



HAL
open science

Mécanismes de la sélection de l'action et de la prise de décision dans les ganglions de la base : approche par un modèle connexionniste.

Charlotte Héricé

► **To cite this version:**

Charlotte Héricé. Mécanismes de la sélection de l'action et de la prise de décision dans les ganglions de la base : approche par un modèle connexionniste.. Neurosciences [q-bio.NC]. Université de Bordeaux, 2016. Français. NNT : 2016BORD0234 . tel-01448194

HAL Id: tel-01448194

<https://theses.hal.science/tel-01448194>

Submitted on 27 Jan 2017

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

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

THÈSE PRÉSENTÉE
POUR OBTENIR LE GRADE DE
**DOCTEUR DE
L'UNIVERSITÉ DE BORDEAUX**

ÉCOLE DOCTORALE SCIENCES DE LA VIE ET DE LA SANTÉ
SPÉCIALITÉ NEUROSCIENCES

Par Charlotte HÉRICÉ

MÉCANISMES DE LA SÉLECTION DE L'ACTION ET DE LA PRISE
DE DÉCISION DANS LES GANGLIONS DE LA BASE
Approche par un modèle connexionniste

Sous la direction de : André GARENNE

Soutenue le 21 novembre 2016 devant le jury composé de :

Pr. Gwendal LE MASSON	P.U-P.H	I.N.S.E.R.M.	Président du jury
Dr. Fabrice WENDLING	D.R	I.N.S.E.R.M.	Rapporteur
Pr. Pierre CHAUVET	P.U	L.A.R.I.S.	Rapporteur
Dr. André GARENNE	M.C.U	I.N.R.I.A.	Directeur de thèse
Pr. Jean-Marie CABELGUEN	P.U	I.N.S.E.R.M.	Invité
Dr. Marie BEURTON-AIMAR	M.C.U	La.B.R.I.	Invité
Dr. Rémi DUBOIS	M.C.U	L.I.R.Y.C.	Tuteur

Résumé

Titre Mécanismes de la sélection de l'action et de la prise de décision dans les ganglions de la base : approche par un modèle connexionniste.

Résumé Les structures du système nerveux responsables des modalités de la prise de décision forment un circuit constitué par les ganglions de la base, le cortex, le thalamus et leurs nombreuses interconnexions. Ce circuit peut être décrit comme un ensemble de boucles fonctionnant en parallèle et interagissant en différents points. Des interactions entre ces boucles et de la plasticité de leurs connexions émergent les choix et donc les actions d'un individu. Ces comportements émergents et les phénomènes d'apprentissage qui en découlent sont abordés à travers une approche en boucle fermée dans laquelle le modèle théorique est en interaction constante avec l'environnement où se déroule la tâche comportementale étudiée. À cette fin, des outils de modélisation neuronale et d'analyse dédiés ont été développés dans le laboratoire d'accueil. Nous explorons donc ici la dynamique des flux d'information au sein de ce circuit à travers un modèle computationnel décrit à l'échelle du neurone et de la synapse. À partir d'observations expérimentales préalables réalisées sur le primate et de modèles computationnels antérieurs nous avons développé de manière incrémentale un réseau capable d'apprendre à réaliser les tâches comportementales dans plusieurs protocoles et conditions. Le résultat obtenu ici est un modèle computationnel d'apprentissage et de prise de décision dans les ganglions de la base qui permet de tester des hypothèses expérimentales et d'effectuer des investigations physiopathologiques ou pharmacologiques *in silico* à l'échelle cellulaire. Le développement de ce modèle computationnel a été mené en parallèle avec l'étude expérimentale d'un protocole de prise de décision et la mise au point d'un modèle de maladie de Parkinson chez la salamandre (*Pleurodeles waltlii*).

Mots-clés Neurosciences computationnelles, connexionnisme, ganglions de la base, prise de décision, boucle fermée

Abstract

Title Mechanism of action selection and decision-making in the basal ganglia through a connectionist model approach

Abstract The nervous system structures involved in decision making constitute a circuit formed by the basal ganglia, the cortex, the thalamus and their numerous interconnections. This circuit can be described as a set of loops operating in parallel and interacting at different points. The decisions and therefore the actions of an individual emerge from the interactions between these loops and the plasticity of their connections. These emerging behaviors and arising learning processes are addressed through a closed-loop approach in which the theoretical model is in constant interaction with the environment of the task. To this end, neural modeling and dedicated analysis software tools were developed in the laboratory. We explore here the dynamics of information flows within this circuit through a computational model described at the neuron and synapse level. Taking into account previous experimental observations from primates and earlier computational models, we incrementally developed a network capable of learning to perform behavioral tasks under several protocols and conditions. The result here is a computational model of learning and decision making in the basal ganglia that allows for the testing of experimental hypotheses and also to conduct *in silico* pathophysiological or pharmacological investigations at the cellular level. The development of this computational model was conducted in parallel with the development of an experimental protocol of decision making and with the adjustment of a model of Parkinson disease in the salamander (*Pleurodeles waltlii*).

Keywords Computational neurosciences, connectionism, basal ganglia, decision-making, closed loop.

Unités de recherche

Institut des Maladies Neurodégénératives (I.M.N.)

Université de Bordeaux

C.N.R.S. U.M.R. 5293 - Bât. 3B 3^e étage

146, rue Léo Saignat

33076 Bordeaux - FRANCE

Équipe Mnémosyne

Institut National de Recherche en Informatique et en Automatique (I.N.R.I.A.)

Centre de Recherche Bordeaux - Sud-Ouest

200, avenue de la Vieille Tour

33405 Talence Cedex - FRANCE

Remerciements

JE voudrais tout d'abord remercier mon directeur de thèse André GARENNE pour m'avoir encadrée et donné cette formidable opportunité de travailler dans un domaine aussi passionnant.

Un grand merci également à tous mes collègues de l'équipe Mnémosyne pour leur soutien sans lequel ces trois années de thèse n'auraient pas été les mêmes. Je remercie également toutes les personnes avec qui j'ai pu collaborer durant ma thèse, notamment Thomas BORAUD, Jean-Marie CABELGUEN pour m'avoir permis d'apporter un aspect plus « biologique » à mon travail de recherche. De même, je tiens à remercier Manon BONNET-SAVE, Mathieu BOURDENX et Brice DE LA CROMPE pour m'avoir aidée lors des travaux avec les salamandres et sans qui ce travail n'aurait pas vu le jour. Un grand merci également à Marie BEURTON-AIMAR pour ses conseils pratiques et son soutien et ce, depuis le Master.

Je tiens aussi à remercier tous mes collègues (doctorants ou non) de l'IMN, de l'INRIA pour nos discussions inspirantes et leur amitié, ainsi que toutes les personnes avec qui j'ai eu la possibilité d'enseigner et qui m'ont guidée lors de cette nouvelle expérience.

Enfin, je remercie ma famille et mes amis pour leurs encouragements incessants et leur compréhension depuis le début de mes études et pour leur support inconditionnel.

Bordeaux, le 11 octobre 2016.

Table des matières

1	Introduction et contexte de l'étude	1
1.1	Structures impliquées dans les processus de prise de décision	2
1.1.1	Introduction	2
1.1.2	Les ganglions de la base	2
1.1.2.1	Organisation anatomique	3
1.1.2.2	Boucles anatomo-fonctionnelles	6
1.1.3	Le cortex	8
1.1.3.1	Anatomie	8
1.1.3.2	Connectivité	10
1.1.3.3	Évolution structurelle	10
1.1.4	Le thalamus	12
1.1.4.1	Anatomie	12
1.1.4.2	Connectivité	12
1.1.5	Les principales pathologies des BG	12
1.2	Approche expérimentale de l'étude des mécanismes de la prise de décision . . .	17
1.2.1	Principes généraux des protocoles utilisés	17
1.2.2	Supports expérimentaux de notre étude	18
1.2.2.1	Travaux de Pasquereau et al. (2007)	18
1.2.2.2	Travaux de Piron et al. (2016)	21
1.2.2.3	Modèles d'étude chez l'amphibien	23
1.3	Les différents types d'apprentissage et leurs corrélats biologiques	25
1.3.1	Apprentissage supervisé	27
1.3.1.1	Principe	27
1.3.1.2	Apprentissage fondé sur l'erreur dans le cervelet	27
1.3.2	Apprentissage par renforcement	29
1.3.2.1	Principe	29
1.3.2.2	Apprentissage par renforcement dans les BG	29
1.3.3	Apprentissage non-supervisé	31
1.3.3.1	Principe	31
1.3.3.2	Apprentissage non-supervisé dans le cortex	32
1.4	Approches théoriques et computationnelles des mécanismes de la prise de décision	33
1.4.1	Méthodes et outils pour la modélisation	33
1.4.1.1	Principe des réseaux de neurones artificiels	33

1.4.1.2	Modèles mathématiques neuronaux et synaptiques	36
1.4.1.3	Outils de simulation	41
1.4.2	Antériorité des modèles computationnels des BG	43
1.4.2.1	Modèle d'Albin	43
1.4.2.2	Modèle de Gurney (2001)	45
1.4.2.3	Modèle de Leblois (2006)	45
1.4.2.4	Modèle de Humphries (2006)	47
1.4.2.5	Modèle de Guthrie (2013)	48
1.4.2.6	Travaux récents	50
1.4.2.7	Synthèse et choix méthodologiques	51
2	Développement et étude d'un modèle connexionniste des BG : application à la simulation d'une tâche d'apprentissage dans un protocole de prise de décision	54
2.1	Introduction	55
2.2	Matériel & Méthodes	56
2.2.1	Tâches comportementales étudiées	56
2.2.2	Développement des outils de simulation et d'analyse	57
2.2.3	Implémentation du circuit	57
2.2.3.1	Architecture du modèle	57
2.2.3.2	Organisation du réseau	59
2.2.3.3	Paramètres du modèle	60
2.2.4	Implémentation de la plasticité synaptique	63
2.2.5	Développement et mise au point du réseau	64
2.2.5.1	Analyse de l'activité corticale	64
2.2.5.2	Paramétrage et comportement du modèle	66
2.2.5.3	Sélection de l'action dynamique rupture de la symétrie	68
2.3	Résultats	70
2.3.1	Étude de sensibilité	70
2.3.1.1	Effet du bruit synaptique	70
2.3.1.2	Amplitude de l'activité basale	70
2.3.2	Tests lésionnels	71
2.3.3	Validation dans le protocole initial (Pasquereau et al., 2007)	73
2.3.3.1	Propriétés d'apprentissage	73
2.3.3.2	Protocole d'inversion de l'apprentissage	77
2.3.3.3	Protocole d'extinction	79
2.3.4	Validation dans le protocole complémentaire (Piron et al., 2016)	81
2.3.4.1	Sélection préférentielle de la cible optimale	81
2.3.4.2	Diminution du délai d'initiation du mouvement avec d'apprentissage	81
2.3.4.3	Effet de la lésion du GPi sur les capacités d'apprentissage du modèle	82
2.4	Discussion	83

2.4.1	Apports et prédictions du modèle	83
2.4.1.1	Existe-t-il une échelle idéale pour suivre les flux d'information ?	84
2.4.1.2	Comparaison des mécanismes de sélection de l'action	86
2.4.1.3	Deux niveaux pour la prise de décision et la sélection de l'action	87
2.4.1.4	Convergence et divergence dans l'architecture des BG	88
2.4.1.5	Apprentissage dans le réseau	88
2.4.2	Limites et perspectives de l'approche	89
3	Mise au point d'un protocole de prise de décision chez <i>Pleurodeles waltlii</i> en condition physiologique et à l'aide d'un modèle parkinsonien	91
3.1	Introduction	92
3.2	Choix d'un modèle animal : l'urodèle	92
3.3	Matériel et méthodes	93
3.3.1	Dispositif d'étude comportementale	93
3.3.1.1	Animaux	93
3.3.1.2	Bassin d'entraînement	94
3.3.1.3	Arène	95
3.3.1.4	Test comportemental	95
3.3.2	Métriques et études des effets des lésions dopaminergiques	97
3.3.2.1	Enregistrements vidéo	97
3.3.2.2	Traitement des données	97
3.3.2.3	Lésions dopaminergiques	98
3.3.3	Immunohistochimie	100
3.4	Résultats	102
3.4.1	Protocole de prise de décision	102
3.4.1.1	Entraînement	102
3.4.1.2	Test en arène	102
3.4.2	Lésions dopaminergiques	105
3.4.2.1	Analyses vidéos	105
3.4.2.2	Observations au microscope à fluorescence	108
3.5	Discussion	109
3.5.1	Synthèse des travaux	109
3.5.2	Points sensibles des méthodes d'analyse	109
3.5.2.1	Effet du champ magnétique terrestre	109
3.5.2.2	Apathie d'après lésion	110
3.5.2.3	Analyses vidéos	110
3.5.2.4	Immunofluorescence	110
3.5.3	Perspectives	110
4	Discussion générale	112
4.1	Résumé des travaux	113
4.1.1	Apports de l'approche de modélisation	113
4.1.1.1	Implémentation du réseau	113

4.1.1.2	Bruit et activité basale dans le réseau	114
4.1.1.3	Apprentissage dans le réseau	114
4.1.1.4	Effet de l'apprentissage sur le temps de réaction du modèle .	114
4.1.1.5	Rôle de l'inhibition striatale latérale	115
4.1.1.6	Protocole d'inversion de l'apprentissage	115
4.1.1.7	Protocole d'extinction	115
4.1.1.8	Tests lésionnels réalisés <i>in silico</i>	116
4.1.2	Apports de l'approche expérimentale	116
4.1.2.1	Conditionnement fonctionnel et reproductible	116
4.1.2.2	Protocole lésionnel fonctionnel	116
4.1.2.3	Mise en place de méthodes d'analyse	117
4.2	Perspectives et compléments d'étude	117
4.2.1	Rôles de l'inhibition latérale au niveau thalamique et cortical	117
4.2.2	Ajout de structures associées aux BG	117
4.2.3	Voie indirecte dans le modèle computationnel	118
4.2.4	Influence du champ magnétique terrestre	118
4.2.5	Amélioration du suivi de la récupération des animaux	118
4.2.6	Apathie des animaux après chirurgie	119
4.2.7	Implémentation d'un modèle computationnel des BG des pleurodèles .	119

5 Conclusion 120

A Annexes xxii

A.1	Glossaire	xxiii
A.2	Circuiterie détaillée du modèle	xxiv
A.3	Tables supplémentaires	xxv
A.4	Publications	xxviii
A.5	Posters	xxviii

Table des figures

1.1	Localisation des BG chez l'humain (d'après Marc SAVASTA, INSERM Grenoble).	3
1.2	Organisation fonctionnelle des BG.	4
1.3	Schéma de la connectivité des BG avec les trois voies principales.	7
1.4	Taille et organisation du cortex chez différentes espèces.	9
1.5	Les interactions entre les ganglions la base et le télencéphale sont proportionnelles au rôle de ce dernier dans le contrôle des comportements.	11
1.6	Relations entre les principaux syndromes neuropathologiques observés dans les BG et les altérations de leur connectivité.	14
1.7	Approche en boucle fermée	17
1.8	Schéma de la tâche comportementale à quatre cibles.	19
1.9	Exemples de trajectoires du mouvement.	20
1.10	Schéma de la tâche comportementale à deux cibles.	22
1.11	L'injection stéréotaxique de 6-OHDA cause une perte cellulaire dopaminergique dans le diencéphale et le mésencéphale.	24
1.12	Spécialisation du cervelet, des BG et du cortex en fonction des types d'apprentissage.	26
1.13	Diagramme du circuit cérébelleux.	28
1.14	Modèle hypothétique de l'apprentissage par renforcement dans les BG.	30
1.15	Diagramme du circuit neuronal cortical.	32
1.16	Comparaison entre un neurone humain et un neurone artificiel.	34
1.17	Comparaison de réseaux de neurones biologiques et artificiels.	35
1.18	Modèle initial du neurone formel.	36
1.19	Circuit équivalent à la membrane du modèle LIF.	38
1.20	Exemples d'activité de neurones d'Izhikevich.	39
1.21	Schéma de la connectivité du modèle des BG de Albin.	44
1.22	Adaptation du schéma de la connectivité du modèle des BG.	46
1.23	Architecture du modèle des BG de Leblois.	47
1.24	Anatomie fonctionnelle des BG du modèle d'Humphries.	48
1.25	Architecture du modèle des BG de Guthrie.	49
2.1	Architecture du modèle computationnel des BG et structures associées.	58
2.2	Exemple de l'architecture détaillée du modèle pour le cortex et le striatum.	60
2.3	Exemple de simulation de deux cellules interconnectées.	62
2.4	Illustration des notions de GC et GD à travers un exemple.	66

2.5	Fréquences moyennes d'activité en fonction des sous-régions.	67
2.6	Exemples de taux de décharge moyens enregistrés dans des sous-régions du cortex.	68
2.7	Taux moyen de récompense et de GC (« Bon Choix »).	69
2.8	Effet du bruit et du courant de base sur les fréquences de prise de décision.	71
2.9	Effet de lésions spécifiques sur l'activité de chaque population du réseau.	72
2.10	Effet de l'apprentissage sur les capacités d'apprentissage du modèle lors du protocole à quatre cibles.	73
2.11	Effet de l'apprentissage sur les poids et le comportement du modèle.	74
2.12	Délai moyen de déclenchement du mouvement.	75
2.13	Effets de l'absence de SLI sur les taux de réussite et sur la performance d'apprentissage du modèle.	76
2.14	Résultats des simulations pour les trois étapes du protocole d'inversion de l'apprentissage	79
2.15	Résultats des simulations du protocole d'extinction	80
2.16	Etude des capacités d'apprentissage du modèle sur un protocole à deux cibles.	82
2.17	Délai moyen de déclenchement du mouvement après apprentissage lors du protocole à deux cibles.	82
2.18	Effet d'une lésion du GPi sur les performances d'apprentissage du modèle.	83
3.1	Schéma du bassin d'entraînement.	94
3.2	Schéma de l'arène.	95
3.3	Arène vue du dessus.	96
3.4	Lésions à la 6-OHDA au niveau des neurones dopaminergiques chez la salamandre.	99
3.5	Modes de locomotion comparés du pleurodèle et de la lamproie.	100
3.6	Groupe d'animaux naïfs testés sous lumière non polarisée et polarisée.	102
3.7	Groupe d'animaux entraînés testés sous lumière non polarisée et polarisée.	103
3.8	Effectifs d'animaux par décile d'angles.	104
3.9	Temps d'exécution de la tâche comportementale en fonction des conditions.	105
3.10	Évolution du nombre de pas moyen après lésion dopaminergique.	106
3.11	Évolution du nombre de cycles de nage moyen après lésion dopaminergique.	106
3.12	Évolution de la durée moyenne des pas après lésion dopaminergique.	107
3.13	Évolution de la durée moyenne des cycles de nage après lésion dopaminergique.	108
3.14	Cliché de microscopie à fluorescence montrant les marquage DAPI/TH des neurones du mésencéphale/diencéphale des pleurodèles en condition contrôle et un mois après une lésion.	109
A.1	Circuiterie détaillée du modèle.	xxiv
A.2	Poster présenté lors du « <i>Fourth International Symposium on Biology of Decision Making</i> » - Paris, France - Mai 2014.	xxix
A.3	Poster présenté lors du « <i>Fifth International Symposium on Biology of Decision Making</i> » - Paris, France - Mai 2015	xxx
A.4	Poster présenté lors de « <i>Multi-disciplinary Conference on Reinforcement Learning and Decision Making</i> » - Edmonton, Canada - Juin 2015	xxxi

A.5	Poster présenté lors des Rencontres INRIA-Industries (R2I) - Bordeaux, France - Octobre 2015	xxxii
A.6	Poster présenté lors du « <i>Sixth International Symposium on Biology of Decision Making</i> » - Paris, France - Mai 2016	xxxiii

Liste des Tables

1.1	Combinaison des couples de cibles affichées à l'écran durant la tâche comportementale.	20
1.2	Résumé des modèles computationnels des BG.	52
2.1	Nombre de neurones et de synapses par sous-populations dans le modèle.	61
2.2	Poids synaptiques et délais utilisés pour la connectivité entre les sous-populations du réseau. M : région motrice ; C : région cognitive ; A : région associative.	63
2.3	Paramètres du protocole d'inversion de l'apprentissage.	78
3.1	Présentation des conditions testées	97
A.1	Détails de l'environnement de travail lors du développement du modèle et de l'analyse des résultats.	xxv
A.2	Détails des numéros des neurones et sous-populations neuronales du modèle (1/3).xxv	
A.3	Détails des numéros des neurones et sous-populations neuronales du modèle (2/3).xxvi	
A.4	Détails des numéros des neurones et sous-populations neuronales du modèle (3/3).xxvii	
A.5	Résumé du modèle.	xxvii

Abréviations

Termes scientifiques (1/2)

Abréviation	Français	Anglais
AFR	Taux de Décharge Moyen	<i>Average Firing Rate</i>
ANN	Réseau de Neurones Artificiels	<i>Artificial Neuron Network</i>
BC	Courant de Base	<i>Basal Current</i>
BG	Ganglions de la Base	<i>Basal Ganglia</i>
CNS	Système Nerveux Central	<i>Central Nervous System</i>
DA	Dopamine	<i>Dopamine</i>
GABA	Acide γ -AminoButyrique	<i>γ-AminoButyric Acid</i>
DBS	Stimulation Cérébrale Profonde	<i>Deep Brain Stimulation</i>
dMSN	Neurone Épineux Moyen direct	<i>direct Medium Spiny Neuron</i>
FNM	Modèle de FitzHugh-Nagumo	<i>FitzHugh-Nagumo Model</i>
FSI	Interneurone à Décharge Rapide	<i>Fast-Spiking Interneurons</i>
GC	Bon Choix	<i>Good Choice</i>
GD	Bonne Décision	<i>Good Decision</i>
Glu	Glutamate	<i>Glutamate</i>
GP	Globus Pallidus	<i>Globus Pallidus</i>
GPe	Segment externe du Globus Pallidus	<i>Globus Pallidus external segment</i>
GPi	Segment interne du Globus Pallidus	<i>Globus Pallidus internal segment</i>
HHM	Modèle de Hodgkin-Huxley	<i>Hodgkin-Huxley Model</i>
HPP	Motif de Projection « Horizontal »	<i>« Horizontal » Projection Pattern</i>
iMSN	Neurone Épineux Moyen indirect	<i>indirect Medium Spiny Neuron</i>
LI	Inhibition Latérale	<i>Lateral Inhibition</i>
LIF	« Intégrateur à Fuite »	<i>Leaky Integrate-and-Fire</i>
LTD	Dépression à Long Terme	<i>Long Term Depression</i>
LTP	Potentialisation à Long Terme	<i>Long Term Potentiation</i>
MAS	Système Multi-Agent	<i>Multi-Agent System</i>
MLM	Modèle de Morris-Lecar	<i>Morris-Lecar Model</i>
MSN	Neurone Épineux Moyen	<i>Medium Spiny Neuron</i>
NAc	Noyau Accumbens	<i>Nucleus Accumbens</i>
NC	Condition de Nouveauté	<i>Novelty Condition</i>
NMR	Résonance Magnétique Nucléaire	<i>Nuclear Magnetic Resonance</i>
ODE	Équation Différentielle Ordinaire	<i>Ordinary Differential Equation</i>
PFC	Cortex Pré-Frontal	<i>Pre-Frontal Cortex</i>
RC	Condition de Routine	<i>Routine Condition</i>
RevL	Apprentissage Inverse	<i>Reversal Learning</i>
RL	Apprentissage par Renforcement	<i>Reinforcement Learning</i>

Termes scientifiques (2/2)

Abréviation	Français	Anglais
SB	Rupture de la Symmétrie	<i>Symmetry Breaking</i>
SLI	Inhibition Latérale Striatale	<i>Striatal Lateral Inhibition</i>
SN	Substance Noire	<i>Substantia Nigra</i>
SNe	Substance Noire compacte	<i>Substantia Nigra pars compacta</i>
SNN	Réseau de Neurones à Spikes	<i>Spiking Neural Network</i>
SNr	Substance Noire Réticulée	<i>Substantia Nigra pars reticulata</i>
spike	Potentiel d'Action	<i>Spike</i>
STDP	Plasticité Dépendante du Temps	<i>Spike-Timing Dependent Plasticity</i>
STN	Noyau Sous-Thalamique	<i>SubThalamic Nucleus</i>
TAN	Neurone à Activité Tonique	<i>Tonically Active Neuron</i>
TDL	Apprentissage par Différence Temporelle	<i>Temporal Difference Learning</i>
TOC	Trouble Obsessionnel Compulsif	<i>Obsessive-Compulsive Disorder</i>
VTA	Aire Ventrale Tegmentale	<i>Ventral Tegmental Area</i>
VPP	Motif de Projection « Vertical »	<i>« Vertical » Projection Pattern</i>

Sigles d'organismes et de conférences

Abréviation	Français	Anglais
CNRS	Centre National de la Recherche Scientifique	<i>National Center of Scientific Research</i>
IMN	Institut des Maladies Neurodégénératives	<i>Institute of Neurodegenerative Diseases</i>
INRIA	Institut National de la Recherche en Informatique et en Automatique	<i>National Institute of Informatics and Automatics Research</i>
INSERM	Institut National de la Santé et de la Recherche Médicale	<i>National Institute of Health and Medical Research</i>
UMR	Unité Mixte de Recherche	<i>Mixed Research Unit</i>

Préface

Cette thèse s'est effectuée au sein de l'équipe Mnémosyne de l'Institut des Maladies Neurodégénératives de Bordeaux. Le travail de recherche de cette équipe s'appuie sur les outils et les concepts des neurosciences intégratives et computationnelles. Son objectif est de modéliser le cerveau comme un système de mémoires actives en synergie et en interaction avec les mondes interne et externe et de le simuler comme un tout et en situation. Ce travail s'inscrit précisément dans ce cadre de recherche. En effet, il a porté sur l'étude des mécanismes de sélection de l'action et de prise de décision dans les ganglions de la base (BG). De façon plus précise, nous nous sommes attachés à mettre en évidence les relations existant entre ces mécanismes et la circuiterie des BG décrite à l'échelle microscopique (neurone et synapse).

