Quoy, M. & Thorpe, S. (2007). Dynamique des réseaux de neurones artificiels biologiquement plausibles en robotique autonome. In *Proceedings of the ACI neurosciences intégratives et computationnelles*. Paris, France.
- Schölkopf, B. & Smola, A. J. (2002). *Learning with kernels : support vector machines, regularization, optimization, and beyond*. MIT press.
- Schöner, G., Dose, M. & Engels, C. (1995). Dynamics of behavior : theory and applications for autonomous robot architectures. *Robotics and Autonomous System*, *16*(2-4), 213–245.
- Schultz, W., Dayan, P. & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593–1599.
- Sejnowski, T. J. (1977). Strong covariance with nonlinearly interacting neurons. *Journal of Mathematical biology*, *4*, 303–321.
- Sen Song, Kenneth D. Miller, L. A. (2000). Competitive hebbian learning through spike-timing dependent synaptic plasticity. *Nature*.
- Shadlen, M. N. & Newsome, W. T. (1998). The variable discharge of cortical neurons : implications for connectivity, computation, and information coding. *The Journal of neuroscience*, *18*(10), 3870–3896.
- Shouval, H. Z., Bear, M. F. & Cooper, L. N. (2002). A unified model of nmda receptor-dependent bidirectional synaptic plasticity. *Proceedings of the National Academy of Sciences*, *99*(16), 10831–10836.
- Siegelmann, H. (1999). *Neural networks and analog computation : beyond the Turing limit*, Volume 20. Springer.
- Simen, P., Balci, F., Cohen, J. D., Holmes, P. et al. (2011). A model of interval timing by neural integration. *The Journal of Neuroscience*, *31*(25), 9238–9253.
- Skaggs, W. E., McNaughton, B. L., Wilson, M. A. & Barnes, C. A. (1996). Theta phase precession in hippocampal neuronal populations and the compression of temporal sequences. *Hippocampus*, *6*(2), 149–172.

- Skarda, C. & Freeman, W. (1987). How brains make chaos in order to make sense of the world. *Behav. Brain Sci.*, *10*, 161–195.
- Softky, W. & Koch, C. (1993). The highly irregular firing of cortical cells is inconsistent with temporal integration of random epsps. *J. of Neuroscience*, *13(1)*, 334–450.
- Sompolinsky, H., Crisanti, A. & Sommers, H. (1988). Chaos in random neural networks. *Phys. Rev. Lett.*, *61*, 259–262.
- Soula, H., Beslon, G. & Mazet, O. (2006). Spontaneous dynamics of asymmetric random recurrent spiking neural networks. *Neural Computation*, *18(1)*, 60–79.
- Stephenson-Jones, M., Samuelsson, E., Ericsson, J., Robertson, B. & Grillner, S. (2011). Evolutionary conservation of the basal ganglia as a common vertebrate mechanism for action selection. *Current Biology*, *21(13)*, 1081–1091.
- Sternad, D. (2008). Towards a unified theory of rhythmic and discrete movements—behavioral, modeling and imaging results. In *Coordination : Neural, behavioral and social dynamics* (pp. 105–133). Springer.
- Stewart, T. C., Bekolay, T. & Eliasmith, C. (2011). Neural representations of compositional structures : Representing and manipulating vector spaces with spiking neurons. *Connection Science*, *23(2)*, 145–153.
- Sutskever, I., Hinton, G. E. & Taylor, G. W. (2009). The recurrent temporal restricted boltzmann machine. In D. Koller, D. Schuurmans, Y. Bengio & L. Bottou (Eds.), *Advances in Neural Information Processing Systems 21* (pp. 1601–1608). Curran Associates, Inc.
- Sutton, R. (1988). Learning to predict by the method of temporal differences. *Machine learning*, *3*, 9–44.
- Sutton, R. S. & Barto, A. G. (1998). *Reinforcement Learning : An Introduction*. Cambridge, MA : MIT Press.
- Tani, J., Ito, M. & Sugita, Y. (2004). Self-organization of distributedly represented multiple behavior schemata in a mirror system : reviews of robot experiments using rnnpb. *Neural Networks*, *17(8)*, 1273–1289.
- Thomas, E., Clerc, M., Carpentier, A., Daucé, E., Devlaminck, D. & Munos, R. (2013). Optimizing p300-speller sequences by rip-ping groups apart. In *Neural Engineering (NER), 2013 6th International IEEE/EMBS Conference on* (pp. 1062–1065).
- Thomas, E., Daucé, E., Devlaminck, D., Mahé, L., Carpentier, A., Munos, R., Perrin, M., Maby, E., Mattout, J., Papadopoulo, T. & Clerc, M. (2014). Coadapt p300 speller : optimized flashing sequences and online learning. In *proceedings of the 6th International Brain-Computer Interface Conference*. Graz, Austria.
- Thorpe, S. J., Fize, D. & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*, 520–522.
- Treves, A. & Rolls, E. T. (1994). Computational analysis of the role of the hippocampus in memory. *Hippocampus*, *4(3)*, 374–391.
- Tsuda, I. (1992). Dynamic link of memory - chaotic memory map in non equilibrium neural networks. *Neural Networks*, *5*, 313–326.