Avancées

Le point de départ de ce travail de modélisation est une étude comportementale menée chez le primate, montrant le rôle des BG dans une tâche de prise de décision [Pasquereau et al., 2007] à l'aide d'enregistrements électrophysiologiques et comportementaux. Suite à cette étude chez l'animal, un premier modèle computationnel implémentant cette tâche a été développé [Guthrie et al., 2013]. Reprenant une partie des hypothèses du modèle de [Leblois et al., 2006a], le modèle de Guthrie a été développé à l'échelle mésoscopique, c'est-à-dire que chacune de ses unités de base est une population neuronale complète dont la dynamique est décrite par une seule équation différentielle. Cette échelle de description intermédiaire permet de développer rapidement des modèles dynamiques de populations interconnectées et surtout de les paramétrer assez facilement en raison d'un nombre relativement limité de paramètres. En revanche elle limite *de facto* les possibilités d'exploration du modèle, notamment car elle rend difficile les investigations sur les mécanismes décrits à l'échelle cellulaire et synaptique (ce que nous nommons ici approche connexionniste). Or c'est à ces échelles qu'il peut devenir particulièrement intéressant de tester les hypothèses neurobiologiques de l'effet des drogues, de la neurodégénérescence de certains types neuronaux ou encore les théories reliant plasticité synaptique et apprentissage. La première étape de notre travail de recherche a donc consisté à implémenter un modèle computationnel de réseau de neurones à *spikes*¹ permettant de décrire au moins partiellement la circuiterie des BG et ses propriétés dynamiques à l'échelle microscopique. La seconde étape de développement a consisté à simuler le réseau ainsi obtenu à l'aide de notre outil logiciel afin de reprendre, d'adapter puis de développer le précédent modèle en utilisant ce nouveau formalisme. Une fois que les paramètres de ce modèle ont été ajustés, il a été soumis à une nouvelle tâche comportementale

1. Spike ou potentiel d'action

récemment étudiée au laboratoire d'accueil [Piron et al., 2016] et a ainsi pu être validé dans un contexte plus large. Dans une troisième étape, des simulations complémentaires ont montré que le modèle se comportait de façon cohérente quand il était soumis à des lésions focalisées et à des variantes des protocoles comportementaux initiaux. Le nouveau modèle ainsi obtenu reprend donc les résultats des versions antérieures, les généralise à de nouveaux protocoles expérimentaux et nous permet de tester les effets des mécanismes observés à l'échelle cellulaire ou synaptique sur la dynamique du réseau. Il contribue donc à combler le gap toujours très vaste qui existe entre les capacités de traitement de l'information des neurones considérés individuellement et les propriétés fonctionnelles des réseaux des BG [Yin and Knowlton, 2006].

Difficultés

Les difficultés inhérentes à ce genre d'étude sont principalement de deux ordres : (i) le choix de ce qu'il est pertinent d'inclure ou non dans le modèle en terme de données biologiques et de degré de description [Wendling et al., 2016] et (ii) la gestion d'un espace des paramètres de grande dimension lors de son implémentation. Comme cette étude prolonge un travail précédemment effectué dans le même laboratoire, le point (i) a été assez rapidement surmonté en s'appuyant sur les publications et les interactions avec les chercheurs du laboratoire d'accueil ainsi que sur nos ressources de calcul. Le point (ii) en revanche est bien connu des modélisateurs pour être délicat et il a nécessité de longues périodes de simulations et d'analyses de données pour pouvoir être résolu.

Perspectives

En parallèle avec le développement du modèle computationnel en cours, nous avons travaillé sur le développement d'un protocole novateur de prise de décision sur le pleurodèle de Waltl, en collaboration avec Thomas Boraud de l'IMN et Jean-Marie Cabelguen de l'Institut Magendie qui possède une longue expertise sur ce modèle animal. Une fois validé, ce protocole doit permettre de mieux comprendre les rôles respectifs qu'y jouent les différents constituants des BG. L'intérêt de poursuivre notre étude sur cet animal est double : (i) tout d'abord celui-ci possède des structures analogues aux BG qui ont une organisation similaire à celle des mammifères mais avec un niveau de complexité bien plus accessible quant à la taille des populations neuronales mises en jeu. Il devient donc théoriquement plus simple d'envisager des modèles computationnels à l'échelle neuronale tout en gardant, au moins partiellement, la singularité de ces structures et de leur rôle dans les mécanismes comportementaux qui nous intéressent. Avec le même modèle animal, il devient du même coup plus envisageable d'inclure secondairement les autres régions du système nerveux central (« *Central Nervous System* », CNS) pouvant jouer un rôle dans les processus de prise de décision et d'avoir ainsi une vue intégrée et dynamique de leur fonctionnement global. De plus (ii) le pleurodèle est désormais un modèle validé pour l'étude de la maladie de Parkinson [Parish et al., 2007]. Nous avons donc entrepris de reproduire à l'Institut Magendie les résultats de l'étude de Parish afin de les utiliser en parallèle avec le développement de la tâche décrite précédemment. L'objectif de ce travail expérimental est le développement d'un

modèle complet (physiologique et pathologique) qui soit plus simple à mettre en œuvre que chez les modèles de mammifères et à plus long terme, en raison de la relative simplicité du système nerveux du pleurodèle, la mise au point d'un modèle computationnel physiopathologique des BG à l'aide d'un formalisme décrit à l'échelle cellulaire.

Introduction et contexte de l'étude

« Neuroscience is by far the most exciting branch of science because the brain is the most fascinating object in the universe. Every human brain is different - the brain makes each human unique and defines who he or she is. »

Stanley B. Prusiner

Sommaire

1.1 Structures impliquées dans les processus de prise de décision	2
1.1.1 Introduction	2
1.1.2 Les ganglions de la base	2
1.1.3 Le cortex	8
1.1.4 Le thalamus	12
1.1.5 Les principales pathologies des BG	12
1.2 Approche expérimentale de l'étude des mécanismes de la prise de décision	17
1.2.1 Principes généraux des protocoles utilisés	17
1.2.2 Supports expérimentaux de notre étude	18
1.3 Les différents types d'apprentissage et leurs corrélats biologiques	25
1.3.1 Apprentissage supervisé	27
1.3.2 Apprentissage par renforcement	29
1.3.3 Apprentissage non-supervisé	31
1.4 Approches théoriques et computationnelles des mécanismes de la prise de décision	33
1.4.1 Méthodes et outils pour la modélisation	33
1.4.2 Antériorité des modèles computationnels des BG	43

1.1 Structures impliquées dans les processus de prise de décision

1.1.1 Introduction

CETTE thèse a pour thème général l'étude des mécanismes de la sélection de l'action et de la prise de décision dans les BG au travers d'un modèle connexionniste. Cette question nécessite, comme la plupart des études en neurosciences computationnelles, une approche multidisciplinaire, mêlant neurosciences et informatique. À ce titre, il sera donc question de modèles théoriques, de neurones artificiels, de tâches comportementales et d'apprentissage.

La faculté d'apprendre est sans aucun doute la propriété la plus attractive des réseaux de neurones. Selon Claude TOUZET, l'apprentissage est « *une phase de développement d'un réseau de neurones durant laquelle le comportement du réseau est modifié jusqu'à l'obtention du comportement désiré.* » [Touzet, 1992]. L'apprentissage via les réseaux de neurones fait donc appel à l'expérience, à des exemples de comportement.

La prise de décision est un processus permettant à un organisme de choisir ses actions afin de maximiser les bénéfices résultants [Lee et al., 2012]. Pour expliquer ce processus, il convient tout d'abord d'envisager les options possibles car sans alternative il n'y a pas de sélection possible. Une fois les possibilités intégrées, un choix doit émerger. À ce moment, différents facteurs sont inclus dans l'analyse comme l'expérience, les connaissances, les contraintes, le but recherché ainsi que d'autres facteurs externes et internes. Enfin, une fois tous ces paramètres analysés, une décision est prise pour conduire - le plus souvent - à une action. L'apprentissage est en général la résultante de choix et de la prise en compte durable des conséquences de ces choix.

Le processus de prise de décision émerge de l'activité d'un réseau complexe formé par plusieurs structures cérébrales parmi lesquelles les BG, le thalamus et le cortex ont été identifiés comme jouant un rôle majeur. Nous allons donc présenter un aperçu synthétique de ces structures et de leurs interrelations afin de clarifier les notions de base qui seront étudiées et modélisées dans notre étude. Nous évoquerons ici des éléments d'anatomie fonctionnelle décrits chez le primate et dans une moindre mesure chez le rongeur.

1.1.2 Les ganglions de la base

Les ganglions de la base (BG) - aussi appelés « *Noyaux Gris Centraux* » dans la littérature scientifique francophone du siècle dernier - sont historiquement connus pour leur implication dans l'exécution et la planification des mouvements. Ils sont désormais également connus pour leurs rôles dans les comportements complexes nécessitant une coordination entre la cognition, la motivation et les mouvements [Bar-Gad et al., 2003]. Localisées au cœur du CNS, les structures des BG ont donc un rôle majeur dans le traitement de l'information. Cependant, même si leur connectivité a été précédemment décrite en grande partie [Gurney et al., 2001a, Hazrati et al., 1995], leur complexité est telle qu'il est difficile de les formaliser clairement.

Dans cette partie, nous verrons tout d'abord l'organisation anatomo-fonctionnelle des BG, les différences structurales en fonction des espèces et enfin les principales pathologies liées à leur dysfonctionnement.

1.1.2.1 Organisation anatomique

Les principaux composants des BG (Figure 1.1) sont le striatum, le Globus Pallidus (segments interne et externe, respectivement GPi et GPe), la substance noire (parties réticulée et compacte, respectivement SNr et SNc) et le Noyau Sous-Thalamique (« *Sub-Thalamic Nucleus* » en anglais, STN).

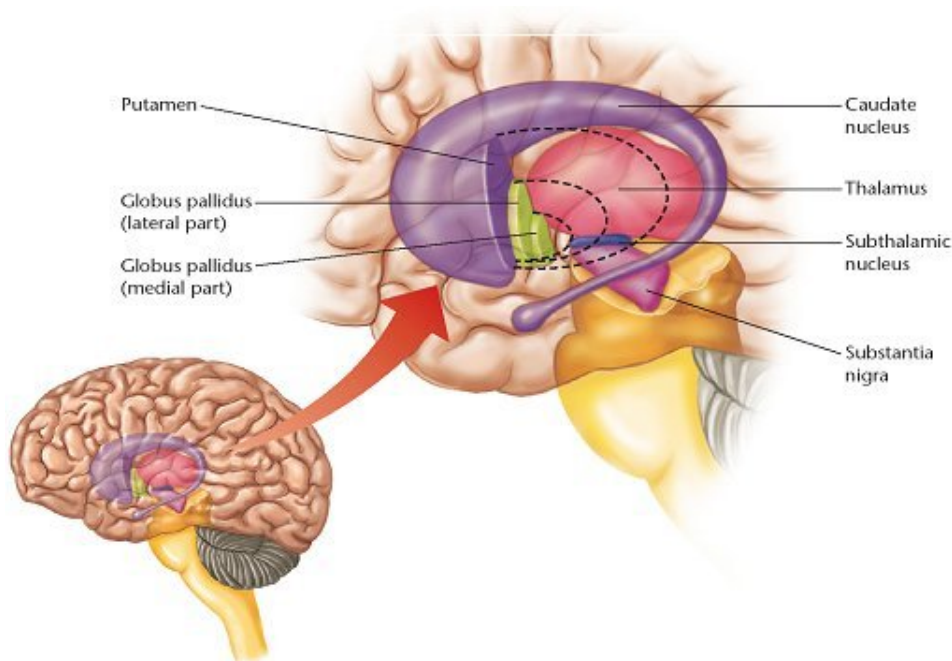


FIGURE 1.1 – Localisation des BG chez l’humain (d’après Marc SAVASTA¹, INSERM Grenoble). Violet foncé : putamen et noyau caudé (constituants du striatum), vert : globus pallidus (segments interne et externe), bleu : noyau sous-thalamique, violet clair : substance noire compacte.

L’organisation somatotopique dans le cortex cérébral est une notion importante dans la compréhension du fonctionnement des aires corticales et chaque structure des BG comporte également à son niveau une organisation et des projections de ce type que nous décrirons ici brièvement (pour de plus amples détails sur ce point il est possible de se référer aux articles de [Alexander et al., 1986] et de [Nambu, 2011]).

Le striatum Il s’agit de la principale structure d’entrée de l’information des BG. Il reçoit des afférences de toutes les aires corticales et est divisé en deux parties : l’une ventrale et l’autre dorsale. La partie dorsale comprend le noyau caudé et le putamen. La partie ventrale est, comme son nom l’indique, l’extension ventrale du striatum incluant notamment le noyau accumbens (NAc) ainsi que les portions médiale et ventrale du noyau caudé et du putamen [Bar-Gad et al., 2003].

Le striatum reçoit des projections glutamatergiques (excitatrices) de la part du cortex, du thalamus [Parent and Hazrati, 1995] ainsi que des projections dopaminergiques mésencéphaliques (SNc et Aire Ventrale Tegmentale - « *Ventral Tegmental Area* » en anglais, VTA). En effet,

1. <http://acces.ens-lyon.fr/acces/ressources/neurosciences/maladies-et-traitements/parkinson/causes/resolveuid/f6bc96ad216c4bdc456f3429f1b7b186>

les projections glutamatergiques du prosencéphale (cerveau antérieur) et les projections dopaminergiques du mésencéphale (cerveau moyen) convergent vers le striatum pour finir sur les projections des neurones épineux [Wickens, 1997]. Ces neurones de projection composent une vaste majorité des neurones striataux et sont nommés neurones épineux moyens (« *Medium Spiny Neurons* » en anglais, MSN). Le neurotransmetteur des MSN est le GABA (« *Gamma AminoButyric Acid* »), ce qui fait que les projections du striatum sont globalement inhibitrices [Bar-Gad et al., 2003].

Le striatum contient également des interneurones [Bar-Gad et al., 2003, Cohen and Frank, 2009, Wilson, 1995] de type cholinergique et gabaergiques. Les interneurones cholinergiques ont une activité de décharge spontanée et tonique et sont donc nommés neurones toniquement actifs (« *Tonically Active Neurons* », TAN), alors que les gabaergiques ont une activité de décharge rapide. Ces derniers semblent d'ailleurs jouer un rôle localement dans la modulation de l'activité des MSN néostriataux [Koós and Tepper, 1999]. Il a en effet été constaté expérimentalement que les interneurones gabaergiques pouvaient ralentir ou bloquer la génération de potentiels d'action dans des populations entières de MSN simultanément.

Le striatum se projette selon différentes voies ou trajets vers les régions de la substance noire et du globus pallidus (Figure 1.3). L'organisation et le rôle de ces voies seront abordés plus loin.

La partie latérale du putamen reçoit des entrées somatotopiques de la part du cortex moteur primaire et de la partie médiale de l'aire motrice supplémentaire. Les projections des régions orofaciales ainsi que celles des membres antérieurs et postérieurs du cortex moteur primaire et de l'aire motrice supplémentaire convergent vers la zone centrale médiolatérale du putamen (Figure 1.2). La partie la plus dorsomédiale reçoit des entrées de la part du cortex préfrontal [Nambu, 2011].

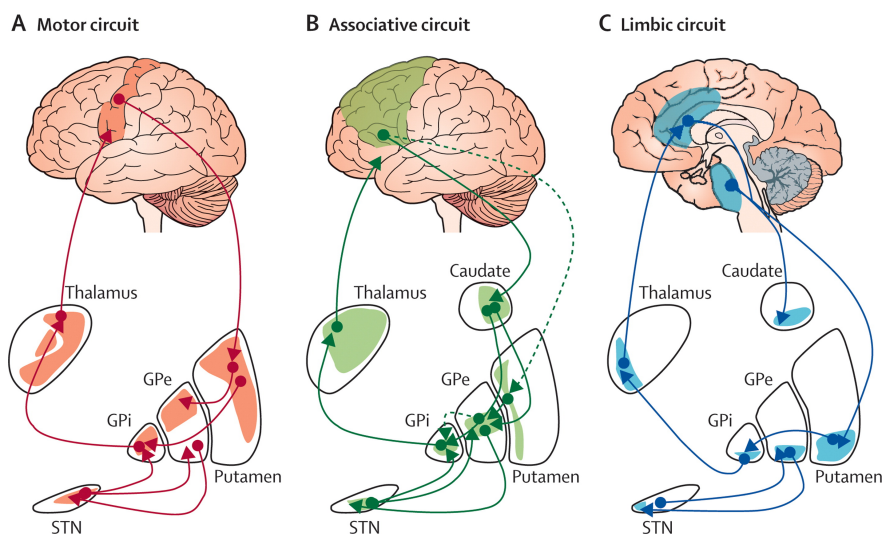


FIGURE 1.2 – Organisation fonctionnelle des BG. Sur cette représentation partielle de leur connectivité interne les BG sont subdivisés en trois sous-régions : motrice (A), associative (B) et limbique (C) qui sont topographiquement séparées en fonction de leurs afférences corticales avec le cortex moteur (rouge), le cortex préfrontal (vert) et le cortex cingulaire antérieur (bleu). GPe=globus pallidus pars externa. GPi=globus pallidus pars interna. STN=subthalamic nucleus (d'après [Obeso et al., 2008] et [Rodriguez-Oroz et al., 2009]).

Le Globus Pallidus Le globus pallidus est divisé en deux parties : *Globus Pallidus pars interna* (GPi) et *Globus Pallidus pars externa* (GPe). Le GPe est la région la plus latérale alors que le GPi est la plus médiale (et est souvent référencée comme le noyau entopédunculaire chez les rongeurs) [Utter and Basso, 2008]. Ces deux parties reçoivent des connexions inhibitrices de type gabaergique de la part des MSN du striatum.

Le GPe va se projeter vers le STN [Albin et al., 1989] ainsi que dans la SNr [Bar-Gad et al., 2003], créant de fortes connexions synaptiques inhibitrices gabaergiques avec le soma des neurones.

Le GPi reçoit donc des projections gabaergiques de la part du striatum et du GPe, ainsi que des projections glutamatergiques du STN. Il se projette sur le thalamus par des noyaux relais (noyaux thalamiques ventro-latéral, ventral antérieur, médio-dorsal et noyau habénulaire latéral) [Bar-Gad et al., 2003, Parent et al., 2001] et est considéré comme la principale voie de sortie des BG.

Tout comme le striatum, le neurotransmetteur majoritairement présent dans les projections des neurones pallidus est le GABA. Les neurones du GPi et du GPe ont beaucoup de similitudes anatomiques, physiologiques et biochimiques : ce sont presque tous des neurones avec une fréquence de décharge rapide mais avec des patterns de décharge différents [Bar-Gad et al., 2003].

Le GPe et le GPi ont tous les deux leurs propres représentations somatotopiques (Figure 1.2). La somatotopie des territoires de l'aire motrice supplémentaire est localisée au niveau rostrodorsal par rapport à son homologue dans les territoires du cortex moteur primaire. Un tiers de la partie dorsale du GPi et du GPe est occupé par le territoire préfrontal [Nambu, 2011].

Le Noyau Sous-Thalamique Le STN représente la deuxième structure d'entrée la plus importante des BG après le striatum, même s'il est plus petit que ce dernier en terme de volume et de nombre de neurones [Oorschot, 1996]. De par sa localisation et ses projections, il est considéré comme la principale station relais dans la voie indirecte [Alkemade et al., 2015] et également dans la voie hyperdirecte où il est impliqué dans l'inhibition de l'action [Alegre et al., 2013].

Il reçoit des projections glutamatergiques excitatrices de la part du cortex frontal et des aires somato-motrices [Nambu, 2011, Bar-Gad et al., 2003]. Il reçoit également toutes les afférences du GPe et projette lui-même à nouveau dans le GPe, le GPi et la SNr. Les projections du STN sont essentiellement excitatrices dans les BG [Utter and Basso, 2008].

La partie dorsolatérale du STN reçoit des entrées somatotopiques du cortex moteur primaire et de la partie médiale de l'aire motrice supplémentaire. Depuis le territoire ventral du STN jusqu'au moteur, nous pouvons retrouver les territoires oculomoteurs et préfrontaux. La partie la plus médiale du STN est occupée par le territoire limbique [Nambu, 2011, Alkemade et al., 2015].

La Substance Noire Les neurones de cette structure des BG sont organisés de façon à créer deux segments : un compact (*substantia nigra pars compacta* ou SNc) et un plus diffus (*substantia nigra pars reticulata* ou SNr).

Les neurones de la SNc ont une activité tonique faible et sont la première source de dopamine (DA) des BG. Ils projettent majoritairement le striatum, mais aussi le GPe, le GPi, le STN et la SNr [Parent et al., 2001, Utter and Basso, 2008].

En revanche dans la SNr, les neurones contiennent majoritairement du GABA. Ils possèdent des afférences striatales massives (via les voies directes et indirectes détaillées plus loin dans ce chapitre) et se projettent vers le thalamus (et secondairement vers le noyau pédonculopontin), vers le colliculus supérieur et le cortex frontal mais également vers la SNc. Ainsi, tout comme le GPi, la SNr est un noyau de sortie des BG. Dans la continuité du STN (via les afférences du cortex moteur primaire et de l'aire motrice supplémentaire) et comme le GPi, la SNr possède une organisation somatotopique mise en évidence notamment au niveau oro-facial [Nambu, 2011].

1.1.2.2 Boucles anatomo-fonctionnelles

Pour des raisons de clarté, nous allons rester dans ce paragraphe à un niveau de description à l'échelle de la région (cortex, thalamus, striatum ...). Le détail des différentes sous-régions sera abordé ultérieurement et au cours de la présentation du modèle.

La plupart des régions cérébrales sont connectées de manière bidirectionnelle les unes aux autres. Autrement dit, s'il y a une connexion partant d'une structure A pour aller vers une structure B, il y en a une autre allant de B vers A [Bar-Gad et al., 2003]. C'est le cas entre le cortex et les BG, ainsi qu'entre le thalamus et le cortex (boucles thalamo-corticales). Ces connexions vont se diviser en trois voies principales (Figure 1.3) pour former les boucles cortico-corticales (boucles fermées). En effet, ces trois voies commencent par le cortex (projections glutamatergiques excitatrices) et se terminent toutes par le GPi, principal point de sortie des BG et par la SNr [Plantinga et al., 2016]. Ces deux structures vont se projeter sur le thalamus par des connexions inhibitrices (gabaergiques) qui renvoient des connexions excitatrices (glutamatergiques) vers le cortex.

La voie directe Cette voie connecte le cortex au principal point d'entrée des BG, à savoir le striatum. Le striatum va ensuite se projeter sur le GPi et la SNr, qui va à son tour se projeter dans le thalamus. Ce dernier va finalement se projeter vers le cortex et réaliser ainsi cette première boucle. Les MSN striatonigraux de la voie directe expriment les récepteurs D1 à la DA [Lambot et al., 2016].

Les projections gabaergiques du striatum vers le GPi sont inhibitrices, et comme celles du GPi vers le thalamus le sont aussi, la voie directe va exercer une rétroaction globalement excitatrice, convergente et focalisée [Percheron et al., 1984]. Par exemple, plus l'inhibition du GPi par le striatum sera importante et moins celle du thalamus par le GPi le sera (et inversement) ce qui amène à dire que la voie directe exerce une action facilitatrice sur les sorties motrices [Dunovan and Verstynen, 2016].

De plus, la rétroaction positive exercée par la voie directe va permettre une désinhibition localisée, permettant l'amplification de toute différence d'activation entre des réseaux corticaux en compétition. Ce dernier élément va mécaniquement permettre la sélection d'une option spécifique parmi plusieurs alternatives [Guthrie et al., 2013].

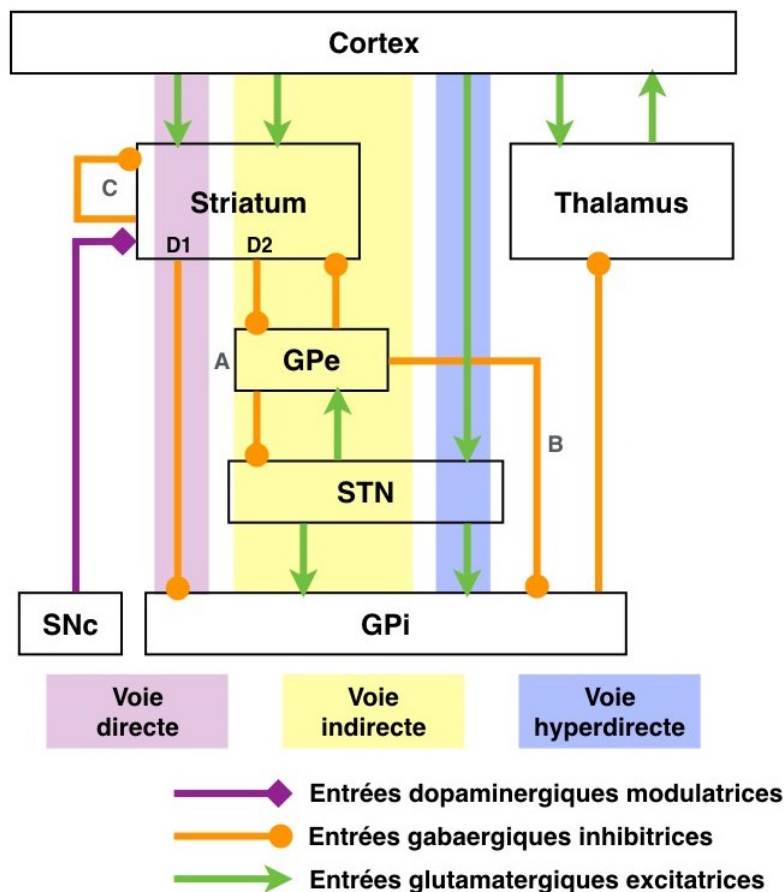


FIGURE 1.3 – Schéma de la connectivité des BG avec les trois voies principales réalisé d'après [Bar-Gad et al., 2003, Hamker, 2013, Smith et al., 1998]. Mauve : voie directe, jaune : voie indirecte, bleu : voie hyperdirecte. Les types de flèches précisent le neurotransmetteur principalement présent dans les connexions entre les structures d'intérêt. (A) Trajet indirect « long » (cortex→striatum→GPe→STN→GPi), (B) Trajet indirect « court » (cortex→striatum→GPe→GPi), (C) Projection inhibitrice du striatum vers lui-même symbolisant les inhibitions latérales présentes dans cette région (d'après [Bar-Gad et al., 2003, Hamker, 2013, Smith et al., 1998]).

La voie hyperdirecte Cette voie va directement relier le cortex au STN et le STN au GPi par des projections excitatrices (glutamatergiques). Elle va exercer une rétroaction globalement inhibitrice sur la boucle cortico-corticale et cet effet sera généralisé en raison de la forte divergence des projections du STN vers le GPi. La voie hyperdirecte va donc exercer une inhibition corticale rapide et généralisée afin de maintenir et contrôler toutes les actions possibles [Dunovan and Verstynen, 2016].

L'interaction entre les voies directe et hyperdirecte a tendance à amplifier la moindre différence entre les réseaux corticaux activant différentes voies directes [Mink, 1996]. En effet, ces deux voies opèrent comme des boucles de rétroaction en compétition. Ce mécanisme serait une des bases du fonctionnement des BG, notamment pour ce qui concerne la sélection de l'action [Leblois et al., 2006a].

La voie indirecte Cette autre voie majeure des BG relie le cortex au GPi mais avec des connexions intermédiaires : le striatum se projette sur le GPe qui se projette à son tour sur le STN, le tout par des projections inhibitrices gabaergiques. Les projections du STN vers le GPi

quant à elles sont excitatrices (glutamatergiques). Les MSN striato-pallidiaux de la voie indirecte expriment les récepteurs D2 à la DA [Lambot et al., 2016].

Il était initialement envisagé que la voie indirecte entraînait une boucle fermée de rétroaction négative, mais avec une délimitation plus précise des connexions anatomiques [Bevan et al., 1997, Kita and Kita, 1994, Smith and Bolam, 1989, Guthrie et al., 2013], il semblerait que cette voie forme en réalité davantage un sous-réseau, ce dernier comprenant des boucles de rétroaction positive et négative au cœur même des BG. Smith a suggéré que le modèle canonique [Albin et al., 1989, DeLong, 1990] de la voie indirecte devait être adapté pour répondre aux observations et à la synaptologie [Smith et al., 1998]. Il semble en effet que cette voie indirecte puisse suivre deux trajets : un trajet long qui est celui qui est initialement retenu (Figure 1.3) qui co-existerait avec un trajet indirect court [Hamker, 2013]. En raison de cette augmentation de la complexité anatomique, il n'est pas certain que la voie indirecte forme une boucle cortico-sous-corticale assurant le transfert précis de l'information spatiotemporelle [Levy et al., 1997]. Les divergences entre les sous-boucles GPe-striatum et GPe-STN [Bevan et al., 1997, Chang et al., 1981, Gerfen, 1985], tout comme celles dans les collatérales des axones du GPe [Bevan et al., 1997, Brown et al., 1998] suggéreraient même que c'est peu probable.

De récentes études [Schroll et al., 2014] ont proposé un rôle spécifique de la voie indirecte quant à son implication dans les processus d'apprentissage. Il a en effet été montré que, lors d'une prise de décision, les MSN de la voie indirecte inhiberaient les alternatives principales à cette décision ayant été précédemment considérées comme incorrectes. Ce même phénomène permettrait également à la voie indirecte d'inhiber les mouvements superflus. En effet il a été montré que des lésions à l'intérieur de cette voie empêchaient l'inhibition des réponses non désirées en conduisant à des mouvements choréïques du type de ceux observés dans les premiers stades de la chorée de Huntington [Schroll et al., 2015].

1.1.3 Le cortex

1.1.3.1 Anatomie

Le cortex est une structure cérébrale située dans le télencéphale commune à de nombreuses espèces animales (Figure 1.4), ce qui peut s'expliquer par le fait qu'il est le siège de la plupart des fonctions cognitives supérieures et de comportements complexes. La taille du cortex varie selon les espèces, mais de manière générale il s'agit de la plus grande structure cérébrale. Il représente environ 80% de la masse cérébrale totale chez l'humain. Le cerveau humain entier compte environ 120 milliards de neurones, le cortex va en contenir entre 21 et 26 milliards (et dix fois plus de cellules gliales) [Herculano-Houzel, 2009].

Il a une structure multi-laminaire, c'est à dire qu'il est constitué de plusieurs couches [Ramon y Cajal, 1894]. Ces couches possèdent des types et une densité de neurones différents, ainsi qu'une organisation propre à chacune. Le néocortex par exemple est divisé en six couches, numérotées en chiffres romains de la plus superficielle à la plus profonde. La couche I dite moléculaire, possède peu de neurones. Les couches II et III sont les couches supra-granulaires, les VI et VI sont les couches infra-granulaires. La couche IV correspond à la couche granulaire et reçoit les connexions thalamo-corticales.

Les couches du cortex sont lisses chez les rongeurs, alors que chez les mammifères elles

sont repliées mais plus étendues (ce qui permet d'optimiser l'occupation du volume crânien) [Hill and Walsh, 2005]. Ces replis forment des sillons plus ou moins profonds, délimitant les circonvolutions cérébrales (ou gyrus). Les sillons les plus profonds vont diviser le cortex en quatre principaux lobes (régions) : occipital, pariétal, temporal et frontal [Felleman and Essen, 1991].

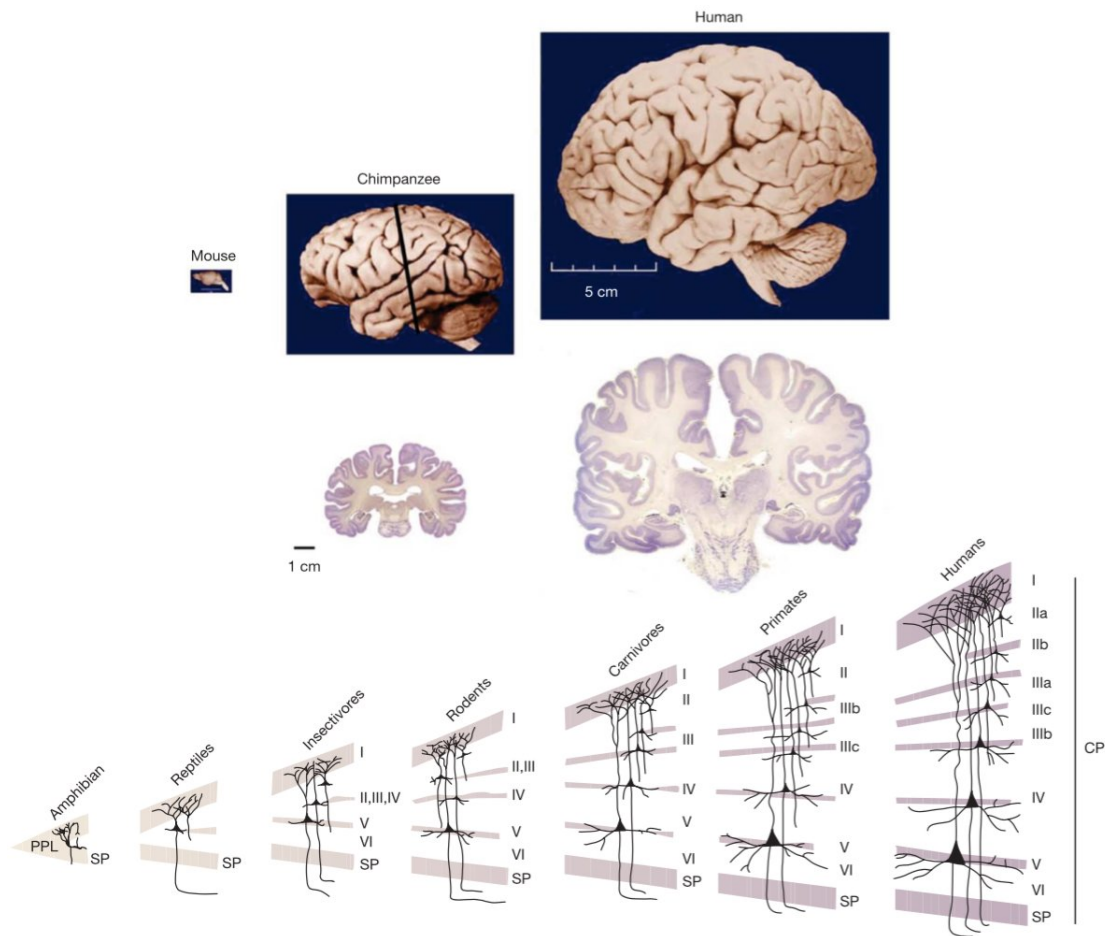


FIGURE 1.4 – Taille et organisation du cortex chez différentes espèces, d'après [Hill and Walsh, 2005]. La partie du haut de la figure, de gauche à droite, montre une vue de côté d'un cerveau de souris, de chimpanzé et d'humain pour illustrer les tailles relatives. La partie du milieu montre une coupe de cerveau de chimpanzé et d'humain. La partie du bas montre la composition cellulaire du cortex chez différentes espèces (de gauche à droite : amphibiens, reptiles, insectivores, rongeurs, carnivores, primates, humains).

Le néocortex possède une organisation fonctionnelle bien spécifique et va donc être divisé en régions fonctionnelles : les aires corticales. Ces dernières vont être communes à tous les individus au sein d'une même espèce, mais n'y auront pas systématiquement la même taille [Rakic, 1988, Alexander et al., 1986]. Les aires corticales se divisent généralement en trois catégories : les aires sensorielles, motrices et d'association. Elles possèdent une organisation physiologique hiérarchique, décrite pour la première fois par [Felleman and Essen, 1991]. Dans le sens rostro-caudal se trouvent successivement les aires préfrontale, pré-motrice, motrice primaire, somato-sensorielle et associatives. Plus le niveau de hiérarchie est élevé et plus les aires sont intégratives, c'est à dire qu'elles vont coordonner l'activité de plusieurs fonctions.

Le cortex possède également une spécialisation locale, indiquant que les aires corticales peuvent être représentées plus essentiellement sous la forme d'une carte des comportements et

capacités essentielles propres à l'espèce en question. Chez les rongeurs, par exemple, les aires impliquées dans l'odorat sont comparativement plus importantes que celles impliquées dans la vision, ce qui est cohérent avec les importances respectives de ces deux domaines sensoriels. En revanche, chez les primates pour lesquels la capacité visuelle joue un rôle majeur, plus d'une dizaine de domaines fonctionnels distincts impliqués dans l'analyse d'une scène visuelle ont été identifiés [Hill and Walsh, 2005].

1.1.3.2 Connectivité

Le cortex représente la structure d'entrée dans l'ensemble des boucles cortico-baso-thalamiques. En effet, le système de projection cortico-striatal représente la première étape de traitement de l'information transitant par les BG [Alexander et al., 1986, Parent and Hazrati, 1995]. Le cortex va envoyer des projections glutamatergiques excitatrices vers le striatum (voies directe et indirecte) et le STN (voie hyperdirecte). Il est également interconnecté par le même type de projections au thalamus [Alexander et al., 1986, Bar-Gad et al., 2003]. Toutes ces projections sont censées imposer une organisation fonctionnelle au niveau cortical qui se répercute au sein des boucles cortico-baso-thalamo-corticales. La connaissance de l'organisation des projections cortico-striatales est donc importante pour déterminer la nature ainsi que le chemin de l'information transitant par le circuit des BG.

1.1.3.3 Évolution structurelle

Chez l'homme, le cortex semble avoir une grande importance dans le processus de prise de décision. Cependant, il n'apparaît que chez les mammifères car les autres vertébrés possèdent un pallium. Ce dernier a une anatomie plus simple que le cortex : il est constitué d'une couche de cellules entourées de part et d'autre par deux couches de fibres de communication. De plus, il représente moins de 20% de la masse cérébrale chez les espèces les plus archaïques [Boraud, 2015a].

Ainsi, des anamniotes² (où il n'est qu'une afférence secondaire des BG) aux mammifères, le télencéphale contenant le cortex va progressivement devenir l'interlocuteur principal des BG. Cependant, le cortex des mammifères n'est pas une simple expansion du pallium des vertébrés inférieurs (Figure 1.5). Il correspond aussi à une modification structurelle significative. Le pallium est constitué de deux ou trois types de neurones organisés en trois couches. Le cortex quant à lui se subdivise en six couches de neurones et comprend une dizaine de types de neurones différents [Shepherd, 2011]. Ils s'organisent de façon modulaire en mini-unités fonctionnelles constituées d'un neurone de projection et d'interneurones. Cette unité fonctionnelle, appelée colonne, fournit des modules de base beaucoup plus complexes et performants que le système archaïque à trois couches. La richesse et le développement du registre comportemental des mammifères doivent probablement autant à cette complexification qu'à la simple augmentation de volume de l'encéphale.

Au cours de l'évolution chez les vertébrés, le cortex olfactif est la région du cortex qui a le moins changé avec l'hippocampe. En effet, chez les mammifères, le cortex à trois couches persiste dans le cortex olfactif et l'hippocampe. Cela conforte l'hypothèse qu'un cortex à trois

2. Essentiellement les poissons et les amphibiens.

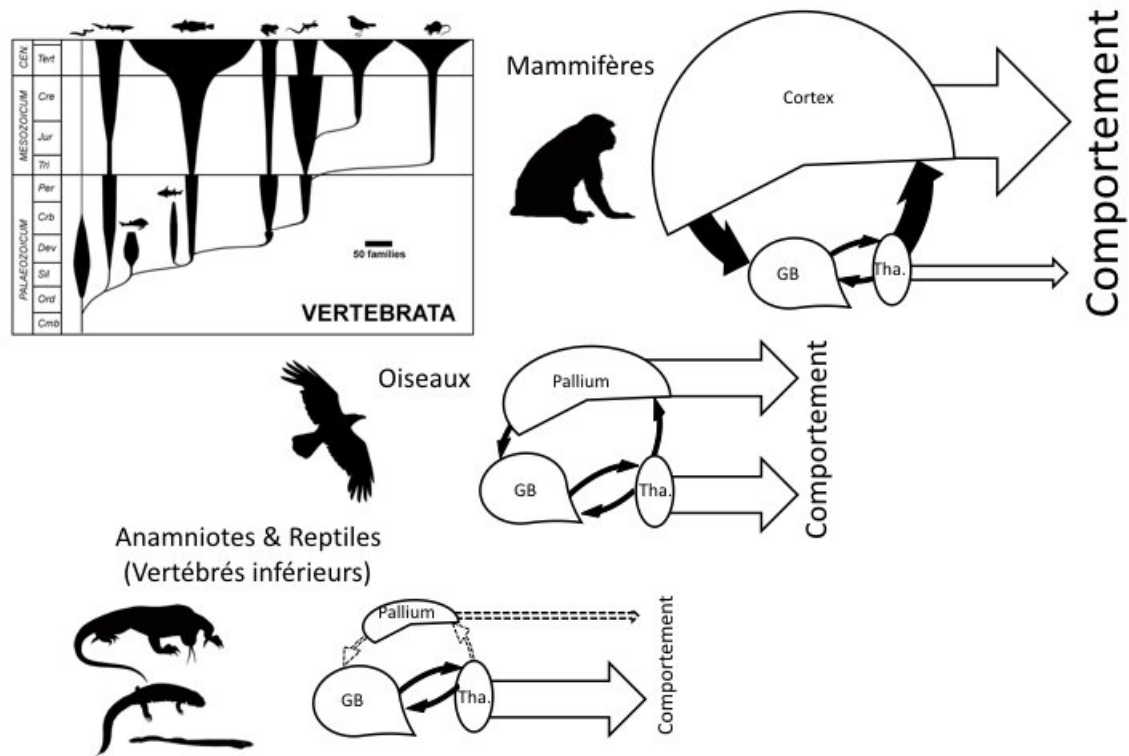


FIGURE 1.5 – Les interactions entre les ganglions la base et le télencéphale sont proportionnelles au rôle de ce dernier dans le contrôle des comportements, [Boraud, 2015b]. Chez les anamniotes, le comportement est principalement contrôlé par le thalamus qui est l’interlocuteur principal des BG. Chez les oiseaux, une transition est observée : le pallium prend en charge une partie de plus en plus importante du comportement et les BG interagissent dans des proportions identiques avec les deux autres structures. Chez les mammifères le cortex prend le contrôle de la plus grande partie des comportements en corrélation avec l’augmentation de sa taille. Il devient aussi la source principale d’afférences des BG et une de leur cible privilégiée.

couches contiendrait le cœur d’un micro-circuit basique, qui aurait été élaboré dans le néocortex à six couches.

De plus, des études sur les connexions efférentes et afférentes du striatum et du NAc des urodèles ont montré que ce groupe de vertébrés amphibiens partage de nombreuses caractéristiques sur l’organisation des BG avec les amniotes³ [Marín et al., 1997a, Marín et al., 1997b]. L’étude des BG chez les amphibiens constitue donc une piste sérieuse et cohérente pour aider à la clarification du fonctionnement sous-cortical des BG chez les vertébrés.

Ces similitudes concernent principalement les caractéristiques générales telles que les entrées et sorties des structures, ainsi que les neurotransmetteurs. En effet, chez les mammifères, le cortex est fortement interconnecté avec le NAc et le striatum. Il envoie des projections directes vers ces structures et en obtient une forte rétroaction via le thalamus (Figure 1.5).

3. Mammifères, reptiles, oiseaux.

1.1.4 Le thalamus

1.1.4.1 Anatomie

Le thalamus est une structure cérébrale symétrique et se présentant sous la forme de noyaux composés de deux moitiés partiellement distinctes. Il est situé dans le diencephale, entre le cortex et le tronc cérébral et est à l'interface entre les systèmes sensoriels et les systèmes moteurs. Le thalamus est une composante essentielle des circuits entre les BG et le cortex et intervient notamment dans la dynamique émotionnelle et motivationnelle, la planification et la cognition au cours du développement et l'expression de comportements dirigés vers un but [Haber and Calzavara, 2009]. Chaque région fonctionnelle du cortex frontal est reliée à des zones spécifiques de chaque structure des BG et du thalamus.

1.1.4.2 Connectivité

Le thalamus a longtemps été perçu comme la structure cérébrale transmettant l'information spinale (véhiculant des informations sensorielles comme la nociception, la température et le toucher) et sous-corticale vers le cortex, formant le dernier maillon dans le circuit entre le cortex et les BG. En effet, dans les premiers modèles fonctionnels des BG, le rôle principal du thalamus était le relai de l'information traitée dans les BG provenant du GPi/SNr et allant vers le cortex [Albin et al., 1989, DeLong, 1990, Bar-Gad et al., 2003] et le STN [Parent and Hazrati, 1995]. Les groupes nucléaires du thalamus associés principalement à cette fonction sont les noyaux ventraux antérieur et latéral, ainsi que le noyau médio-dorsal. D'après la revue de [Bar-Gad et al., 2003], il existe des connexions réciproques complexes entre le thalamus et le cortex. En effet, en comparaison des projections thalamo-corticales directes (« feed-forward »), les rétro-projections cortico-thalamiques vers les noyaux ventraux antérieur, latéral et médio-dorsal sont plus importantes. Cela montre la réciprocité partielle des connexions dans le relai de l'information thalamo-cortico-thalamique.

Le thalamus reçoit des entrées glutamatergiques excitatrices massives de la part du cortex et de la moelle épinière [Parent and Hazrati, 1995]. Le GPi et la SNr lui envoient également des projections unidirectionnelles gabaergiques inhibitrices. Cela va permettre de créer une boucle de connectivité partiellement fermée entre le thalamus, le cortex et les BG comme cela est illustré dans la (Figure 1.3).

Il possède une somatotopie marquée (comme chaque noyau des BG [Nambu, 2011]) notamment dans sa partie orale où le noyau ventrolatéral reçoit les projections pallidales, dans sa région médiale où il reçoit les projections de la SNr ainsi que dans la région caudale de son noyau ventroposterolatéral qui reçoit des projections cérébelleuses. En plus de constituer une région de convergence de ces deux structures jouant un rôle majeur dans les processus sensori-moteurs en général, ces deux noyaux ont chacun leur propre représentation somatotopique, qui se prolonge dans le sens rostrocaudal [Nambu, 2011].

1.1.5 Les principales pathologies des BG

Comme nous l'avons vu précédemment, les BG sont un ensemble de structures cérébrales au cœur du CNS. Une atteinte de l'une ou plusieurs de ces structures peut donc avoir d'importantes

répercussions sur la physiologie et l'état de l'individu. Certains troubles des BG ne se manifestent que par le biais de difficultés motrices, alors que d'autres combinent les difficultés motrices et cognitives, voire des troubles comportementaux [Chiu et al., 2016]. D'une façon générale, il est intéressant d'associer un modèle physiologique comme celui qui est présenté ici à des troubles ou à des effets lésionnels. Cette approche complémentaire permet en effet (i) de renforcer la cohérence globale du modèle et également (ii) d'en obtenir des prédictions, voire dans certains cas de tester *in silico* des hypothèses sur des pistes thérapeutiques. La partie qui suit se propose donc de présenter brièvement les principales pathologies en relation avec les circuits des BG vus précédemment.

Maladie de Parkinson Cette maladie a été décrite pour la première fois au début du XIX^e siècle par James PARKINSON. Elle se caractérise principalement par des symptômes moteurs comme la rigidité, les tremblements, l'akinésie (lenteur d'initiation des mouvements, avec tendance à l'immobilité) ou la bradykinésie (lenteur d'exécution, ralentissement des mouvements) [Bar-Gad et al., 2003, Utter and Basso, 2008].

La cause exacte de la maladie de Parkinson n'a pas encore été identifiée à ce jour et plusieurs facteurs, aussi bien d'ordres environnementaux [Hu and Wang, 2016] que génétiques, semblent impliqués dans son étiologie [Coetzee et al., 2016]. Cependant, nous savons que les symptômes les plus évocateurs - à savoir essentiellement les troubles moteurs - sont dus à une dégénérescence des neurones dopaminergiques dans la SNc [Albin et al., 1989]. La DA a un effet ambivalent selon ses sites d'action. Dans la voie directe, les récepteurs D1 ont un effet globalement excitateur sur l'effet des projections du striatum vers le GPi [Planert et al., 2013] alors que dans la voie indirecte, les récepteurs D2 ont un effet inhibiteur sur le GPe [Galvan et al., 2015, Calabresi et al., 2014]. Le manque de DA va entraîner un déséquilibre entre les voies directe et indirecte en faveur de cette dernière. L'activation du GPe va donc être diminuée, donc son rôle inhibiteur sur le STN va être atténué et celui-ci va donc exercer une activité amplifiée sur le GPi. Dans le même temps, la baisse d'activité des projections striatales inhibitrices sur le GPi va être diminuée (effet de la perte de DA sur la voie directe) et cet effet va s'ajouter à l'augmentation d'activité liée au STN. La conséquence est que le GPi va inhiber fortement le thalamus, ce qui va se traduire par une inhibition motrice, une perte de la facilitation, une baisse de l'initiation des mouvements et plus globalement un déficit moteur constaté en clinique (Figure 1.6A).

Le premier traitement de la maladie de Parkinson a été développé en s'intéressant à la molécule précurseur de la DA, à savoir la L-Dopa (ou lévodopa) [Wichmann and DeLong, 1996]. Cette molécule va pouvoir, après administration, franchir la barrière hémato-encéphalique et permettre de restaurer le taux de DA. Cependant, après un traitement prolongé à la L-Dopa, beaucoup de patients présentent des effets secondaires très invalidants du type dyskinésie dopa-induite. Ces troubles secondaires peuvent à leur tour être pris en charge par des techniques de Stimulation Cérébrale Profonde (« *Deep Brain Stimulation* » en anglais, DBS) [Umemura et al., 2016]. C'est à la fin du XX^e siècle que cette technique a été effectuée avec succès sur des patients parkinsoniens [Benabid, 2003]. Cette technique, qui s'appuie souvent sur l'insertion d'électrodes de stimulation dans le STN, permet d'améliorer l'état du patient en réduisant les phases akinétiques (quand la L-Dopa n'a temporairement plus d'effet), les phases « off », les états dyskinétiques et va permettre de normaliser l'activité des BG.