- Tuckwell, H. (1988). *Introduction to Theoretical Neurobiology*. Cambridge University Press.
- Van Rossum, M. C., Bi, G. Q. & Turrigiano, G. G. (2000). Stable hebbian learning from spike timing-dependent plasticity. *The Journal of Neuroscience*, *20*(23), 8812–8821.
- Van Vreeswijk, C. & Sompolinsky, H. (1998). Chaotic balanced state in a model of cortical circuits. *Neural Computation*, *10*, 1321–1371.
- Vapnik, V. N. (1995). *The nature of statistical learning theory*. Springer Verlag.
- Vapnik, V. N. (1998). *Statistical learning theory*, Volume 1. Wiley New York.
- Varela, F., Lachaux, J.-P., Rodriguez, E. & Martinerie, J. (2001). The brainweb : phase synchronization and large-scale integration. *Nature reviews neuroscience*, *2*(4), 229–239.
- Varela, F., Thompson, E. & Rosch, E. (1991). *The Embodied Mind*. MIT Press.
- Van de Ville, D., Britz, J. & Michel, C. M. (2010). Eeg microstate sequences in healthy humans at rest reveal scale-free dynamics. *Proceedings of the National Academy of Sciences*, *107*(42), 18179–18184.
- Voges, N. & Perrinet, L. (2012). Complex dynamics in recurrent cortical networks based on spatially realistic connectivities. *Frontiers in computational neuroscience*, *6*.
- Waitzman, D. M., Silakov, V. L. & Cohen, B. (1996). Central mesencephalic reticular formation (cmrf) neurons discharging before and during eye movements. *J Neurophysiol*, *75*(4), 1546–1572.
- Wald, A. (1947). *Sequential analysis*. Mineola, NY : Dover Phoenix Editions.
- Wang, X.-J. (2010). Neurophysiological and computational principles of cortical rhythms in cognition. *Physiological reviews*, *90*(3), 1195–1268.
- Wang, X.-J. & Buzsáki, G. (1996). Gamma oscillation by synaptic inhibition in a hippocampal interneuronal network model. *The journal of Neuroscience*, *16*(20), 6402–6413.
- Weisbuch, G. (1989). *Dynamique des systèmes complexes*. InterEditions/Editions du CNRS.
- Wessberg, J., Stambaugh, C. R., Kralik, J. D., Beck, P. D., Laubach, M., Chapin, J. K., Kim, J., Biggs, S. J., Srinivasan, M. A. & Nicolelis, M. A. (2000). Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. *Nature*, *408*(6810), 361–365.
- Wiener, N. (1965). *Cybernetics or Control and Communication in the Animal and the Machine*, Volume 25. MIT press.
- Williams, R. J. (1992). Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine Learning*, *8*, 229–256.
- Wilson, H. R. & Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical journal*, *12*(1), 1–24.
- Wilson, M. A. & Bower, J. M. (1989). The simulation of large-scale neural networks. *Methods in neuronal modeling : From synapses to networks*, MIT Press (pp. 291–333).

- Wolpert, D. M. & Ghahramani, Z. (2000). Computational principles of movement neuroscience. *nature neuroscience*, *3*, 1212–1217.
- Wolpert, D. M., Miall, R. C. & Kawato, M. (1998). Internal models in the cerebellum. *Trends in cognitive sciences*, *2(9)*, 338–347.
- Wright, W. A., Smith, R. E., Danek, M. & Greenway, P. (2001). A generalisable measure of self-organisation and emergence. In Dorffner, G., Bischof, H. & Hornik, K. (Eds.), *Artificial Neural Networks - ICANN 2001*.
- Wurtz, R. H. (2008). Neuronal mechanisms of visual stability. *Vision research*, *48(20)*, 2070–2089.