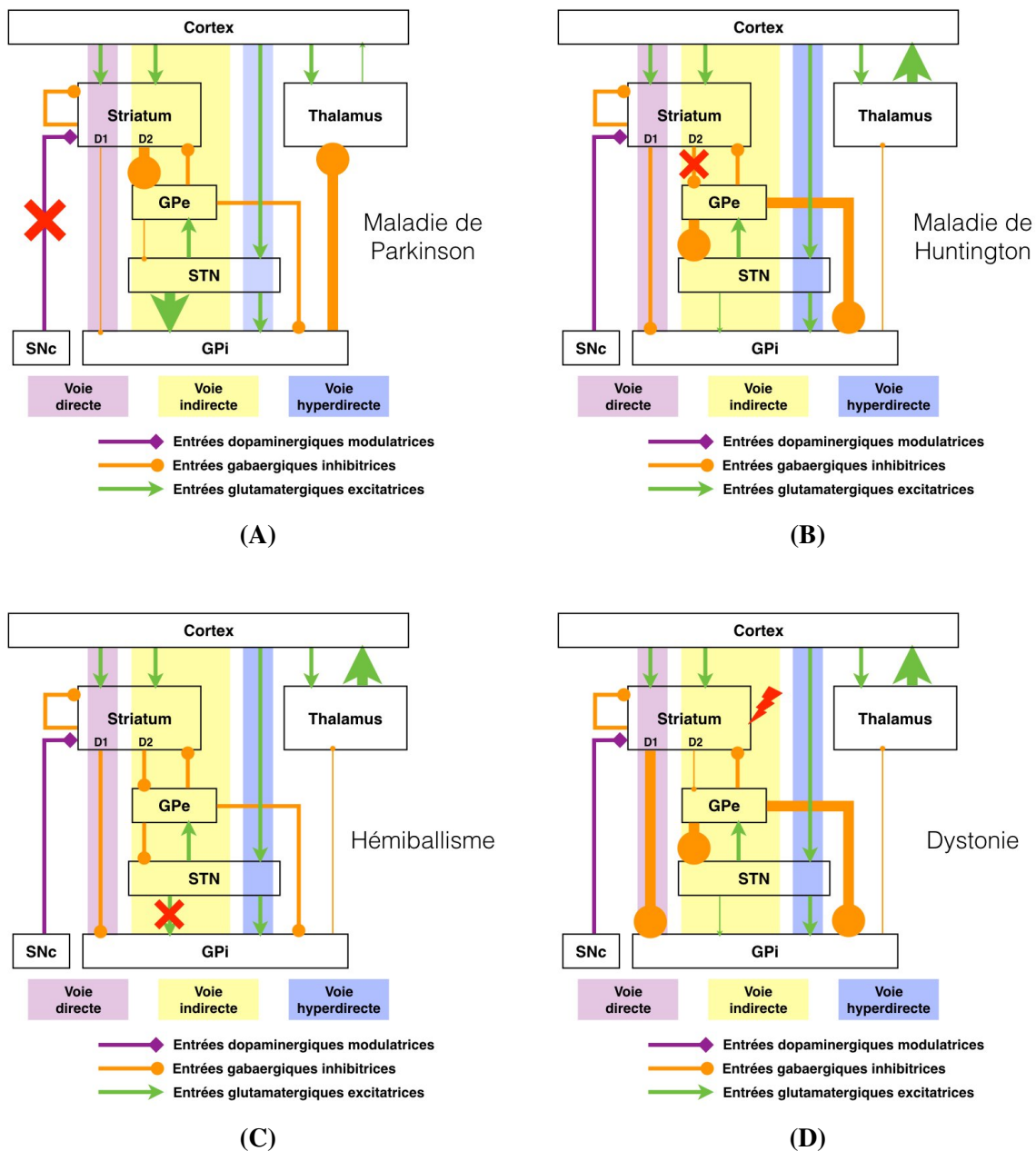


FIGURE 1.6 – Relations entre les principaux syndromes neuropathologiques observés dans les BG et les altérations de leur connectivité. La croix rouge montre la zone touchée selon la pathologie et l'épaisseur des flèches symbolise la force de connexion entre deux zones. (A) Cas de la maladie de Parkinson où une dégénérescence des projections de la SNc vers le striatum provoque une augmentation de l'inhibition de ce dernier vers le GPe, entraînant une diminution de l'inhibition du GPe vers le STN, puis une augmentation de l'excitation du STN vers le GPi, une augmentation de l'inhibition du GPi vers le thalamus et enfin une diminution de l'excitation du cortex par le thalamus. (B) Cas de la maladie de Huntington où une dégénérescence des projections du STN vers le GPe (voie indirecte) va provoquer une diminution de l'inhibition tonique du GPi vers le thalamus, ce qui aura pour effet une augmentation de l'excitation du cortex par le thalamus. (C) Cas de l'hémiballisme où une diminution de l'excitation du GPi par le STN va provoquer une diminution de l'inhibition du thalamus par le GPi en entrainer une activation plus forte du cortex par le thalamus. (D) Cas de la dystonie où les projections du striatum vers le GPe (voie indirecte) et le GPi (voie directe) vont être affectées. L'inhibition du GPi sera augmentée alors que celle du GPe sera diminuée, provoquant une augmentation de l'inhibition du STN par le GPe, puis diminution de l'excitation du GPi par le STN. Cela va entrainer une diminution de l'inhibition du thalamus par le GPi et donc une augmentation de l'excitation du cortex par ce dernier.

Maladie de Huntington La chorée de Huntington se caractérise par des mouvements choréiformes (mouvements spasmodiques involontaires des extrémités) sans but apparent. Le plus souvent, les signes précurseurs de cette maladie consistent en une subtile altération de l'état émotionnel et de l'humeur par des modifications du caractère [Armstrong et al., 2007, Albin et al., 1989].

La maladie de Huntington se manifeste habituellement à la moitié de l'âge adulte entre 40 et 60 ans et est due à une mutation génétique autosomique dominante [Gusella et al., 1983]. Par un mécanisme encore méconnu, cette mutation va provoquer une dégénérescence des projections gabaergiques [Utter and Basso, 2008] des MSN projetant sur le GPe (voie indirecte) (Figure 1.6B). Cette altération des projections striatales va entraîner une activation anormale du GPe et donc une diminution dans l'excitation qu'envoie le STN au GPi. Cela va provoquer une diminution des efférences inhibitrices des BG et donc les troubles moteurs et cognitifs observés.

À l'heure actuelle, il n'existe pas de remède ni de traitement pour la chorée de Huntington. Cependant, il est possible de traiter les symptômes en utilisant des antagonistes de la DA afin de régulariser l'activité des projections striato-pallidales ou la stimulation cérébrale profonde.

Hémiballisme Il s'agit d'un syndrome rare provoquant un trouble moteur associant des mouvements unilatéraux des membres supérieurs de façon involontaire, rapide et désordonnée. Ces mouvements sont en général de grande amplitude et assez violents [Postuma and Lang, 2003]. L'hémiballisme est une sous-catégorie d'un syndrome plus étendu : le ballisme [Albin et al., 1989]. Ce dernier, plus rare, a les mêmes symptômes mais sur l'ensemble du corps.

Le mécanisme de cette maladie est assez similaire à celui de la chorée de Huntington : une diminution de l'activité du STN va entraîner une diminution de l'activité du GPi et provoquer par la suite une désinhibition des projections thalamo-corticales [Wichmann and DeLong, 1996]. Cela va donc empêcher l'inhibition tonique des BG (Figure 1.6C) et déclencher une décharge anormale des neurones moteurs.

Dystonie Il s'agit d'une maladie musculaire, parfois héréditaire, se traduisant par des positions spontanées extrêmes et contorsionnées d'une ou plusieurs parties du corps, pouvant durer de quelques secondes à quelques minutes [Albin et al., 1989]. Ces contractions résultent de la co-contraction soutenue du muscle squelettique, au repos ou lors d'un mouvement précis. Le cas le plus connu de dystonie localisée se produit dans les mains d'écrivains ou musiciens professionnels (« crampe de l'écrivain » par exemple).

La dystonie va affecter les efférences du striatum et donc impacter les voies directe et indirecte des BG (Figure 1.6D). Dans ce cas, le système moteur va continuer à fonctionner mais sans la modulation imposée par les BG [Albin et al., 1989]. Une mutation génétique peut être la cause de certaines formes de dystonies, mais la façon dont se produit cette mutation n'est pas encore claire de nos jours [Utter and Basso, 2008].

Syndrome de Gilles de la Tourette Cette maladie se caractérise principalement par des tics convulsifs, pouvant être accompagnés de coprolalie⁴ dans certains cas.

4. L'individu profère à haute voix des mots qu'il voudrait bien retenir [de la Tourette, 1899]

Comme Albin le fait remarquer dans son article sur l'anatomie fonctionnelle des troubles des BG [Albin et al., 1989], le syndrome de Gilles de la Tourette offre un contraste intéressant entre la chorée de Huntington (mouvements hyperkinétiques) et la maladie de Parkinson (dégénérescence des neurones dopaminergiques dans la SNc). En effet, une hyperactivité du système dopaminergique (ou une hypersensibilité des récepteurs) désinhiberait le striatum par rapport aux conditions normales ce qui aboutirait à une trop forte excitation thalamo-corticale, provoquant une sur-stimulation des centres moteurs dans le cortex et des mouvements hyperkinétiques [Cohen et al., 1997].

Les patients adultes atteints de ce syndrome peuvent être traités avec des antagonistes de la DA mais leur efficacité reste limitée. Des essais de DBS dans la thalamus [Mink, 2004, Utter and Basso, 2008] ont été réalisés avec des résultats encourageants.

Troubles Obsessionnels Compulsifs (TOC) Les TOC sont caractérisés par des pensées ou images intrusives formant des obsessions, qui vont augmenter l'anxiété. Dans la plupart des cas, le fait de réaliser une action rituelle ou répétitive va permettre de ramener le niveau d'anxiété à la normale [Stein, 2002].

Les patients souffrant de troubles obsessionnels compulsifs peuvent également avoir des altérations au niveau neuropsychiatrique (problèmes d'identification olfactive ou inhibition intracorticale par exemple) ou neuropsychologique (problèmes avec les fonctions exécutives ou de mémorisation visuelle). Ces bouleversements sont cohérents avec un dysfonctionnement et une diminution de l'inhibition au niveau des boucles cortico-striato-thalamico-corticales.

Une pharmacothérapie est possible dans certains cas, avec l'introduction d'inhibiteurs de recaptage de la sérotonine dans l'organisme afin de réguler le niveau de l'inhibition dans les BG. Dans le cas où ce traitement ne fonctionnerait pas, il est possible d'utiliser des techniques de stimulation cérébrale profonde ou, en dernier recours, la neurochirurgie [Utter and Basso, 2008]. De récents travaux en optogénétique ouvrent également la voie pour un nouveau type de traitement des comportements compulsifs [Burguière et al., 2013].

Toutes ces maladies et troubles cités précédemment impliquent une altération plus ou moins importante des BG et malgré les récentes découvertes quant à leur fonctionnement, il n'existe pas de traitement ou de remède pour chacune d'entre elles. À terme, leur prise en charge passera certainement par un meilleur contrôle des pharmacothérapies, des modalités de DBS, ou encore dans certains cas, d'interventions chirurgicales.

De nombreux modèles - animaux ou théoriques - ont été et sont encore actuellement développés afin d'apporter une compréhension plus fine de la dynamique interne des BG et de leurs structures associées, ainsi que de leurs interactions. En effet, l'ambition d'un modèle est généralement d'aider à mieux comprendre les mécanismes à l'origine des troubles conduisant à l'expression des symptômes observés lors de la maladie en question [Féger et al., 2009]. L'amélioration de ces modèles demeure toujours une nécessité car ils permettent des investigations ciblées (*in vivo*, *in vitro*, *in silico* et même *post-mortem*) afin d'apporter des éléments nouveaux aux hypothèses de travail orientées vers un développement thérapeutique.

1.2 Approche expérimentale de l'étude des mécanismes de la prise de décision

1.2.1 Principes généraux des protocoles utilisés

Cette partie va succinctement présenter les méthodes, principes et métriques fréquemment utilisés pour étudier sur des modèles animaux les principes de la prise de décision et de l'apprentissage par renforcement. Les métriques utilisées dans les approches expérimentales qui nous inspirent sont de deux ordres : (i) comportemental et (ii) électrophysiologique. Comme nous l'avons mentionné en préambule, notre objectif (à l'instar de celui de nombreux autres modèles computationnels) est de lier l'activité des différentes régions des BG (exprimée dans notre cas à l'aide de trains de spikes) avec les comportements observables des animaux en action. Cette approche constitue une étude en boucle fermée selon un schéma classique en neurosciences computationnelles et est présentée dans la Figure 1.7.

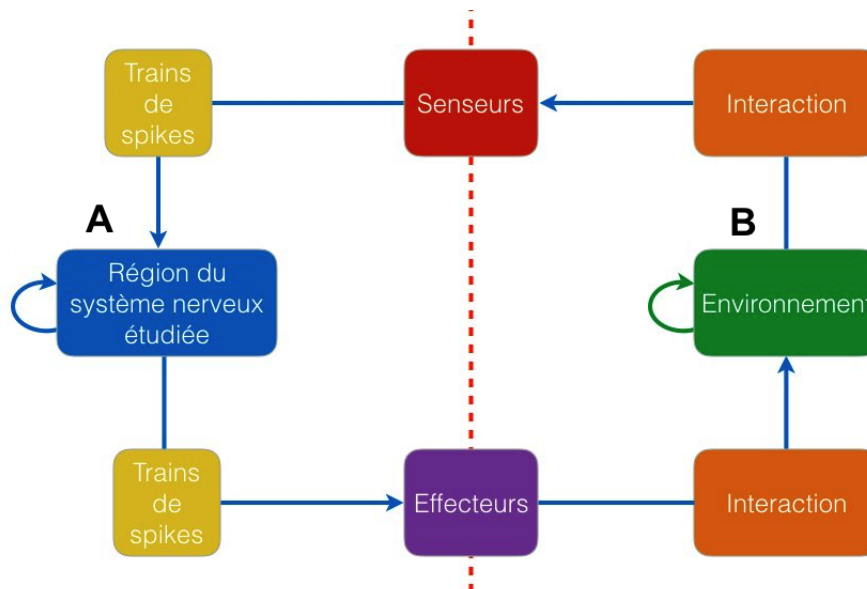


FIGURE 1.7 – Approche en boucle fermée Le modèle computationnel est symbolisé par l'encadré bleu qui représente la région du CNS qui est étudiée. Cette région est susceptible d'être modifiée (A) lors de processus d'apprentissage et est séparée de son environnement par une interface sensori-motrice représentée ici par la ligne rouge en pointillés. Le modèle émet des trains de spikes (spontanément, en réponse à des stimulations ou les deux) et ces trains de spikes seront encodés de façon à entraîner des effecteurs (représentation des jonctions neuro-musculaires et activité motrice résultante). Les effecteurs matérialisent l'interaction (essentiellement mécanique) avec l'environnement dans lequel évolue l'organisme. L'environnement est éventuellement dans ce cas modifiable (B). L'environnement en retour interagit sur le sensorium du modèle (de façon active ou passive) et cette action est traduite en trains de spikes (ou en une autre modalité de stimulation) ce qui ferme la boucle

Cette représentation simple met le modèle en situation et permet d'étudier ses propriétés émergentes et donc sa cohérence globale et par voie de conséquence, sa validité. Elle est particulièrement indiquée lorsque l'on étudie les phénomènes d'apprentissage dans la mesure où c'est le modèle seul sans superviseur qui doit accomplir *in silico* ce dont est capable son homologue biologique. Un autre avantage de ce formalisme est que, pour généraliser les capacités du modèle, il suffit de modifier la partie « environnement » et de voir ainsi s'il se comporte

encore de façon cohérente avec les résultats expérimentaux. Ce cadre de travail est suffisamment souple pour s'appliquer à beaucoup de protocoles utilisant les modèles animaux aussi bien que les modèles computationnels. En effet, la partie concernant les enregistrements électrophysiologiques correspond aux trains de spikes entrant et sortant du modèle et la partie comportement correspond essentiellement aux interactions mécaniques émanant des effecteurs (Figure 1.7). Les modes d'encodage des signaux comportent toutefois de nombreux blancs et certaines hypothèses doivent être posées en différents points de la boucle. Ces hypothèses seront abordées et justifiées dans les paragraphes suivants.

1.2.2 Supports expérimentaux de notre étude

Afin de d'implémenter un modèle computationnel, il est nécessaire de disposer au préalable des résultats expérimentaux adéquats (bien que partiels) afin s'assurer de sa cohérence. Dans cette partie nous allons détailler les moyens d'investigation mis en œuvre au niveau expérimental nous ayant aidé dans la construction du modèle.

1.2.2.1 Travaux de Pasquereau et al. (2007)

Cette tâche est issue d'une étude sur les BG visant à mettre en place chez l'animal des réponses motrices en fonction de différents stimuli associés à des récompenses [Pasquereau et al., 2007]. Lors de cette étude, les neurones du striatum dorsal et du GPi (respectivement entrée et sortie des BG) ont été enregistrés simultanément chez deux singes en train de réaliser une tâche motrice dans laquelle des cibles visuelles avaient été associées à différentes probabilités de récompense.

Thème général de l'étude Une tâche motrice visuellement orientée va vraisemblablement impliquer les BG comme « système d'assistance ». De précédentes études sur le striatum suggèrent une modulation de la récompense liée à l'activité des neurones dopaminergiques du mésencéphale. Le striatum serait donc en mesure de déterminer une probabilité ainsi qu'une quantité de récompense potentielle, associée à un stimulus spécifique. De plus, cette information pourrait à son tour être utilisée dans le choix d'une réponse comportementale appropriée. L'encodage d'une valeur d'incitation pour un stimulus spécifique par les neurones striataux et la modulation des neurones dopaminergiques étant bien documenté, les auteurs se sont focalisés sur les interactions entre les informations cognitives (valeur de récompense) entrantes et un comportement moteur résultant du processus de prise de décision.

Animaux Deux singes femelles (*Macaca mulatta*) de 5.6 et 4 kg ont participé à ces travaux, en périodes de jour uniquement. Les primates ont été maintenus en restriction hydrique (mais avec nourriture à volonté) afin d'accroître leur motivation durant les sessions d'apprentissage et d'enregistrement. Les animaux ont été entraînés quotidiennement afin d'apprendre la tâche comportementale. Au bout de neuf et six mois d'entraînement respectivement, les primates avaient atteint un taux de succès supérieur à 95% pour des séries de plus de 400 essais. Dès lors, une chambre d'enregistrement leur a été implantée pour la suite des tests. Tous ces travaux ont été réalisés en respectant toutes les règles de l'éthique de l'expérimentation animale.

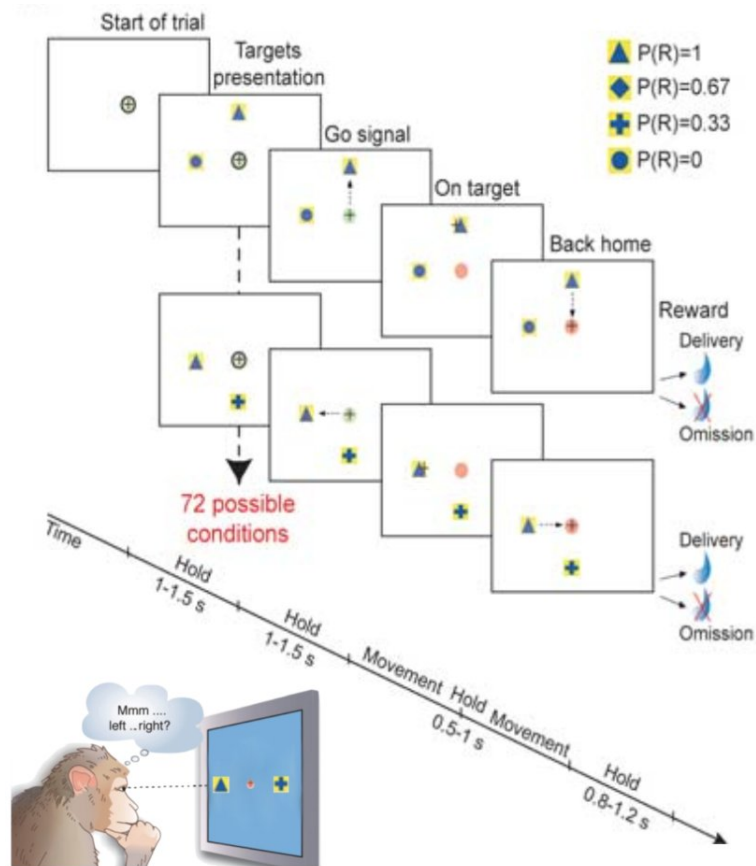


FIGURE 1.8 – Schéma de la tâche comportementale à quatre cibles d’après [Pasquereau et al., 2007]. Cette tâche de prise de décision en condition d’incertitude est fondée sur une probabilité de récompense. À chaque essai, deux cibles différentes ($P1$ et $P2$) choisies parmi quatre sont présentées simultanément à l’écran. Chaque cible peut apparaître à quatre positions possibles et ces dernières sont sélectionnées aléatoirement à chaque essai. Il y a donc six combinaisons de cibles (deux choisies parmi les quatre) et six possibilités de positionnement (deux cibles parmi quatre emplacements) ce qui représente un total de 72 combinaisons possibles. L’animal doit ensuite suivre une séquence précise pour réaliser la tâche convenablement et obtenir une récompense.

Tâche comportementale Les singes ont été entraînés à actionner une manette dans un plan horizontal (26 x 26 cm) de la main droite. Cette manette permettait de déplacer un curseur sur l’écran d’un ordinateur placé à 50 cm en face du visage des animaux. Ces derniers devaient initier un essai en gardant le curseur à l’intérieur d’un cercle plein vert (ce cercle étant lui-même entouré d’un autre cercle vide noir), au centre de l’écran (Figure 1.8). Après un laps de temps aléatoire (de 1 à 1.5 secondes), deux différentes cibles parmi les quatre possibles étaient choisies aléatoirement et affichées simultanément à l’écran (Table 1.1). Chaque cible pouvait apparaître dans l’une des quatre directions possibles (0° , 90° , 180° et 270°) et une probabilité de récompense différente (0, 0.33, 0.67 et 1) était attribuée à chacune. La cible ainsi que la direction associée étaient choisies aléatoirement à chaque essai afin d’éviter tout biais dans l’expérience. Dans le but de toujours avoir une situation où il y aurait un « meilleur » choix, il ne pouvait pas y avoir deux fois le même cible affichée à l’écran, ou deux cibles à la même position dans un même essai. Après un autre laps de temps aléatoire (1-1.5 s également), le signal de départ était donné par la disparition du cercle noir central et les singes devaient initier le mouvement vers l’un des deux cibles (Figure 1.9). Le curseur devait être maintenu sur la cible pendant une période aléatoire

(0.5-1 s), avant d'être ramené au niveau du cercle central. Pour compléter et finir l'essai, les singes devaient maintenir le curseur à l'intérieur de ce cercle central pendant une durée minimum aléatoire (0.8-1.2 s). La disparition du cercle central indiquait à l'animal qu'il avait réussi l'essai. La récompense (0.3 ml du jus de fruit) était donnée en fonction de la probabilité associée à la cible sélectionnée. À chaque essai, si un singe choisissait la cible associée à la plus haute probabilité de récompense, son choix était considéré comme optimal. Dans tous les cas, une récompense était fournie ou non à l'aide d'un tirage aléatoire et en suivant la probabilité associée à la cible choisie. Les essais étaient espacés d'intervalles inter-essais (« *InterTrial Interval* » en anglais, ITI) de 2 à 2.5 secondes, durant lesquels l'écran devenait noir. Dans le cas d'une erreur, les essais étaient momentanément abandonnés et suivis d'un intervalle de temps avant l'essai suivant.

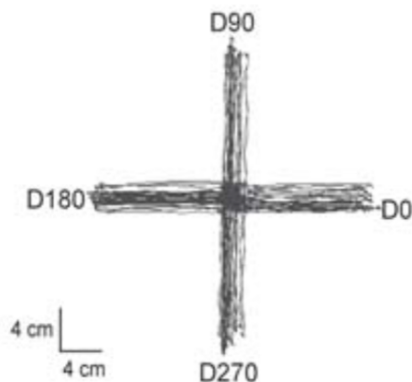


FIGURE 1.9 – Exemples de trajectoires du mouvement exécutées durant 50 essais réussis. Les valeurs 0, 90, 180 et 270 font référence aux valeurs des angles d'affichage des cibles par rapport au centre [Pasquereau et al., 2007].

TABLE 1.1 – Combinaison des couples de cibles affichées à l'écran durant la tâche comportementale.

Combinaisons des cibles	
P(R) = 1	vs P(R) = 0
P(R) = 1	vs P(R) = 0.33
P(R) = 1	vs P(R) = 0.67
P(R) = 0.67	vs P(R) = 0.33
P(R) = 0.67	vs P(R) = 0
P(R) = 0.33	vs P(R) = 0

Entraînement Pendant les premières étapes de l'apprentissage, une seule cible était affichée pour les deux singes (celle à $P(R) = 1$). Elle pouvait apparaître aléatoirement dans chacune des quatre positions possibles. Une fois les mouvements appris par les singes, une autre cible associée à une autre probabilité de récompense ($P(R) = 0.67$, puis 0.33 , puis 0) était présentée. Cette répétition permettait aux animaux d'apprendre à relier une cible à une probabilité de récompense. Enfin, les six combinaisons de cibles étaient utilisées aléatoirement et les singes étaient libres de faire la sélection de leur choix.

Conclusions de l'étude Les auteurs ont montré que l'activité de certains neurones d'entrée et de sortie des BG pouvait être modulée simultanément par des paramètres de mouvement (direction) et cognitifs (probabilité de récompense), avant et après une directive de mouvement. Ils ont également confirmé l'encodage dynamique des paramètres d'une tâche motrice par les neurones des BG.

1.2.2.2 Travaux de Piron et al. (2016)

Afin d'aller plus loin dans l'étude expérimentale des BG et de leur implication dans les processus d'apprentissage et de prise de décision, une étude plus récente [Piron et al., 2016] a utilisé une variante de la tâche vue précédemment. Le modèle computationnel développé au cours de notre étude a également été soumis à une partie de cette tâche afin de tester son efficacité et ses résultats seront présentés ultérieurement.

Thème général de l'étude Les données expérimentales ont tendance à montrer que les BG sont impliqués dans les comportements orientés vers un but ainsi que dans les comportements routiniers (automatiques), mais il arrive parfois que les observations cliniques lors des tests expérimentaux ne concordent pas. En effet, des lésions ou perturbations par stimulation cérébrale profonde du GPi ont été utilisées dans des buts thérapeutiques variés, allant de l'amélioration de la dystonie au traitement du syndrome de Gilles de la Tourette mais aucune de ces approches n'a signalé de déficience particulière dans les mouvements automatiques ou orientés vers un but auxquels on pourrait s'attendre.

Pour faire face à cette contradiction apparente entre la théorie et l'observation, les auteurs ont conçu un paradigme expérimental combinant le comportement de choix routinier, la prise de décision volontaire et l'apprentissage procédural. Les expériences ont été menées sur des primates non humains avec une inactivation pharmacologique spécifique du GPi. Les résultats ont été reproduits *in silico* en ajoutant une compétition latérale et un apprentissage Hebbien au niveau des couches corticales du modèle computationnel.

Animaux Deux singes femelles (*Macaca mulata*) de 4.9 et 5.6 kg ont participé à cette étude. Les macaques ont été, comme lors de l'étude précédente, gardées en conditions de restriction hydrique et avec nourriture à volonté afin d'accroître leur motivation à travailler.

Les animaux ont été entraînés à réaliser une variante d'un « bandit à deux bras » (« *2-armed bandit task* »), mais avec des probabilités de récompense différentes. Durant les essais suivants, les possibilités de choix entre les cibles à récompenses familières (favorisant la routine) ont été alternées avec de nouvelles cibles afin de faire intervenir un processus d'apprentissage volontaire.

Tâche comportementale Les singes ont été entraînés de jour dans des salles expérimentales qui leurs sont familières. Dans cette salle, quatre boutons sont placés sur un tableau de bord à quatre différentes positions (0°, 90°, 180° et 270°). Un bouton supplémentaire est placé en position centrale et sert à détecter le contact avec la main des singes (Figure 1.10). Ces quatre premiers boutons correspondent aux quatre positions possibles du curseur sur l'écran. Les primates sont assis sur des sièges en face de cet écran vertical, à une distance de 50 cm. Elles initialisent un essai en gardant leur main appuyée sur le bouton central, ce qui fait apparaître le curseur en position centrale de l'écran. Après une période de durée aléatoire (0.5-1.5 s), deux cibles apparaissent à deux des quatre positions possibles (les dites positions sont déterminées aléatoirement à chaque essai). Les deux conditions expérimentales (Condition de Routine RC et Condition Nouvelle NC) sont alternées par blocs de 10 essais. Dans les RC, les animaux ont été entraînés avec les deux cibles RC1 et RC2 (8 mois pour l'animal 1 et 13 mois pour l'animal 2).

Chaque cible était associée à une probabilité de récompense fixe ($P_{RC1}=0.75$ et $P_{RC2}=0.25$) qui ne changeait pas pendant et entre les sessions. Dans les NC, deux nouvelles cibles sont présentées et ont également une probabilité de récompense fixe ($P_{NC1}=0.75$ et $P_{NC2}=0.25$). Les associations cible-probabilité sont inchangées durant les sessions, mais changées entre les sessions. Une fois que les cibles sont affichées, les singes ont une fenêtre de temps aléatoire (0.5-1.5 s) pour appuyer sur le bouton associé à la cible choisie. Cela déplace le curseur sur la cible sélectionnée et les animaux doivent maintenir la pression sur le bouton pendant 0.5 à 1.5 s. Après ce délai, les singes sont récompensés de 0.3 ml d'eau ou non suivant la probabilité de récompense associée à la cible choisie. À la fin de l'essai, un signal spécifique correspondant à la disparition du curseur est donné, indiquant aux primates que l'essai est terminé et qu'ils peuvent en commencer un nouveau, avec un délai inter-essai de 0.5-1.5 s.

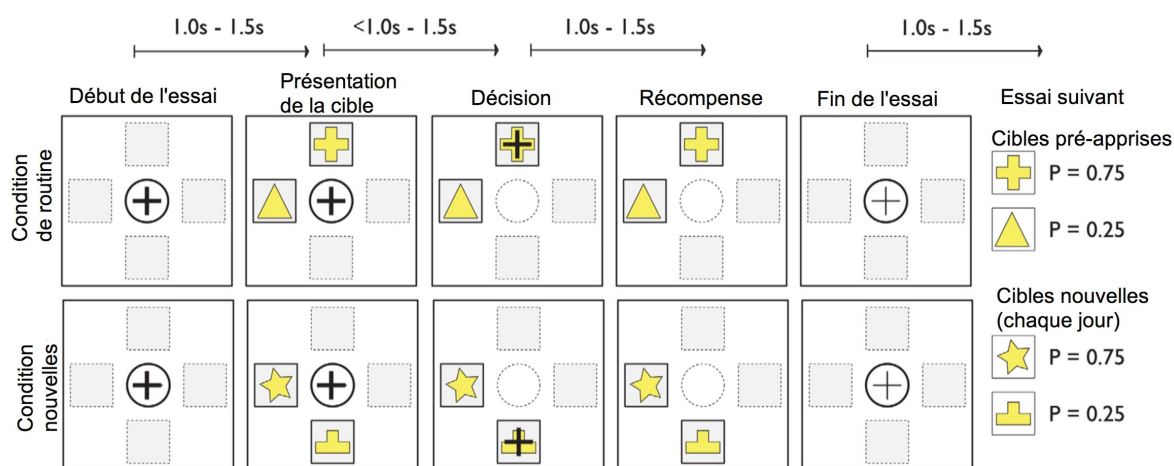


FIGURE 1.10 – Schéma de la tâche comportementale à deux cibles adaptée de [Piron et al., 2016]. Une session consiste en une série de 250 essais au minimum, découpés alternativement en blocs de 10 essais de routine (haut de la figure) ou de conditions nouvelles (bas de la figure). À chaque essai, deux cibles sont présentées simultanément et aléatoirement à deux des quatre positions possibles sur l'écran. Le singe doit exprimer son choix en bougeant un curseur vers l'une des cibles et est récompensé par 0.3 ml d'eau avec une probabilité prédéfinie fixe, dépendante du choix. Dans les conditions de routine (haut de la figure), les cibles ($P_{RC1}=0.75$ et $P_{RC2}=0.25$) sont celles avec lesquelles les singes ont été entraînés et sont familiarisés. Cette condition permet d'avoir des conditions de sélection de l'action automatiques et réflexes. Dans les nouvelles conditions (bas de la figure), les cibles ($P_{NC1}=0.75$ et $P_{NC2}=0.25$) ont les mêmes probabilités de récompense, mais les paires sont changées (nouvelles formes et couleurs) à chaque session. Cela permet d'inclure à l'étude des capacités d'apprentissage et de prise de décision délibérées, orientées vers un but.

Inactivation bilatérale du GPI Pour les deux animaux, une micro-injection de muscimol hydrobromide (agoniste du $GABA_A$) ou de solution saline (NaCl 9%) est réalisée bilatéralement 15 minutes avant chaque session. Cette inactivation bilatérale du GPI empêche les animaux d'apprendre de nouvelles associations cible-récompense, alors que les performances acquises avant l'injection demeurent intactes bien que plus lentes.

Modèle computationnel Le modèle employé est de celui de Guthrie [Guthrie et al., 2013] auquel est ajouté un apprentissage hebbien (potentialisation à long terme, « *Long Term Potentialisation* », LTP) au niveau cortical qui se produit entre les parties cognitives et motrices. Il

est implémenté grâce à la librairie DANA [Rougier and Fix, 2012] sur un formalisme de type modèle fréquentiel (par opposition aux modèles à « *spiking neurons* »).

Ce modèle théorique des BG se montre capable de reproduire les résultats observés chez les singes dans ce protocole, en conditions normales et lorsque le GPi est lésé. Dans ce dernier cas, les retours cortico-basaux sont supprimés et le modèle est alors incapable d'apprendre en Condition Nouvelle. Cependant, il est encore capable (mais plus lentement) de sélectionner lorsque la valeur de récompense associée à la cible a été apprise au niveau cortical par des mécanismes Hebbiens.

Conclusion de l'étude Les auteurs ont montré que des conditions nouvelles augmentent le temps de réaction, tout comme l'inactivation du GPi (mais en conditions normales et de routine). De plus, l'inactivation du GPi ne nuit pas au comportement de routine car le taux de succès à la tâche ne diminuait pas significativement. Cette inactivation du GPi porte en revanche atteinte à l'apprentissage des nouvelles conditions.

Les BG jouent donc un rôle critique dans les processus délibératifs supportant l'apprentissage, mais ne semblent donc pas nécessaires pour l'expression des mouvements de routine. L'approche des auteurs appliquée ici prédit qu'après une pallidotomie ou durant la simulation, les patients devraient avoir des difficultés avec les processus de prise de décision complexes ou avec l'apprentissage de nouveaux comportements orientés vers un but.

1.2.2.3 Modèles d'étude chez l'amphibien

La physiopathologie des BG (à travers l'étude de la maladie de Parkinson) a été également étudiée chez un autre modèle animal peu répandu : la salamandre [Parish et al., 2007]. Les auteurs ont voulu vérifier ici la possibilité d'une régénération des neurones dopaminergiques dans le cerveau de vertébrés adultes. Cette espèce présente l'avantage d'avoir d'importants mécanismes de plasticité cellulaire et de réparation des tissus fonctionnels. En effet, ces animaux ont la capacité de régénérer de grandes parties de leurs corps, comme des membres entiers ou encore leur moelle épinière. Cette régénération pourrait impliquer à la fois une activation des cellules souches et une reprogrammation des cellules différenciées, en fonction du site de la blessure. Afin de mieux comprendre ce phénomène, les capacités de régénération cérébrale et de neurogenèse des salamandres ont été étudiées chez des animaux sains et rendus parkinsoniens (au travers de lésions dopaminergiques réalisées avec de la 6-hydroxydopamine (2,4,5-trihydroxy phenylethylamine ou 6-OHDA en abrégé). Cette drogue est une neurotoxine utilisée classiquement dans la recherche scientifique pour léser sélectivement les neurones dopaminergiques [Breese et al., 2005]). Les lésions à la 6-OHDA vont ainsi induire un déficit de DA et provoquer un syndrome parkinsonien chez les animaux traités.

Thème général de l'étude Comme nous l'avons déjà évoqué, la maladie de Parkinson se caractérise notamment par la mort et le manque de régénération fonctionnelle des neurones dopaminergiques du mésencéphale, la diminution des entrées dopaminergiques dans le striatum, ainsi que des mouvements anormaux. Un moyen pour favoriser la récupération serait d'induire ou d'améliorer la neurogenèse dopaminergique. Ici, les auteurs décrivent la création d'un modèle

de salamandre 6-OHDA parkinsonienne afin d'étudier la régénération dopaminergique dans le mésencéphale.

Animaux Les animaux utilisés ici sont des urodèles (*Notophthalmus viridescens*) adultes, également appelés tritons verts à points rouges. Ils étaient gardés dans un environnement à température fixe (25 °) et nourris de façon hebdomadaire.

Protocole expérimental Le comportement des salamandres a été observé dans des conditions similaires à celles de leur environnement naturel, en période de jour uniquement. Les animaux étaient placés sur des assiettes en plastiques de couleur claire (situation aversive pour ces animaux), de 15 cm de diamètre et avec des bords de 1,5 cm de haut. Un quadrillage de 3 cm a été dessiné au fond des assiettes avant d'y ajouter 50 ml d'eau.

Les salamandres ont été divisées en trois groupes : témoins, sham et lésés. Dans les deux derniers cas, elles étaient anesthésiées avant de subir une micro-injection (solution saline ou 6-OHDA suivant le groupe) dans le troisième ventricule, juste au niveau des neurones dopaminergiques diencéphaliques et mésencéphaliques.

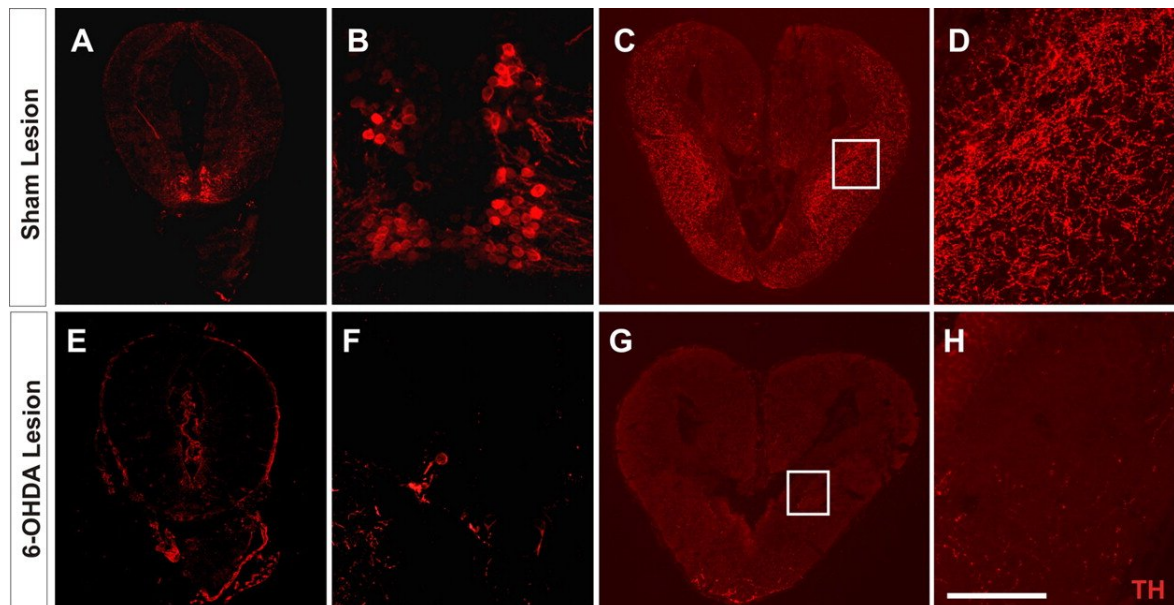


FIGURE 1.11 – L'injection stéréotaxique de 6-OHDA cause une perte cellulaire dopaminergique dans le diencéphale et le mésencéphale, d'après [Parish et al., 2007]. Les coupes du haut sont réalisées chez un animal du groupe sham (lésé à la solution saline) et celles du bas chez un animal lésé à la 6-OHDA. Ici l'analyse immunohistochimique a été réalisée avec un anticorps anti tyrosine hydroxylase (TH, enzyme impliquée dans la biosynthèse de la DA). La présence de cellules ou fibres TH⁺ indique celle de neurones dopaminergiques. (A, B) Photomicrographies des cellules TH⁺ au niveau de l'aire tegmentale du mésencéphale des tritons. (C) Fibres TH⁺ au niveau du striatum d'un animal lésé à la solution saline. (D) Agrandissement de la zone encadrée en C. (E, F) même région que pour A et B mais chez un triton lésé à la 6-OHDA. (G) Fibres TH⁺ au niveau du striatum chez un animal lésé à 97% trois jours après l'injection de 6-OHDA. (H) Agrandissement de la zone encadrée en G. Barre d'échelle : 1 mm en A C, E et G, 150 μm en B, F et J, 250 μm en D et H.

Le jour des tests comportementaux, les animaux étaient placés au centre de l'assiette et enregistrés par caméra vidéo. Ils étaient observés pendant 1 minute toutes les 10 minutes,

pendant une période de 50 minutes. Afin de suivre l'évolution de la perte dopaminergique, les mouvements des salamandres ont été quantifiés en comptant le nombre de fois où leur queue franchissait une ligne du quadrillage pendant la nage. La nage est un mouvement très caractéristique de l'état des tritons et peut être décrit comme une onde sinusoïdale commençant au niveau de la tête et se propageant jusqu'à la queue. De plus, le groupe cellulaire lésé avec la 6-OHDA ne projette pas dans la moelle épinière des salamandres et la lésion dopaminergique n'est donc pas censée affecter le mouvement de la queue.

Résultats Les tests immunohistochimiques ont montré une forte diminution du nombre des neurones dopaminergiques après lésion avec la 6-OHDA. En effet, les tritons lésés avec cette neurotoxine avaient une amplitude de mouvement moindre par rapport à ceux lésés avec la solution saline ou par rapport au groupe témoin. L'activité motrice a été mesurée pendant un mois, et la plupart des salamandres atteignaient un score moteur normal à la fin de la période de test. La cinétique de l'amélioration du comportement a suivi celle de la réapparition du nombre de neurones (Figure 1.11). Les réponses comportementales étaient en forte corrélation avec le nombre de neurones lésés, ce qui fait que le degré de lésion pouvait être estimé grâce à ces réponses. Ainsi, les tritons étaient capables de régénérer leur système dopaminergique, ce qui a été confirmé par des analyses cellulaires et fonctionnelles.

Conclusions de l'étude Les auteurs ont démontré une régénération robuste et complète du système dopaminergique après élimination de ses neurones. Cette neurogénèse dopaminergique conduit à une restauration histologique en permettant la régénération, ainsi qu'à une récupération complète du comportement moteur. Des analyses moléculaires du motif d'expression temporel des déterminants dopaminergiques indiquent que la régénération des neurones dopaminergiques se fait lors d'un long protocole de développement, de façon similaire à celui de leurs homologues mammifères lors de l'embryogenèse. Le modèle de salamandre donne un aperçu du mécanisme de régénération/neurogénèse dopaminergique et pourrait contribuer au développement de nouvelles stratégies de régénération pour le cerveau des mammifères.

Nous avons décidé d'inclure cette étude dans notre synthèse bibliographique car elle a servi de point de départ à l'élaboration d'une tâche comportementale de prise de décision chez ce même modèle animal. La mise au point de cette tâche qui a débutée avant que nous y apportions notre contribution, sera détaillée plus loin dans le manuscrit.

1.3 Les différents types d'apprentissage et leurs corrélats biologiques

Comme nous l'avons déjà mentionné en introduction, l'apprentissage est une des propriétés les plus intéressantes des réseaux de neurones biologiques. C'est toujours le cas lorsque ce processus devient applicable à des réseaux de neurones artificiels (« *Artificial Neural Networks* » ou ANN, en anglais). Nous allons donc étendre la définition initiale de l'apprentissage que nous avons mentionnée précédemment (cf. 1.1.1) afin de l'appliquer aux ANN. Le processus d'apprentissage avec les ANN peut en effet se traduire par la mise à jour de l'architecture d'un

réseau et des poids de ses connexions pour que ledit réseau puisse effectuer une tâche cognitive, comportementale ou motrice spécifique [Jain et al., 1996]. En général, le réseau doit apprendre les poids des connexions synaptiques à partir de modèles d'entraînement, et les performances sont améliorées au cours du temps en mettant ces poids à jour à l'intérieur du réseau. En d'autres termes, les ANN sont capables d'apprendre plus ou moins automatiquement à partir d'exemples, ce qui les rend très attractifs (voir la section 1.4.1.1 pour plus de détails sur les ANN).

Pour comprendre ou concevoir les processus d'apprentissage, il faut tout d'abord définir un environnement de travail, c'est à dire le type d'ANN utilisé et ses caractéristiques. La conception de cet environnement de travail est en général guidée par la nature et les contraintes de la tâche à effectuer. Ensuite, il faut comprendre la façon dont les poids sont mis à jour dans ce réseau, c'est à dire les règles d'apprentissage à appliquer dans le contexte étudié. Un algorithme d'apprentissage consiste ensuite en une procédure qui met à jour les poids dans le réseau, en fonction des règles d'apprentissage appliquées.

L'apprentissage machine au sens large peut être subdivisé en trois grandes catégories :

- l'apprentissage supervisé
- l'apprentissage par renforcement
- l'apprentissage non-supervisé

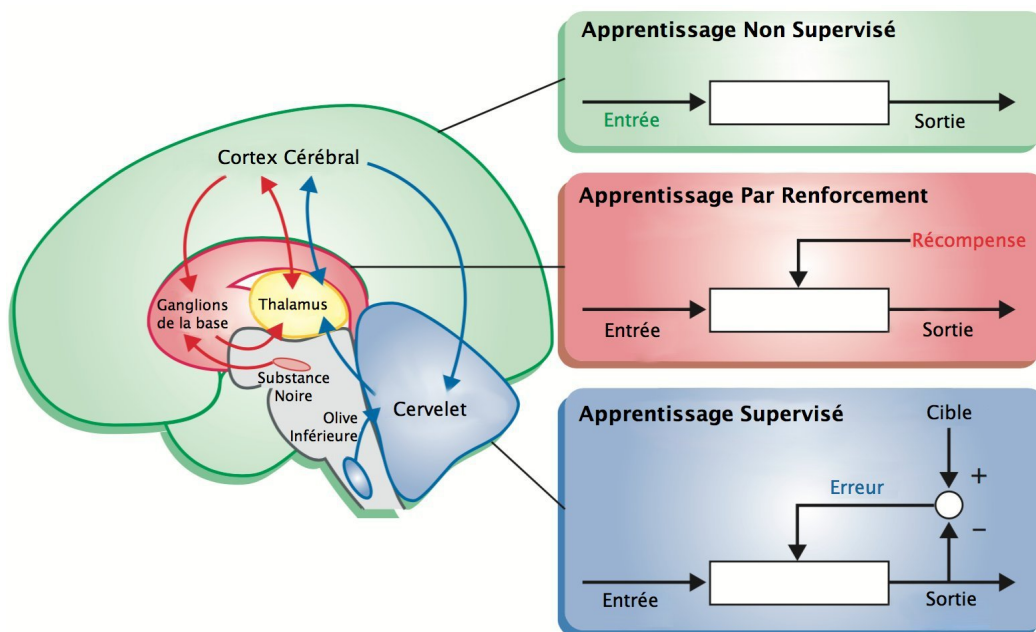


FIGURE 1.12 – Spécialisation du cervelet, des BG et du cortex en fonction des types d'apprentissage, d'après [Doya, 2000]. Le cervelet (en bleu) serait spécifique de l'apprentissage supervisé, qui est guidé par un signal d'erreur encodé par les entrées des fibres grimpantes provenant du noyau olivaire inférieur. Les BG (en rouge) seraient spécialisés dans l'apprentissage par renforcement, guidé par un signal de récompense encodé par les entrées dopaminergiques provenant de la substance noire. Le cortex (en vert) serait spécifique de l'apprentissage non-supervisé, qui lui est guidé par les propriétés statistiques du signal d'entrée lui-même. Il peut également être régulé par les entrées neuromodulatrices ascendantes.

Ces trois catégories possèdent chacune au moins un corrélat biologique (Figure 1.12) qui sont respectivement le cervelet, les BG et le cortex [Doya, 1999, Doya, 2000]. Dans l'apprentissage supervisé, la réponse exacte attendue en sortie par rapport à une entrée donnée est indiquée à l'algorithme. Dans l'apprentissage non-supervisé au contraire, aucune réponse exacte n'est

suggérée et le système apprend « à l'aveugle ». L'apprentissage par renforcement quant à lui, peut être décrit comme un mode intermédiaire d'apprentissage automatique dans lequel le système apprenant va devoir s'adapter pour répondre en fonction d'informations partielles à un contexte dans lequel il devra optimiser ce qu'il va considérer comme des récompenses. Cette brève description va être maintenant détaillée dans les sections suivantes.

1.3.1 Apprentissage supervisé

1.3.1.1 Principe

Dans l'apprentissage supervisé, le réseau va faire des prédictions fondées sur un ensemble d'exemples vus lors de son entraînement. Pour cela, il va recevoir un exemple de réponse correcte désirée en sortie, en fonction d'une entrée donnée [Doya, 1999, Jain et al., 1996, Bar-Gad et al., 2003]. Le réseau va donc créer une sorte de carte d'entrées-sorties correspondantes, en utilisant des couples de vecteurs d'entrées-sorties de la forme (x_t, y_t) . Pour chaque vecteur d'entrée x_t , il y aura un vecteur de sorties désirées y_t correspondant. La carte d'entrées et sorties sera donc construite d'après les exemples de sorties désirées en fonction des entrées données. Les poids synaptiques dans le réseau vont être modifiés pour que ce dernier puisse produire des réponses les plus proches possibles des réponses exactes attendues, comme lors de la phase d'entraînement. Il va donc y avoir une constante mise à jour des paramètres du réseau pour modifier les valeurs des poids synaptiques afin de minimiser l'erreur réalisée par le réseau vis-à-vis de la sortie désirée. Après la phase d'apprentissage, le réseau va être en mesure de prédire la sortie par rapport aux données d'entrée avec un risque d'erreur minimum dont la valeur dépendra des conditions d'apprentissage. Cette caractéristique est particulièrement intéressante et va au delà de l'architecture d'une simple table de correspondance par exemple (en anglais « *look-up table* » ou LUT). En effet, un ANN correctement supervisé est capable par exemple de donner un bon résultat à partir d'une information partielle en entrée. L'apprentissage supervisé appliqué à un ANN peut ainsi le rendre apte à la reconnaissance des invariants et donc de formes partielles ou déformées (visages, écriture ...) de la même façon qu'un individu est capable d'identifier le contenu d'un texte manuscrit malgré la grande variabilité dans la reproduction des caractères qui le composent.

1.3.1.2 Apprentissage fondé sur l'erreur dans le cervelet

Le cervelet se caractérise par une structure de type « *feed-forward* » et contient différents types de cellules (Figure 1.13), parmi lesquelles les cellules de Purkinje, les fibres moussues et les fibres grimpanes [Raymond et al., 1996, Albus, 1971, Marr, 1969, Doya, 1999, Kawato and Gomi, 1992]. Les cellules de Purkinje représentent la sortie du cortex cérébelleux et projettent, au travers de connexions inhibitrices, dans les noyaux cérébelleux profonds, qui vont eux-même servir de sortie aux autres régions cérébrales vers le reste du CNS. Le cervelet reçoit des entrées de la part des fibres grimpanes et des fibres moussues, formant deux voies bien distinctes physiologiquement et anatomiquement. Les entrées des fibres grimpanes et parallèles vont converger vers les cellules de Purkinje. L'activation des entrées de ces deux types de fibres va induire une modification des poids synaptiques entre les fibres parallèles et les cellules de

Purkinje, diminuant ou augmentant à court ou long terme l'activité de ces cellules de Purkinje. La variation des poids des synapses est sous le contrôle de mécanismes de plasticité dépendante du temps (« *Spike-Timing Dependent Plasticity* », STDP). La variation à long terme de la force des synapses des fibres parallèles projetant vers celles des cellules de Purkinje serait à la base du processus d'apprentissage supervisé du cervelet [Doya, 2000, Houghton, 2014, Piochon et al., 2012]. Ce sont ici les signaux issus des fibres grimpantes provenant de l'olive inférieure qui délivreraient un signal d'erreur en cas de mauvaise correspondance entre les entrées-sorties.

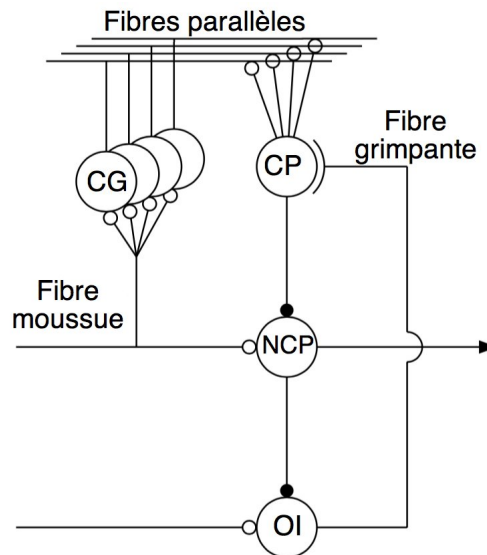


FIGURE 1.13 – Diagramme du circuit cérébelleux, d’après [Doya, 1999]. CG : Cellules Granulaires ; CP : Cellules de Purkinje ; NCP : Noyaux Cérébelleux Profonds ; OI : Olive Inférieure. Les flèches vides représentent les connexions excitatrices et les flèches pleines les connexions inhibitrices.

Les modules⁵ d'apprentissage supervisé dans le cervelet peuvent être utilisés comme un modèle interne de l'environnement. Un tel module peut servir à convertir et interpréter les processus physiques du monde extérieur sous la forme d'entrées pouvant être traitées et analysées. Dans un modèle computationnel, cela peut représenter les entrées sensorielles qui détermineront l'action à réaliser par exemple. Les modules d'apprentissage peuvent également être utilisés comme des raccourcis pour cartographier les entrées-sorties acquises dans d'autres parties du cerveau. En résumé, l'apprentissage supervisé est caractérisé par la mise à jour de paramètres (souvent le poids de la synapse) sur la base de la corrélation entre l'erreur de sortie post-synaptique et l'entrée pré-synaptique [Doya, 1999].

Parmi les ANN à apprentissage supervisé, nous pouvons citer par exemple les perceptrons (mono ou multi-couches) auxquels sont appliqués des règles d'apprentissage de type Hebb ou Widrow-Hoff ainsi que les algorithmes de rétro-propagation. Le modèle initial du CMAC⁶ [Albus, 1975] mais aussi les récentes techniques de « *deep learning* » font également appel à ce mode d'apprentissage supervisé et [LeCun et al., 2015]. On pourra consulter la revue de [Jain et al., 1996] pour avoir de plus amples détails sur ce sujet.

5. Le terme « module » fait ici référence à une fonction d'apprentissage, un élément qui pourra par la suite être instancié et intégré dans un modèle computationnel.

6. Cerebellar Model Articulation Controller

1.3.2 Apprentissage par renforcement

1.3.2.1 Principe

L'apprentissage par renforcement est une variante de l'apprentissage supervisé, dans laquelle le réseau est entraîné à recevoir des indications sur la réponse à produire en fonction des entrées données⁷ [Sutton, 1984]. La réponse correcte attendue par rapport à l'entrée ne lui est pas donnée, mais suggérée, le plus souvent sous la forme d'une récompense lorsque le réseau produit une bonne réponse [Jain et al., 1996]. Lors de la phase d'entraînement, le sujet va apprendre à prédire les résultats de ses actions au travers d'essais, de récompenses et d'erreurs, ce qui est l'essence même de l'apprentissage par renforcement [Lee et al., 2012]. Le concept d'incertitude est donc consubstantiel de celui d'apprentissage par renforcement [Sutton and Barto, 1998].

Dans l'apprentissage par renforcement, l'algorithme choisit une action (sortie) en réponse à chaque stimulus (entrée) ou « spontanément » s'il s'agit d'un agent opérant. L'algorithme d'apprentissage reçoit ensuite un signal de récompense qui indique la qualité de la décision [Bar-Gad et al., 2003]. En fonction de ce signal, il va modifier son état afin de maximiser la fréquence des récompenses.

Quand le renforcement a lieu juste après l'action (« *immediate-reinforcement task* ») la notion de délai n'entre pas en ligne de compte (du moins dans l'algorithme) [Sutton, 1984]. Toutefois, une récompense peut, et c'est souvent le cas, nécessiter soit un temps d'obtention long, soit un grand nombre d'étapes intermédiaires pour être obtenue. En effet, un gain moyen à court terme peut avoir une valeur moindre si un gain beaucoup plus important est obtenu avec un délai supplémentaire. Le phénomène est alors nommé « *delayed-reinforcement task* » qui aboutit au problème bien connu du « *Temporal Credit Assignment Problem* ». Ce problème est résolu algorithmiquement par les méthodes de « *Temporal Difference learning* » [Sutton, 1984, Schultz et al., 1997, Bar-Gad et al., 2003]. Cet algorithme a été inspiré par des données comportementales sur la façon dont les animaux apprennent les prédictions et il permet notamment de fournir une interprétation de l'activité des neurones dopaminergiques au travers notamment de l'apprentissage dépendant de la récompense.

L'apprentissage par renforcement est une approche courante en robotique, où les signaux que reçoivent les capteurs représentent les données d'entrée pour l'algorithme qui doit choisir la prochaine action du robot. D'une façon générale, c'est une approche intéressante pour enseigner à une machine à réaliser une tâche dont l'objectif nous est connu, mais pas forcément tous les paramètres et toutes les étapes de sa réalisation.

1.3.2.2 Apprentissage par renforcement dans les BG

En première analyse, l'apprentissage par renforcement dans les BG est un processus fondé sur la recherche de récompenses. Il n'est du reste pas possible d'assimiler complètement la recherche des récompenses avec l'évitement des punitions (punition signifiant dans ce contexte tout ce qu'un individu cherche à éviter d'une façon générale) bien que le résultat soit globalement positif dans les deux cas. En effet ces deux processus, bien que voisins, n'empruntent pas exactement les mêmes circuits. En revanche, ils font tous les deux appel à la modulation dopaminergique

7. <https://webdocs.cs.ualberta.ca/sutton/book/ebook/the-book.html>

qui s'exerce aussi bien sur la plasticité synaptique [Lambot et al., 2016] (notamment au niveau cortico-striatal) que sur les propriétés dynamiques des neurones [Planert et al., 2013]. De plus, ils s'influencent réciproquement [Abraham et al., 2014]. Comme les supports expérimentaux de notre étude font appel uniquement à la notion de récompense directe (par exemple le jus de fruit donné au singe) nous laisserons de côté ce qui concerne les processus aversifs.

La récompense est un concept décrivant une valeur positive qu'un être peut assimiler à un objet, un comportement ou un état physique. La capacité d'associer une valeur prédictive positive ou négative à un contexte ou à une action est un élément clé de la survie. Même s'il est admis que certains stimuli sont recherchés ou évités *a priori* (ce qui relève de la nociception par exemple), ce concept est plastique et peut évoluer au cours du temps chez un individu à travers l'apprentissage par renforcement. Il semble en effet que les projections de l'amygdale baso-latérale (BLA) vers le NAc et vers l'amygdale centro-médiale soit le support des associations qui se créent entre un stimulus et sa « *valence* » positive ou négative [Namburi et al., 2015].

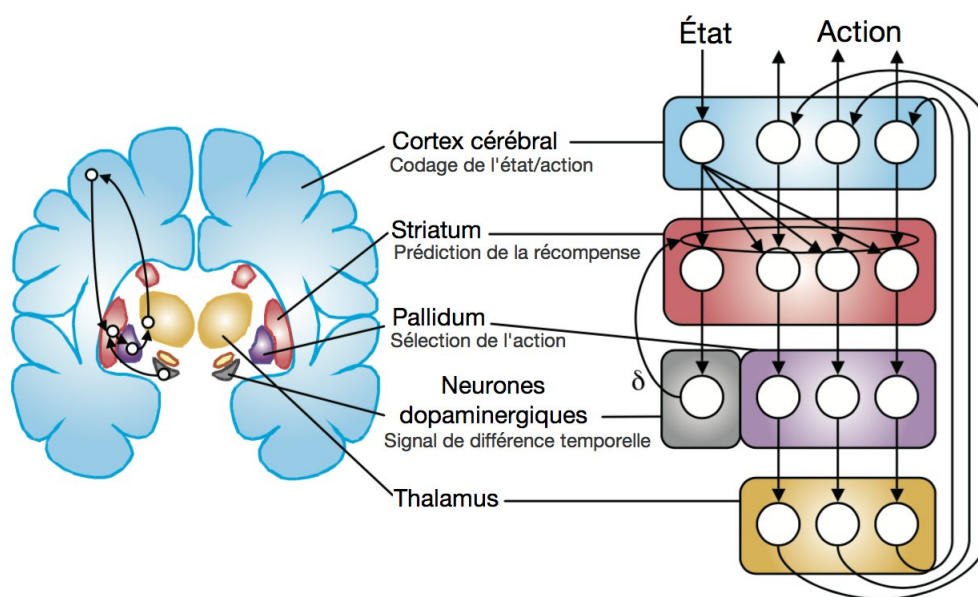


FIGURE 1.14 – *Modèle hypothétique de l'apprentissage par renforcement dans les BG, d'après [Doya, 2008]. Gauche* Vue coronale du cerveau. *Droite* modèles fonctionnels, où δ représente l'erreur de prédiction de la récompense véhiculée par les neurones dopaminergiques du mésencéphale.

Les récompenses peuvent être décrites en fonction du comportement désiré. Elles jouent également un rôle dans les renforcements positifs quand elles augmentent la fréquence des réactions comportementales pendant l'apprentissage et maintiennent les comportements appétitifs après l'apprentissage [Schultz et al., 1997, Doya, 1999, Sutton and Barto, 1998]. Comme cela a été précisé plus haut, la valeur de la récompense associée au stimulus n'est pas une propriété statique, intrinsèque au stimulus. En effet, différentes valeurs appétitives peuvent être associées à un même stimulus en fonction de l'état interne du sujet.

Dans les circuits des BG, les neurones dopaminergiques sont impliqués dans de nombreux processus de contrôle de l'apprentissage et de la mémoire comme le codage et la prédiction de la récompense, la motivation, la sélection de l'action, la consolidation mnésique et les comportements [Houk et al., 1995, Montague et al., 1996, Abraham et al., 2014, Doya, 2008].

Les neurones dopaminergiques de la VTA et de la SNc envoient leurs axones vers les

structures cérébrales impliquées dans la motivation et les comportements dirigés vers un but, comme le striatum, le NAc et le cortex frontal [Schultz et al., 1997]. Les neurones corticaux auraient la capacité de préserver les informations des entrées sensorielles [Doya, 2008], rendant le cortex responsable de leur analyse et de leur utilisation (Figure 1.14).

Ces neurones dopaminergiques possèdent également des connexions réciproques avec les neurones épineux striataux. Ces derniers sont impliqués dans la prédiction de la récompense et ce, avec un court intervalle de temps grâce au signal δ de différence temporelle des neurones dopaminergiques. Ce signal δ est également appelé signal de renforcement effectif car il va déterminer à lui seul s'il y aura un renforcement des circuits à l'origine de l'action ayant permis cette récompense [Rivest, 2009]. À travers ce mécanisme récursif, les neurones dopaminergiques seraient donc responsables de l'apprentissage de nouveaux comportements [Doya, 2000] en prédisant les récompenses.

1.3.3 Apprentissage non-supervisé

1.3.3.1 Principe

Lors de l'apprentissage non-supervisé, il n'y a pas de possibilité d'une association de la bonne réponse avec chaque motif d'entrée dans le jeu de données d'entraînement [Jain et al., 1996, Doya, 1999, Le et al., 2013]. Dans ce cas il n'existe en effet pas de « bonne réponse » *a priori*. L'objectif d'un algorithme d'apprentissage non-supervisé est d'organiser spontanément les données d'une certaine façon ou de trouver des relations pour décrire et généraliser leur structure.

Cela peut se manifester par des catégorisations des données d'un espace d'entrée en « *clusters* » ou encore par la recherche des différentes manières de visualiser des données complexes afin d'en simplifier l'affichage ou de l'organiser plus efficacement. Le réseau explore les structures sous-jacentes des données, les corrélations entre les motifs et les données et organise les motifs en catégories à partir de ces corrélations. Un exemple classique d'apprentissage non-supervisé est donné par les cartes auto-organisatrices de Kohonen⁸. Ces réseaux sont une des alternatives destinées à réduire les dimensions d'un espace de données en le transformant en un espace de représentation de plus faible dimension. L'auto-apprentissage s'effectue en modifiant les poids des neurones de la couche d'entrée vers la couche de sortie tout en conservant la topologie entre les deux espaces. En pratique, les connexions des neurones de la couche d'entrée à N dimensions sont modifiées de façon à ce que leur proximité en entrée les amène à converger vers un neurone commun en sortie. Pour éviter les biais liés à l'ordre de leur présentation, ces mises à jour des poids se font en général avec une permutation avant chaque itération (cette dernière technique est d'ailleurs appliquée à beaucoup d'algorithmes faisant intervenir des structures simulées en parallèle comme les modèles multi-agents [Breton et al., 2000]).

Les données servant à l'apprentissage seront donc uniquement des entrées, et leurs distributions impacteront le résultat du traitement dans la mesure où il n'y a pas d'indication sur la réponse souhaitée. En résumé, il s'agit pour le réseau de découvrir des structures cachées dans les données et de les généraliser à partir d'exemples et de retours connus.

8. <http://eric.univ-lyon2.fr/rias2006/presentations/VincentLemaire.pdf>

1.3.3.2 Apprentissage non-supervisé dans le cortex

Le cortex possède une organisation laminaire (Figure 1.15) et de nombreuses connexions récurrentes. Il est divisé en plusieurs aires fonctionnelles permettant de traiter les informations sensorielles, motrices ou contextuelles qui lui arrivent (cf. section 1.1.3). Les neurones corticaux sont généralement caractérisés par la spécificité de leurs réponses (orientation et couleurs des stimuli visuels dans le cortex visuel par exemple) [Doya, 1999, Le et al., 2013]. De la même façon, les synapses corticales suivent les règles de la plasticité Hebbienne : elles sont potentialisées quand l'entrée pré-synaptique est associée à la réponse post-synaptique, mais déprimées quand cette association ne se fait pas.

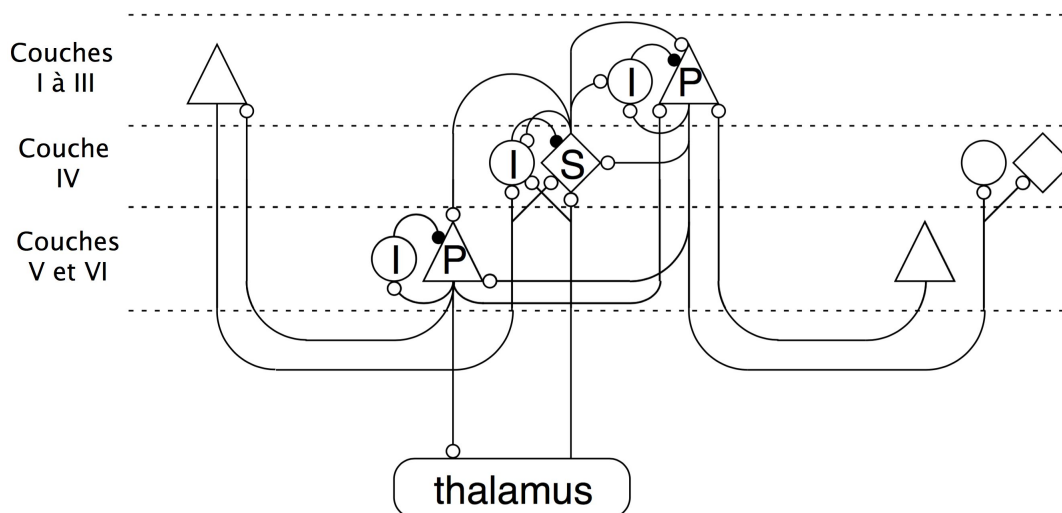


FIGURE 1.15 – *Diagramme du circuit neuronal cortical, d'après [Doya, 1999]. P : Neurones pyramidaux ; S : Neurones épineux étoilés ; I : Interneurones inhibiteurs. Les flèches vides représentent les connexions excitatrices et les flèches pleines les connexions inhibitrices.*

Toutes ces propriétés suggèrent que l'apprentissage non-supervisé permet le codage de l'information dans les aires corticales en mettant constamment les poids synaptiques à jour suivant les règles de la plasticité Hebbienne [Doya, 1999, Doya, 2000, Von der Malsburg, 1973]. Les modules d'apprentissage non-supervisé dans le cortex permettraient d'avoir une représentation simplifiée mais suffisamment caractéristique des états de l'environnement et du système comportemental (contexte).

En résumé, l'apprentissage non-supervisé est caractérisé par la plasticité Hebbienne et la réciprocité des connexions entre les aires corticales [Doya, 1999, Doya, 2000]. Dans ce type de plasticité, la force de la connexion entre deux neurones repose sur la corrélation temporelle de leur activité. Le fait d'utiliser ce type d'apprentissage dans un ANN va permettre à l'algorithme d'entraîner chaque couche du réseau « indépendamment » des couches précédentes [Sanger, 1989] et d'encoder les caractéristiques d'un espace d'entrée tout en réduisant la dimensionalité.

1.4 Approches théoriques et computationnelles des mécanismes de la prise de décision

Le cerveau humain a acquis de nombreuses spécialisations au cours de l'évolution et peut à ce stade traiter des données de représentation complexe. Mais l'individu reste plus que jamais soumis au problème du choix et à la gestion de l'incertitude.

Dès lors que des machines ont été capables de réaliser des comparaisons et des calculs et donc de prendre des décisions, elles ont été appliquées à suppléer le cerveau humain dans de nombreuses tâches mécaniques ou pouvant être automatisées ou formalisées. Les systèmes experts ont été parmi les premières implémentations de ces outils d'aide à la décision. Dès 1965 le logiciel Dendral aidait les chimistes à analyser les constituants d'un matériau à partir de données sur la spectrométrie de masse et la NMR. Des versions dérivées et améliorées de ces premiers algorithmes furent ensuite développées [Burlingame et al., 1973] puis appliquées à d'autres domaines comme MYCIN, outil d'aide à la décision pour les praticiens concernant l'identification des germes pathogènes et le choix de l'antibiothérapie la plus adaptée [Shortliffe, 1976]. Il a été relativement simple d'adjoindre à ces systèmes la possibilité de se mettre à jour spontanément en fonction de nouvelles données et cette évolution est, en apparence, assez similaire à ce qui se produit dans le cerveau humain quand celui-ci apprend au cours du développement de l'individu.

Toutefois, si l'on considère le CNS à l'échelle du neurone, celui-ci peut être assimilé à un système collaboratif multi-agent (MAS), massivement parallèle et capable de s'adapter en fonction de contraintes internes et externes à l'organisme qu'il habite afin de le faire survivre et prospérer. Si l'on veut pouvoir comprendre en quoi ce système collaboratif est nécessaire et suffisant pour accomplir cette tâche, il devient indispensable de se pencher sur des modèles capables de le décrire aussi précisément que possible et c'est ici que les ANN entrent en jeu.

1.4.1 Méthodes et outils pour la modélisation

1.4.1.1 Principe des réseaux de neurones artificiels

Les ANN sont initialement inspirés des réseaux de neurones biologiques et ont été créés à la fin des années 1950 dans le but de comprendre et d'imiter le fonctionnement du cerveau et de donner à la machine des éléments qu'elle ne possédait pas à savoir une capacité à apprendre et à généraliser reposant sur des paradigmes biologiques. Les travaux initiaux de Rosenblatt [Rosenblatt, 1958] sur le perceptron appliquaient pour la première fois la règle de Hebb à la plasticité synaptique et à l'apprentissage au sein d'un réseau de neurones formels. Ils connurent ensuite au cours des quatre dernières décennies successivement des pertes puis des regains d'intérêt et ont été appliqués jusqu'à aujourd'hui à des domaines variés tels que l'apprentissage machine, la tolérance d'erreur, le traitement du signal, la reconnaissance automatique et plus précisément au domaine qui nous intéresse : la modélisation de tout ou partie du CNS. Les ANN offrent ainsi une flexibilité et une puissance de calcul complémentaires des approches plus conventionnelles pour aborder ces questions [Jain et al., 1996] ainsi qu'un outil permettant d'explorer les caractéristiques et les possibilités des réseaux de neurones biologiques.

Les méthodes de construction des ANN utilisent des modèles mathématiques fondés sur le

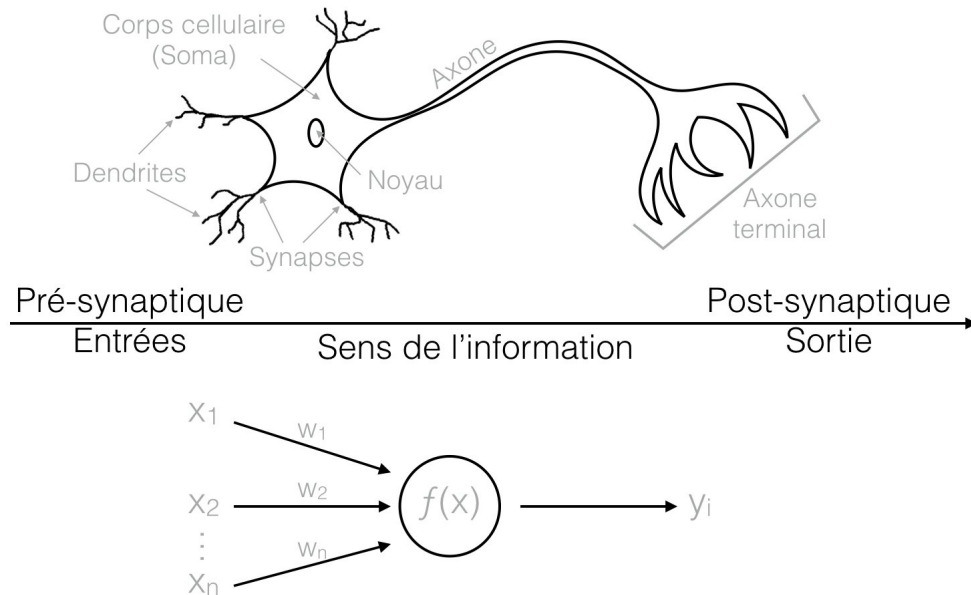


FIGURE 1.16 – Comparaison entre un neurone humain et un neurone artificiel. *Haut* un neurone est composé d'un corps cellulaire (ou soma) contenant un noyau, qui contrôle l'activité cellulaire. Les dendrites et l'axone vont émerger de ce soma et vont contribuer au transfert de l'information électrochimique entre les neurones (sauf cas particulier, les jonctions électriques ne sont que rarement prises en compte). D'une façon générale, l'information est émise au niveau initial de l'axone sous la forme d'un spike (composante électrique) qui va circuler jusqu'à son extrémité où il va déclencher la libération de neurotransmetteurs (composante chimique) ayant un effet sur les cellules cibles. Ces cellules cibles sont en général d'autres neurones mais peuvent être également d'autres types de cellules excitables, par exemple des cellules musculaires. *Bas* un neurone artificiel est un modèle mathématique, fondé sur le fonctionnement du neurone biologique. Par analogie avec ce dernier, il est composé d'une fonction de sommation des activités entrantes (dendrites principalement), d'une fonction d'activation contrôlant la propagation du signal (segment initial de l'axone) et prend en compte d'autres paramètres comme le nombre n de neurones x auxquels il est connecté (x_n) ou le poids w_n des synapses formant ces connexions.

fonctionnement du cerveau humain [Agatonovic-Kustrin and Beresford, 2000]. Tous les ANN ont au départ en commun le concept de « neurone » en tant qu'unité de traitement de l'information. Un neurone artificiel va être caractérisé par ses entrées x_i et sorties y_i , et va posséder un paramètre de poids synaptique w_i pour chacune de ses entrées x_i (Figure 1.16). Ces paramètres de poids vont représenter la force des connexions synaptiques (i) reliant une cellule source (pré-synaptique) et la cellule cible (post-synaptique). Ces neurones artificiels vont inclure également une fonction d'activation (ii) générant la sortie y_i du neurone, qui va, en fonction de cette activation, envoyer ou non (loi du tout ou rien) un signal à la cellule cible. Les différents formalismes de (i) et (ii) seront présentés dans 1.4.1.2.

Les ANN possèdent généralement une couche de neurones d'entrée, une ou plusieurs couches de neurones intermédiaires (couche cachée) et une couche de neurones de sortie. Les fonctions d'activations propres à chaque neurone du réseau vont contrôler la propagation du signal d'un neurone d'une couche à un neurone de la couche suivante (Figure 1.17). Cette fonction peut prendre en compte les poids et l'activation des synapses entre les neurones entre les différentes couches, un facteur signifiant ce poids en fonction de la nature excitatrice ou inhibitrice de la synapse, ou encore le seuil d'activation du neurone.

Pour traiter les informations d'entrées et de sorties à l'intérieur de ces ANN, des équations

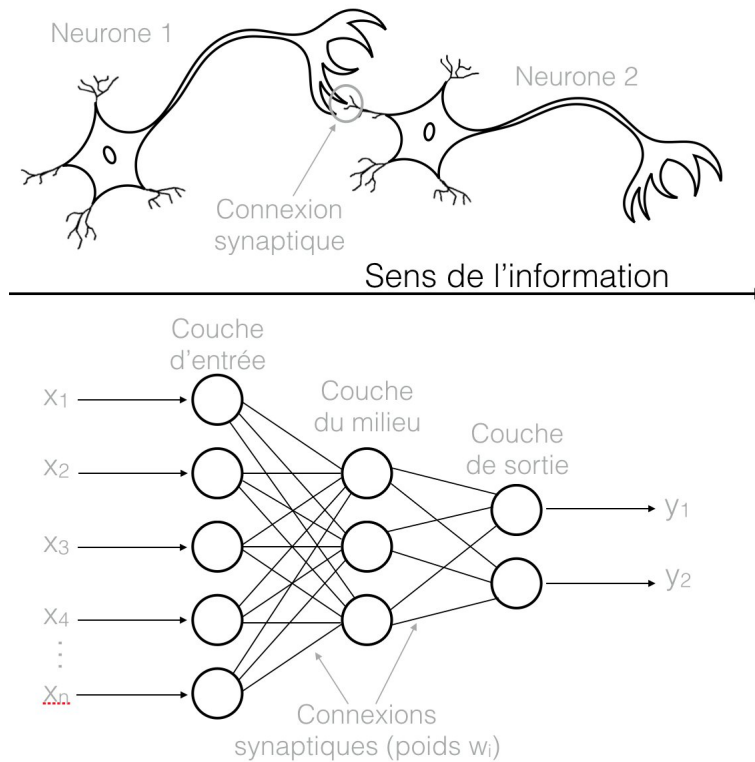


FIGURE 1.17 – Comparaison de réseaux de neurones biologiques (haut) et artificiels (bas). *Haut* Pour avoir un transfert de l'information entre les neurones 1 et 2, il va y avoir une connexion synaptique entre l'axone terminal (pré-synaptique) du neurone 1 et les dendrites (post-synaptique) du neurone 2. *Bas* Un réseau de neurones artificiels comporte une couche d'entrée, une ou plusieurs couches au milieu (cachées) et une couche de sortie. Ces couches sont composées d'un nombre donné de neurones et sont reliées par des connexions synaptiques de poids variable.

mathématiques sont utilisées. Ces équations vont comporter des paramètres (entrées) et générer un comportement (sortie) modélisant en cela le principe de l'association stimulus-réponse. Plusieurs fonctions d'activation (*i.e.* plusieurs équations) peuvent être utilisées dans un ANN pour représenter les différents types de neurones composant ce réseau. Il existe par ailleurs différents modes de connexion à l'intérieur des ANN qui peuvent aider à les catégoriser [Jain et al., 1996]. Ainsi les ANN sont souvent classés en deux grandes catégories : les réseaux « *feed-forward* » et les réseaux récurrents (ou de rétroaction, « *feedback* » en anglais). Dans les réseaux de type « *feed-forward* », les neurones sont organisés en couches avec des connexions unidirectionnelles entre elles, ce qui fait qu'il n'y aura pas de boucles dans ce genre de réseau (Figure 1.17). De plus, ils auront tendance à être statiques car ils ne produisent qu'une seule série de données de sortie pour une séquence de plusieurs valeurs données en entrée. Les réseaux récurrents vont présenter une dynamique plus complexe. En effet, les connexions dans ce type de réseau pouvant être réciproques, ils vont générer des boucles de rétroaction et donc une plus grande complexité et une plus grande richesse de configurations. Les sorties de certains neurones du réseau vont modifier dynamiquement les entrées d'autres neurones, amenant le réseau dans un nouvel état. Suivant l'architecture du réseau, il convient alors d'utiliser un algorithme d'apprentissage approprié pour gérer les mises à jour des différents paramètres du dit réseau suivant son état et en fonction de la tâche considérée.

1.4.1.2 Modèles mathématiques neuronaux et synaptiques

Comme cela a été évoqué précédemment, il existe de nombreux formalismes permettant de décrire et d'implémenter les ANN. Il est possible de les classer de plusieurs manières en fonction de leur utilisation ou en fonction des mécanismes qu'ils décrivent. Nous avons choisi de les classer ici en fonction de la complexité de leur fonction d'activation (ou fonction de transfert cf. 1.4.1.1) afin de présenter ceux qui nous sembleront les plus adaptées pour aborder notre étude.

Les neurones formels Les premiers modèles neuronaux implémentés datent comme nous l'avons mentionné précédemment des années 1940-1950. A cette époque la puissance de calcul disponible était modeste comparée à ce que l'on possède aujourd'hui mais déjà l'idée de décrire le CNS grâce à des montages électroniques était apparue comme une option intéressante à de nombreux mathématiciens, psychologues, physiciens et ingénieurs. Le premier modèle de neurone artificiel implémenté était un objet mathématique possédant une ou plusieurs entrées modélisant les synapses afférentes [McCulloch and Pitts, 1943]. Un exemple est représenté sur la (Figure 1.18)⁹.

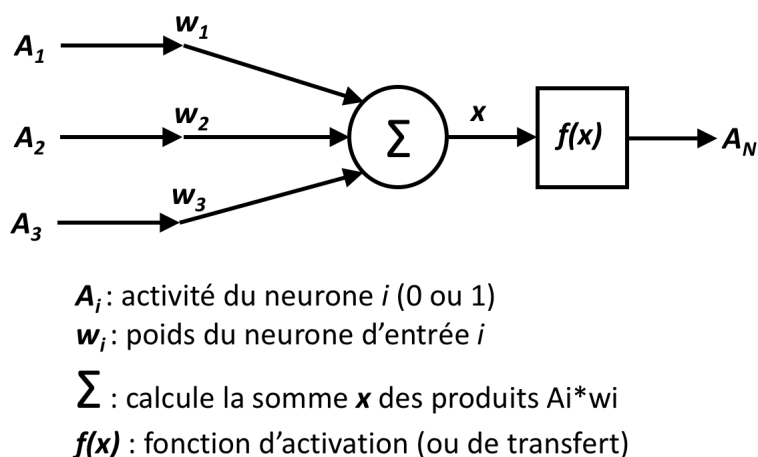


FIGURE 1.18 – Modèle initial du neurone formel : à un instant t donné la somme des entrées pondérées et signées ($\sum(A_i * w_i)$) est réalisée dans le soma d'un neurone N . Si cette somme x répond à un certain critère donné par la fonction d'activation $f(x)$ alors le neurone est à son tour activé et sa sortie A_N est mise à 1 ce qui correspond pour un neurone biologique à l'émission d'un potentiel d'action. Par défaut $A_N = 0$.

Le premier modèle de neurone formel réalisait donc la somme de ces entrées selon qu'elles étaient actives ou non et pondérées par leur poids synaptique. La somme obtenue était alors soumise à une fonction de transfert qui était souvent une fonction à seuil et qui, si et uniquement si le seuil était franchi, activait le neurone qui pouvait alors à son tour transmettre l'information en aval vers les autres neurones auxquels il était connecté. Ce mécanisme basique permet de mimer très simplement (mais de façon assez cohérente) ce qui se produit dans un neurone biologique et notamment la notion de « tout ou rien » pour l'émission d'un potentiel d'action. Les modèles synaptiques accompagnant ces premiers réseaux étaient simplement décrits par des poids signés (positifs pour des synapses excitatrices et négatifs pour des synapses inhibitrices) et l'absence ou

9. <http://www.ece.utep.edu/research/webfuzzy/docs/kk-thesis/kk-thesis-html/node14.html>

la présence d'une activité comme cela est indiqué dans la (Figure 1.18). Ce paradigme du neurone formel a donné naissance à plusieurs formalismes dérivés, à des applications en traitement du signal ainsi qu'à des modèles biologiques.

Le principal écueil à son utilisation pour décrire des modèles biologiques résidait cependant dans la gestion du temps. En effet, les réseaux de neurones formels tels qu'ils avaient été conçus dans les années 1940 étaient avant tout des automates logiques dont les états et transitions suivaient des séquences d'évènements discrets. Les évènements de sommation et d'émission de signaux (ou impulsions) ainsi que les mécanismes synaptiques n'y étaient donc pas directement superposables au temps mesuré dans les évènements biologiques. Cette question était d'ailleurs évoquée dans [McCulloch and Pitts, 1943] où les auteurs discutaient du délicat problème de la prise en compte du temps biologique dans leur modèle. Dans l'introduction, ils évoquaient la très grande vitesse de transmission des influx nerveux le long des axones, la durée très brève des potentiels post-synaptiques et des délais synaptiques et le problème de la période réfractaire. Ils mettaient de côté le rôle des propriétés cellulaires intrinsèques et finalement aboutissaient à un formalisme très rigoureux mais somme toute assez rigide du fonctionnement des circuits du tissu nerveux. Pour pouvoir introduire un temps « biologique » explicite dans les modèles une autre catégorie d'ANN a été progressivement développée que nous nommerons ici les réseaux de neurones à spikes (SNN) pour les distinguer des réseaux formels comme le perceptron, l'Adaline ou encore les réseaux de Hopfield. Ces modèles que nous décrirons succinctement dans les parties suivantes seront capables de simuler efficacement les effets du temps sur la dynamique des réseaux.

Le neurone Intègre et tire (« *Leaky Integrate and Fire* ou *LIF* ») Bien avant les travaux mentionnés dans le paragraphe précédent, le physiologiste français Louis Lapicque avait dès 1905 étendu les concepts existants à l'époque sur l'excitabilité cellulaire¹⁰ et posé les bases du premier modèle mathématique de cellule excitable [Lapicque, 1905]. Il fut à ce titre considéré comme l'un des pionniers de l'ère cybernétique et malgré les limites imposées par son principe de chronaxie¹¹, ses travaux et sa vision globale du CNS influencèrent durablement les neurosciences computationnelles¹². Ses travaux de neurophysiologiste et de médecin l'amènèrent d'ailleurs en 1952 à aborder la question de la conscience en tant que fonction cellulaire à travers les prémisses des neurosciences computationnelles qu'il avait contribuées à poser¹³.

Issu de ces travaux, le modèle LIF assimile la membrane cellulaire à un circuit R-C dans lequel $V(t)$ représente la valeur courante du potentiel de membrane, R_m représente la résistance membranaire au passage du courant et C_m la capacité de la membrane neuronale à la séparation de charges. Une batterie V_m est ajoutée au circuit pour simuler la capacité du neurone à maintenir la différence de potentiel de sa membrane à une valeur négative souvent proche de -65 mV. Enfin, une entrée en courant dans le circuit est ajoutée qui permet de simuler les stimulations du neurone (aussi bien les effets des synapses que les courants pouvant être imposés par les contingences expérimentales) (Figure 1.19).

10. http://www.persee.fr/doc/rhs_0048-7996_1960_num_13_3_3855

11. Constante de temps de l'excitabilité

12. <http://www.histcnrs.fr/ColloqDijon/Dupont-Lapicque.pdf>

13. http://wikilivres.ca/wiki/La_conscience_fonction_cellulaire

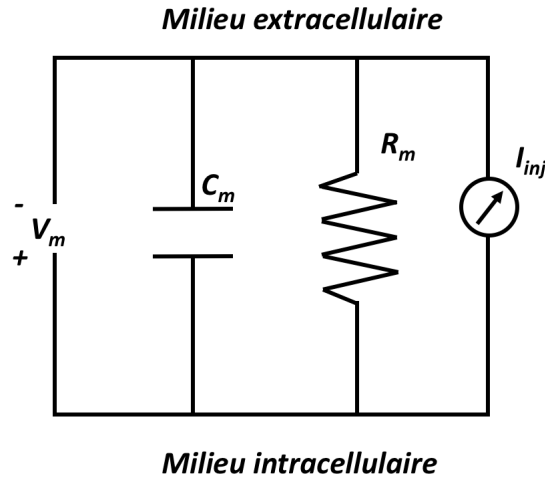


FIGURE 1.19 – Circuit équivalent à la membrane du modèle LIF : V_m représente le potentiel de repos de la membrane (V), C_m sa capacité (F), R_m sa résistance (Ohm), I_{inj} le courant pouvant être appliqué au neurone (A). D’après [Mahdi et al., 2013].

Ce modèle peut se résoudre numériquement et très simplement car l’équation différentielle qui simule le circuit est une équation différentielle ordinaire (ODE) avec une seule variable : $V(t)$. Dans le cas présenté dans la (Figure 1.19) l’équation peut s’écrire :

$$C_m \cdot \frac{dV(t)}{dt} = I(t) - \frac{V(t)}{R_m} \quad (1.1)$$

$$\text{si } V_m(t) \geq V_{seuil} \text{ alors } V(t) \leftarrow V_m \quad (1.2)$$

L’équation 1.1 rend compte de l’activité de la membrane mais n’est pas capable en l’état de simuler la genèse des potentiels d’action (ce circuit mime uniquement la charge et la décharge d’une capacité à travers un circuit résistif). La notion de seuil d’excitabilité évoquée par Lapique prend donc ici tout son sens. Un test est alors rajouté et décrit dans l’équation 1.2. Quand le seuil fixé par l’expérimentateur V_{seuil} est franchi, un spike est alors émis et le potentiel $V(t)$ est remis à sa valeur de repos V_m .

Les modèles neuronaux à deux variables Le neurone LIF présente de nombreuses qualités mais en l’état et avec une seule variable il n’est pas capable de reproduire les propriétés cellulaires intrinsèques des neurones. Les facultés d’adaptation, les effets rebond ou encore l’activité en « *bursts* » (bouffées de spikes) sont autant de propriétés de décharge des neurones biologiques que le modèle initial du LIF est incapable de manifester par lui-même. Comme dans la théorie neuronale l’information est véhiculée par les spikes (et les intervalles entre les spikes) et par ailleurs, d’importants évènements à l’échelle cellulaire et synaptique sont contrôlés également par les instants des spikes [Brette, 2015] il apparaît judicieux de pouvoir mimer ces activités au mieux afin que les modèles fondés sur des « *spiking neurons* » puissent être le plus plausible possible.

Des modèles prenant en compte une seconde variable ont donc été développés pour répondre à ce besoin. Les modèles d’Izhikevich [Izhikevich, 2003] et de Brette [Brette and Gerstner, 2005] pour ne citer qu’eux, incluent une seconde variable souvent nommée « *recovery variable* » et qui

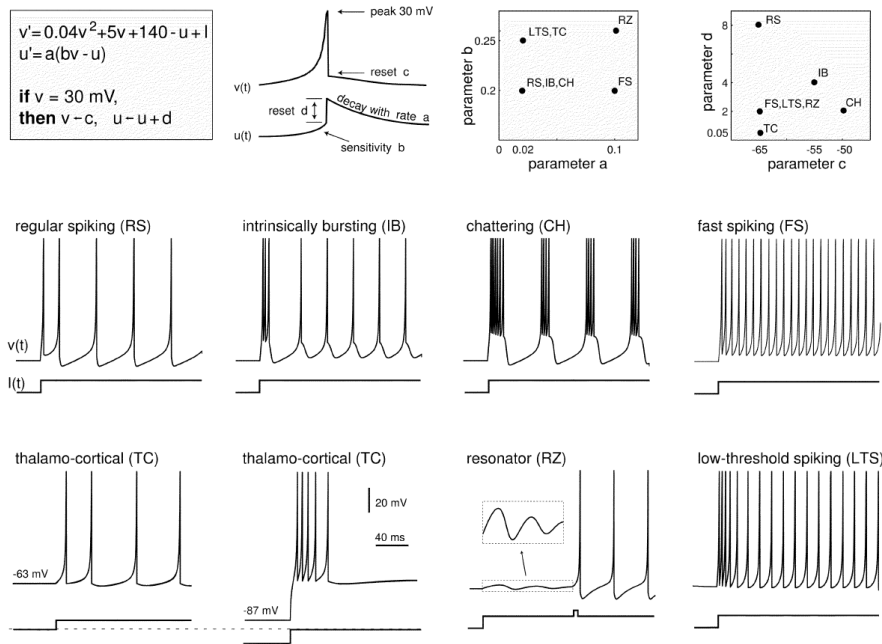


FIGURE 1.20 – Exemples d’activité de neurones d’Izhikevich : en haut à gauche dans l’encadré les équations qui permettent de modéliser le comportement d’un neurone. v' est la dérivée première du potentiel de membrane par rapport au temps, u' la dérivée première de la « recovery variable », I le courant appliqué et a , b , c et d les paramètres du modèle. Les trois figures à droite de l’encadré représentent l’effet d’un spike sur v' et u' ainsi que l’effet des valeurs des paramètres sur le mode de décharge du neurone. Les huit figures situées sous la première ligne représentent des exemples d’activité de décharge caractéristiques de neurones biologiques et obtenus après modification des valeurs des quatre paramètres. D’après [Izhikevich, 2003].

permet de simuler des patterns de décharge complexes. C’est grâce à une analyse mathématique de la dynamique de ces neurones que ces auteurs ont développé des modèles légers à implémenter et à calculer (seulement deux variables et assez peu de paramètres) et qu’ils ont calculé des valeurs pour lesquelles les neurones se comportaient d’une façon très proche de celle de leurs homologues biologiques. Un exemple bien connu est illustré dans la figure 1.20.

FitzHugh-Nagumo [Fitzhugh, 1955] et Morris-Lecar [Morris and Lecar, 1981] sont des cas particuliers de modèles neuronaux à deux variables. Ils ont du reste, inspiré les travaux ultérieurs d’Izhikevich et de Brette mais leur genèse est quelque peu différente car ils ont été conçus comme des simplifications du modèle de Hodgkin-Huxley (HHM). En effet ce dernier, dans sa version mono-compartimentale la plus simple, comprend tout de même quatre variables et un nombre conséquent de paramètres (cf. section 1.4.1.2). Les modèles de FitzHugh-Nagumo (FNM) et de Morris-Lecar (MLM) permettent d’approcher le comportement du HHM notamment dans l’espace des phases. Ces modèles restreints présentaient un intérêt tout particulier à la fois parce que leur approche analytique était simplifiée mais aussi et surtout parce que leur implémentation numérique était allégée à une époque où la puissance de calcul disponible était sensiblement plus faible mais en pratique ils se prêtent assez mal à l’étude que nous menons ici.

Il existe par ailleurs une différence importante entre ces deux modèles (FNM et MLM) et les précédents. En effet, tout comme le HHM ils émettent des spikes sans qu’il soit nécessaire de rajouter une condition de seuil. Alors que dans les précédents des équations différentielles étaient accompagnées d’un test logique pour déterminer la survenue d’un potentiel d’action, pour

le FNM et le MLM ce potentiel d'action est généré par les équations différentielles elle-mêmes. Il en résulte que, si le modèle apparaît plus cohérent d'un point de vue biologique (un vrai spike possède notamment une certaine durée), s'agissant de modèles de réseaux il est plus intéressant de pouvoir générer directement les instants des potentiels d'action. En effet pour les FNM, MLM et HHM, comme cet instant n'est pas explicite, il faut ajouter en plus une méthode de détection du potentiel d'action (comme un dépassement de seuil). Outre la définition toujours délicate de « l'instant » exact d'un potentiel d'action, cela ajouterait de l'ambiguïté, de la complexité et du temps de calcul sans bénéfice particulier *a priori* et c'est notamment pourquoi nous n'avons pas envisagé de nous tourner vers ce formalisme pour notre étude.

Hodgkin-Huxley et ses modèles dérivés En 1952 Alan Lloyd Hodgkin et Andrew Fielding Huxley publient dans le journal de physiologie le modèle qui porte désormais leur nom et qui leur vaudra le prix Nobel de médecine en 1963. Ce modèle [Hodgkin and Huxley, 1952], qui s'appuie sur un formalisme assez proche de celui de Lapicque pour ce qui est de sa représentation (un circuit RC équivalent à la membrane) est capable de reproduire le mécanisme de la genèse des spikes en prenant en compte les propriétés de conductance dynamique de la membrane du neurone en fonction de sa valeur de potentiel. Initialement développé sur l'axone de calmar pour des raisons de facilité expérimentale, il a été par la suite étendu à de nombreux types cellulaires et notamment appliqué à la modélisation de neurones d'autres espèces animales et à des conditions physiopathologiques complexes [Masson et al., 2014]. Dans le cas qui nous intéresse ici le HHM pose d'emblée des problèmes méthodologiques. Très performant dès lors qu'il s'agit de l'étude détaillée de structures cellulaires et sub-cellulaires, ce modèle devient complexe à manipuler dans le cas de grands réseaux. En effet un modèle type HH mono-compartimental nécessite déjà un temps de calcul sensiblement plus long que celui qui est requis pour un LIF ou un neurone d'Izhikevich en raison du grand nombre de calculs à effectuer. Mais en passant à l'échelle du réseau, il faut souvent envisager d'ajouter d'autres compartiments (axone et dendrite notamment) ainsi que des modèles synaptiques qui soient suffisamment détaillés (de préférence avec un degré de description équivalent à celui envisagé pour les neurones). Tous ces éléments impliquent un espace des paramètres de très grande dimension ainsi que des temps de calcul prohibitifs pour les structures que nous étudions qui comportent de grandes populations neuronales en interaction. Il semble donc que cette approche ne soit pas adaptée au travail que nous poursuivons ici et aux ressources dont nous disposons.

Modèles synaptiques et axonaux Les synapses peuvent être classées comme des connexions chimiques ou électriques. Le cas, par ailleurs très intéressant des jonctions électriques ne sera pas abordé ici. Ces éléments jouent certainement un rôle dans le sujet qui nous intéresse mais comme nous n'avons pas d'accès suffisamment précis à ses propriétés et comme notre approche repose essentiellement sur la description des propriétés dynamiques d'un circuit neuronal « classique », nous avons fait le choix de nous limiter à l'étude des synapses chimiques.

Comme cela a été évoqué dans le paragraphe 1.4.1.2 le degré de description des modèles synaptiques doit être cohérent avec celui des modèles neuronaux. Il apparaîtrait assez curieux en effet de décrire des neurones avec un soin extrême pour ensuite étudier les propriétés du réseau en utilisant des modèles synaptiques grossiers. Il nous faut donc envisager une certaine

homogénéité entre les modèles de synapses et de neurones.

L'autre élément à prendre en considération est celui du nombre. Nous disposons en effet de ressources de calcul limitées à la fois en puissance de calcul et en mémoire disponible. Or il est bien connu que, si une grande population neuronale peut être longue à simuler, la prise en compte des synapses rend la modélisation informatique parfois tout simplement inenvisageable, à la fois pour des raisons de complexité dans la connectivité [SandraDiaz-Pier et al., 2016] mais aussi en raison de la quantité d'éléments de connexion à prendre en compte dans les calculs et qui a tendance à croître de façon exponentielle avec le nombre de neurones impliqués. Nous allons donc devoir opter pour un modèle synaptique validé mais qui puisse être utilisé sur des réseaux de grande taille afin de pouvoir mener dans un temps raisonnable les nombreuses simulations, nécessaires à la mise au point des paramètres du modèle.

Les modèles computationnels fondamentaux de synapses chimiques, sauf s'ils nécessitent une description extrêmement détaillée afin de pouvoir élucider leurs mécanismes internes [Nakano et al., 2010], reposent souvent sur les éléments suivants [Schutter, 2009] :

- à l'instant de l'arrivée du spike pré-synaptique une quantité donnée de neuromédiateur est libérée dans la fente synaptique ;
- le courant post-synaptique (pour les synapses ionotropiques) est proportionnel à la quantité de neuromédiateur disponible et provoque une variation dans le potentiel de la membrane post-synaptique selon le sens du courant apporté ;
- la quantité de neuromédiateur libérée disparaît progressivement (catabolisme, recapture ...) selon une décroissance exponentielle.

Ce principe est très souvent utilisé et les synapses pondérées à décroissance exponentielle sont utilisées dans des modèles de grands réseaux avec de très bons résultats. Comme nous le verrons ultérieurement notre modèle comporte néanmoins un très grand nombre de synapses et comme leur effet sur la membrane post-synaptique peut également être modélisé avec un simple saut de potentiel nous emploierons cette dernière approche qui constitue une combinaison adaptée et performante [Brette et al., 2007].

Les axones constituent le dernier élément à prendre en compte dans un modèle de réseau. En effet les délais de transmission des spikes le long des axones, même s'ils sont souvent brefs, se surajoutent au délai synaptique proprement dit et influencent sensiblement les instants des spikes post-synaptiques [Brette et al., 2007]. Comme ces délais semblent jouer un rôle dans la question qui nous intéresse, c'est-à-dire les mécanismes neuronaux de la prise de décision [Humphries et al., 2006] et que par ailleurs, le temps est explicitement pris en compte dans les comportements que nous modélisons, il en sera tenu compte dans nos simulations.

1.4.1.3 Outils de simulation

Il existe plusieurs outils dédiés à la simulation des SNN. Une revue accompagnée de nombreux tests a été publiée en 2007 [Brette et al., 2007] et depuis de nouveaux outils ont encore fait leur apparition. Nous avons donc dans un premier temps recensé les outils qui, en 2013, nous permettait de construire notre modèle de BG. Idéalement, il fallait un outil permettant de gérer des réseaux de grande taille de façon efficace, décrivant explicitement l'activité dans le temps de chacun de ses éléments (neurones, axones et synapses). Le site du CCNLab de l'Université du

Colorado propose une recension récente des logiciels de simulation dédiés aux SNN ¹⁴. Après avoir écarté certains de ces logiciels car ils ne correspondaient pas complètement à ce dont nous avons besoin, nous nous sommes focalisés sur NEURON [Carnevale and Hines, 2009] et Brian [Goodman and Brette, 2009]. Ces deux logiciels possèdent en effet a priori tous les éléments nécessaires. De plus ils sont couramment utilisés au laboratoire d'accueil à des fins pédagogiques et de recherche. Enfin, il est toujours préférable de ré-utiliser un logiciel déjà validé pour des raisons de reproductibilité des résultats et aussi pour éviter d'avoir à ré-implémenter un code source complet. Mais malgré la grande souplesse d'utilisation de ces outils, nous nous sommes dans les deux cas heurtés à des difficultés techniques qui nous ont semblé rédhibitoires.

NEURON est effectivement capable de gérer des réseaux de grande taille avec beaucoup d'efficacité [Migliore et al., 2006, Hines et al., 2008]. Mais de par sa conception originelle, il reste spécialisé dans la simulation de neurones complexes multi-compartimentaux et l'utilisation des modèles de SNN (ArtificialCell dans NEURON) est lourde et difficile à mettre en place. D'une part, l'interface graphique utilisateur (GUI) ne nous aurait pas permis de développer notre modèle de réseau en raison de sa taille et de sa complexité. D'autre part, le langage de script initial (hoc) se prête mal à la conception et la manipulation de ce type de réseaux, malgré l'existence dans la librairie de base de trois types de neurones proches du LIF. Il semble que NEURON soit destiné à terme à utiliser Python comme langage de script mais quand nous avons entrepris cette étude, cette portabilité était très partielle et difficile à utiliser [Hines et al., 2009]. Après avoir implémenté par nous-même des classes adéquates à l'aide de NMODL nous avons finalement renoncé car le temps de mise au point prévisible était trop important pour des résultats incertains.

Brian nous a donc semblé la meilleure alternative. C'est une librairie Python permettant la conception et la simulation de réseaux de grande taille ¹⁵ avec des outils de calcul optimisés qui permettent, tout en ayant la facilité d'implémentation et la grande richesse de Python, de simuler rapidement des réseaux dont les modèles neuronaux et synaptiques peuvent intégralement être paramétrés. Nous avons donc testé des prototypes de notre modèle avec cette librairie. A l'usage nous avons eu des difficultés à bien contrôler les éléments individuellement (synapses et neurones) et notamment les délais de propagation individuels (axones) variables. Le code permettant de générer le modèle atteignait 500 lignes de Python auxquelles il fallait ajouter 500 lignes de code pour les visualisations diverses. De plus, le temps mis pour 1s de simulation (avec un pas de temps de simulation de 0.1ms) était proche de 400s (contre 30s pour notre code). Ces contre-performances auraient sans doute pu être améliorées voire évitées avec un code Brian optimisé mais nous n'avons pas été en mesure de le réaliser sans devoir renoncer à certains aspects de notre modèle qui nous semblaient importants. De notre expérience certes restreinte, il semble que Brian s'avère extrêmement performant sur de grandes populations homogènes mais que ces avantages se trouvent amoindris quand il faut gérer de nombreuses sous-populations hétérogènes en parallèle.

Nous avons également envisagé XPP [Ermentrout, 2002] (ou XPPAUT) ¹⁶ mais celui-ci nous a semblé d'un abord obscur et ne semblait du reste pas se prêter à des simulations contrôlées

14. https://grey.colorado.edu/emergent/index.php/Comparison_of_Neural_Network_Simulators

15. <http://briansimulator.org/>

16. <http://www.math.pitt.edu/~bard/xpp/xpp.html>

en boucle fermée comme nous l'envisagions. Enfin, nous avons testé NEST (et Pynest son implémentation Python [Eppler et al., 2009]) qui était sans doute, Brian mis à part, le logiciel le plus adapté à notre problématique mais nous avons été dérouté par la syntaxe initiale du langage SLI¹⁷ et voyant mal comment l'intégrer dans notre travail il nous a fallu sur le moment laisser de côté cette initiative pourtant prometteuse.

1.4.2 Antériorité des modèles computationnels des BG

Dans le but de mieux comprendre les mécanismes et fonctions des BG ainsi que leurs implications dans certaines maladies, plusieurs modélisations ont été proposées au fil des années, à différentes échelles d'intégration. Les modèles adaptatifs des BG sont donc très nombreux et une revue exhaustive de chacun d'entre eux serait fastidieuse et pas nécessairement pertinente. Nous avons donc choisi de présenter chronologiquement les modèles qui à notre sens constituaient des étapes importantes dans ce domaine et plus particulièrement ceux dont la filiation nous amenait logiquement à développer l'approche que nous avons suivie dans cette étude.

1.4.2.1 Modèle d'Albin

Les représentations en boîtes et flèches (« *Box-and-Arrows* ») sont les précurseurs des modèles computationnels modernes. Leur principe est simple : les structures sont représentées par des boîtes et leurs projections neuronales par des flèches. Les représentations schématiques peuvent varier en fonction de l'information à transmettre : par exemple un code couleur pour les boîtes et/ou les flèches peut signifier un neurotransmetteur différent, ou encore des extrémités de flèches différentes peuvent signifier des effets de modulation différents (excitateur/inhibiteur). Ce type de représentation en boîtes et flèches permet donc de simplifier la visualisation de la région du CNS étudiée, d'en formaliser le graphe et globalement d'aller plus loin qu'une description purement anatomique.

Le modèle d'Albin-DeLong [Albin et al., 1989, DeLong, 1990] est l'un de ces modèles en boîtes et flèches (Figure 1.21). Son objectif était de mieux comprendre et d'expliquer les symptômes des maladies de Parkinson et de Huntington au travers des activités neuronales dans les BG et dans les structures associées. Publié il y a plus de 25 ans, ce modèle est l'un des premiers modèles modernes des BG. En combinant les connaissances de l'époque à ce sujet, les auteurs ont pu faire émerger une vue unifiée et cohérente des interactions entre les différentes régions considérées. Sur un même graphe il devenait possible d'expliquer simplement et d'illustrer de façon logique la physiopathologie des maladies hypokinétiques et hyperkinétiques.

Le concept de base de ce modèle repose sur le contrôle de l'excitation et de l'inhibition corticale par les BG, au travers de deux voies distinctes : la voie directe et la voie indirecte déjà évoquées dans la section 1.1.2.2. En effet, la première (Striatum D1 → GPi/SNr) faciliterait l'activité motrice corticale alors que la deuxième (Striatum D2 → GPe → STN → GPi/SNr) inhiberait la décharge cortico-motrice. Dans ce modèle, ces deux voies agiraient au travers de trois neurotransmetteurs : le glutamate pour un effet excitateur, le GABA pour un effet inhibiteur et la DA pour un effet modulateur (au niveau des MSN).

17. Il s'agit ici du langage de script de Nest

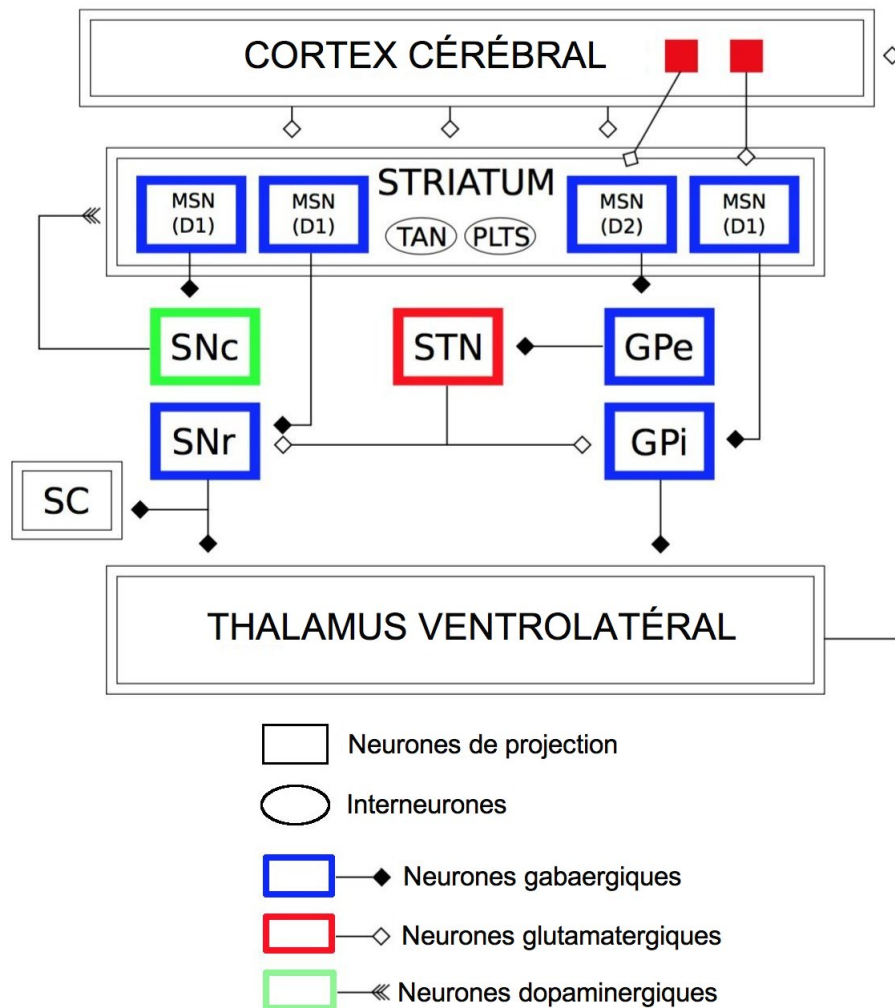


FIGURE 1.21 – Schéma de la connectivité du modèle des BG de Albin à l'état normal, d'après [Liénard, 2013]. Pour plus de lisibilité, les abréviations des structures originales ont été remplacées par celles communément utilisées de nos jours. Les voies directe (Striatum D1 → GPi/SNr) et indirecte (Striatum D2 → GPe → STN → GPi/SNr) sont incluses mais non explicitement représentées. Les entrées du striatum se font par le cortex (projections excitatrices glutamatériques) et la SNc (projections modulatrices dopaminergiques). Notons la présence de synapses dopaminergiques au niveau des interneurones striataux cholinergiques (ACh) fournissant une entrée excitatrice aux neurones de projection striataux. (SC : colliculus supérieur, « Superior Colliculus », PLTS : Persistent Low-Threshold Spiking.)

Dans le cas de sujets sains, une excitation des neurones épineux de la voie directe (dMSN) désinhiberait leurs cibles thalamo-corticales pour initier l'action voulue alors qu'une excitation des neurones épineux de la voie directe indirecte (iMSN) inhiberait ces mêmes cibles pour réprimer les actions non désirées. Selon ce modèle, la sur-activité des neurones de sortie des BG (provoquée par une sous-activation des neurones épineux de la voie directe) inhiberait l'activité thalamo-corticale et entraînerait des manifestations cliniques hypokinétiques observées dans la maladie de Parkinson.

En partant de ce modèle, les auteurs ont ensuite été capables de représenter de façon cohérente d'autres pathologies touchant les BG comme l'hémiballisme ou encore les stades précoces de la maladie de Huntington (voir la section 1.1.5 pour plus de détails). Avec ce modèle, ils pouvaient présenter la relation entre les symptômes moteurs et les dysfonctionnements des activités neuronales afin de mieux comprendre les causes et manifestations de ces maladies.

1.4.2.2 Modèle de Gurney (2001)

Ce modèle est communément appelé modèle GPR1/2 en références à ses créateurs : K. Gurney, T. J. Prescott et P. Redgrave. Il propose une anatomie fonctionnelle des BG novatrice à la fois sous la forme d'un modèle (GPR1) de boîtes et de flèches [Gurney et al., 2001a], mais également avec une analyse mathématique sous la forme d'une implémentation computationnelle (GPR2) [Gurney et al., 2001b].

Ce modèle pose l'hypothèse que les BG participent à la sélection des réponses motrices. En effet, il suit l'hypothèse de [Mink, 1996] suivant laquelle c'est la voie directe qui facilite spécifiquement les réponses désirées plutôt que l'activité motrice en soi. Selon ce modèle, cette facilitation motrice spécifique serait possible car les BG consistent en (i) un système de sélection (structures des voies directe et hyperdirecte) couplé à (ii) un système de contrôle (structures de la voie indirecte). Ce système de contrôle (Figure 1.22) n'aurait pas de fonction propre mais supporterait (i) dans le choix des réponses. Ce contrôle s'exercerait par la régulation de l'activité du STN et, par extension, de celle du GPi. En effet, dans ce modèle, la sélection motrice aurait besoin que l'intensité globale de l'inhibition motrice ne soit ni trop forte pour ne pas annuler toute facilitation d'un programme moteur spécifique via le striatum, ni trop faible pour que des réponses différentes ne soit pas sélectionnées simultanément. En régulant l'activité du STN ainsi que la quantité d'inhibition motrice, le système de contrôle assurerait donc un équilibre adéquat entre excitation et inhibition de façon à produire la réponse la plus adaptée à un moment donné.

Le modèle GPR1 fournit donc une interprétation différente de l'architecture computationnelle des BG, notamment en ce qui concerne le rôle du GPe. En effet, dans ce modèle les auteurs ont délégué la sélection de l'action à la voie directe et sa dé-sélection à la voie indirecte du cortex au STN. Ces deux voies ont donc été conçues autour du GPe, formant une boucle de commande qui améliore la sélectivité.

1.4.2.3 Modèle de Leblois (2006)

Dans cette étude [Leblois et al., 2006a] réalisent un modèle des BG pour expliquer la sélection de l'action et dans lequel deux circuits interagissent au niveau des connexions entre le STN-GPi. Le modèle montre également l'implication de la DA dans la potentiation de la transmission

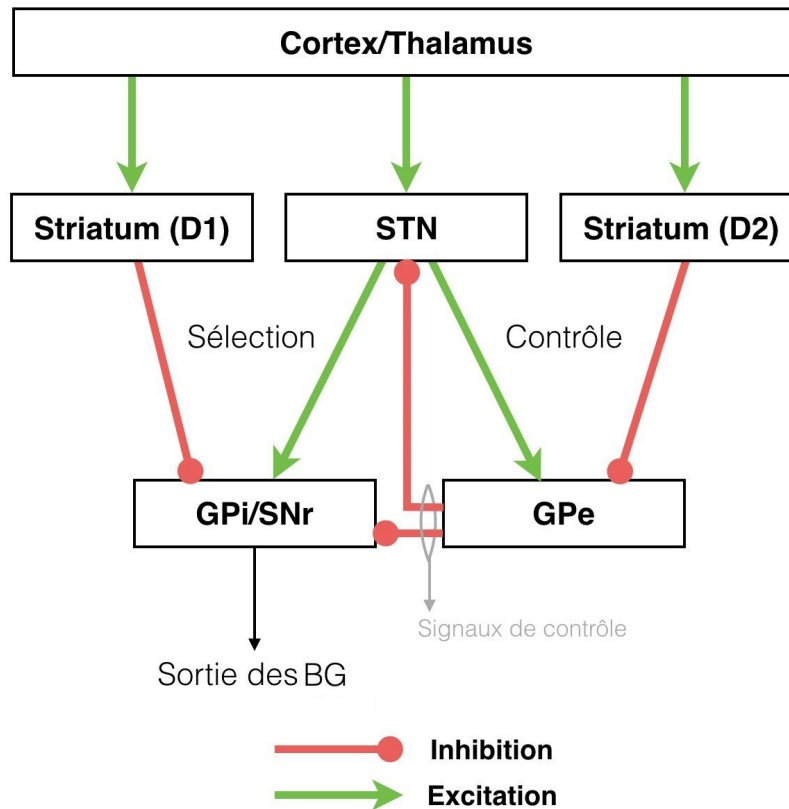


FIGURE 1.22 – *Adaptation du schéma de la connectivité du modèle des BG d’après [Gurney et al., 2001a]. Le système de sélection est constitué de la voie directe et de la voie hyperdirecte. Cette dernière inhiberait globalement tous les programmes moteurs, mêmes ceux étant désirés, inhibition qui se distinguerait uniquement par son activation supplémentaire de la voie directe. Le système de contrôle (voie indirecte) régulerait le nombre de représentations corticales pouvant être activées simultanément par un processus similaire.*

cortico-striatale. En effet, il développe l’étude de l’interaction entre deux canaux compétitifs et montre qu’un manque de DA empêche la mise en place du processus de sélection. Le modèle prédit également que les oscillations anormales observées dans la maladie de Parkinson sont conduites au travers de la voie hyperdirecte et ne se produisent que quand il n’y a plus d’inhibition striatale.

Cette étude se focalise sur les implications des voies directe et hyperdirecte dans la physiologie (à l’état normal et pathologique) des BG. Notons que le GPe ainsi que la voie indirecte ont été volontairement écartés en raison de leur rôles secondaires dans l’émergence des oscillations pathologiques du GPi ainsi que dans la détérioration des mouvements. Ce modèle ne nécessite donc pas la prise en compte complète de la voie indirecte pour fonctionner et avoir des capacités de sélection. Cet élément sera pris en compte dans notre étude.

Comme pour le modèle GPR1, chacune des structures des BG décrites représente une population de neurones et dans ce cas, chacune est séparée en deux canaux (sous-populations) en compétition (Figure 1.23). Ces deux circuits vont interagir au niveau des connexions entre le STN et le GPi pour contrôler les programmes moteurs, exploitant le fait que la connectivité entre ces deux structures est plus divergente que celle qui est observée au sein de la voie directe. Le processus de sélection s’effectue au travers d’une rupture de la symétrie dans l’activité de la boucle cortico-baso-thalamique.

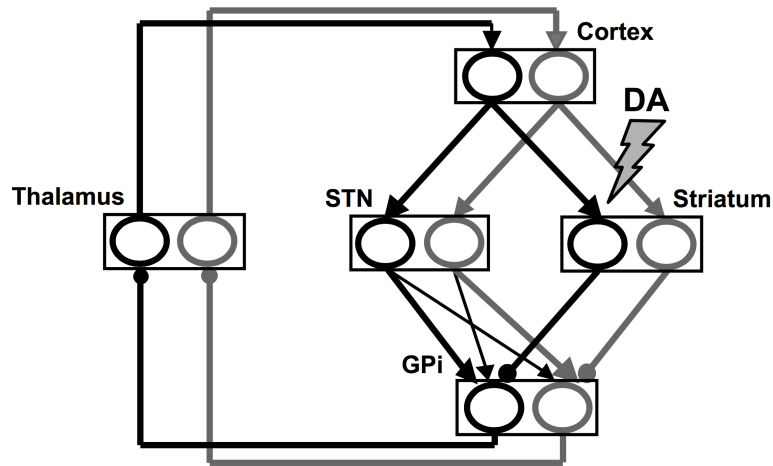


FIGURE 1.23 – Architecture du modèle des BG de Leblois [Leblois et al., 2006a]. Ce réseau est composé de deux circuits, chacun comprenant une population corticale, striatale, thalamique, sous-thalamique et pallidale. Ces deux circuits interagissent via des connexions sous-thalamo-pallidales diffuses. Les flèches en pointe représentent les connexions excitatrices et les flèches rondes les connexions inhibitrices.

1.4.2.4 Modèle de Humphries (2006)

Ce modèle [Humphries et al., 2006] est une amélioration du modèle de GPR1 cité précédemment [Gurney et al., 2001b]. En utilisant ici un modèle de neurones impulsionnels (« *spiking model* ») pour modéliser les BG, les auteurs ont fourni une analyse approfondie des capacités de sélection et des processus oscillatoires, tout en prenant en compte les effets de la DA dans le STN et GPe (boucle sous-thalamico-pallidale), les délais de transmission entre les neurones et les distributions spécifiques des entrées synaptiques sur les dendrites ainsi que d'autres paramètres déterminés au travers d'études expérimentales.

Dans ce modèle (Figure 1.24), l'entrée corticale glutamatergique excitatrice va projeter les cellules vers les deux noyaux d'entrée des BG, le striatum et le STN. Les neurones de projection striataux gabaergiques peuvent être subdivisés en deux populations suivant le type de récepteur dominant à la DA (type D1 ou D2) modulant leur décharge.

Les cellules à dominante D1 vont de préférence envoyer des projections inhibitrices vers les noyaux de sortie, la SNr et le GPi (le modèle présenté ne comprend que la SNr car les deux noyaux sont similaires concernant leurs connexions intra-BG). En retour, ces cellules gabaergiques vont envoyer des sorties inhibitrices toniques vers leurs cibles dans le thalamus et le tronc cérébral. C'est la suppression de cette inhibition tonique qui est supposée signaler la sélection d'une action : la combinaison des cellules striatales D1 et des noyaux de sortie est donc appelée la voie de sélection. L'encadré sur la droite de la Figure 1.24 montre le réseau formé par cette voie, avec à chaque fois une activité hypothétique des populations neuronales des canaux.

Les cellules à dominante D2 du striatum envoient des projections inhibitrices vers les cellules gabaergiques du GP. En retour, ces dernières envoient des projections inhibitrices vers le STN et la SNr/EP (EP : noyau entopédonculaire) et sont une source signal de contrôle pour les BG (ce circuit est donc appelé la voie de contrôle). Les cellules glutamatergiques du STN envoient des projections excitatrices vers le GP et la SNr/GP, en contribuant à la fois à l'activité tonique de la SNr/EP et à l'activité au travers de la voie de contrôle

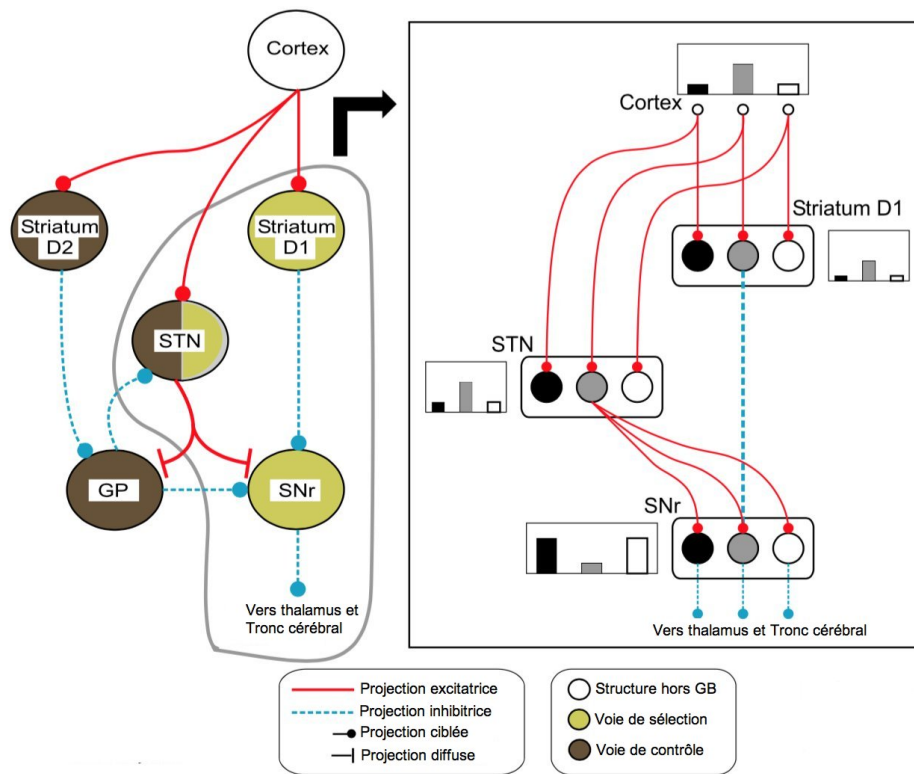


FIGURE 1.24 – Anatomie fonctionnelle des BG du modèle d’Humphries [Humphries et al., 2006]. Tous les composants de l’anatomie fonctionnelle des BG illustrés ici représentent une population neuronale du modèle (5 populations au total). Chacune sera divisée en trois canaux de façon à former trois sous-populations par structure. Le nombre de neurones par sous-population a été fixé à 64, ce qui fait 192 neurones par population.

Ce modèle utilise un formalisme décrit à l’échelle du neurone et a été testé sur ses capacités de sélection de l’action. Il fournit donc l’une des descriptions les plus précises des BG à ce jour. En plus d’expliquer les variations oscillatoires résultantes de l’activité de la boucle STN-GPe, il est capable de réaliser un processus de sélection de l’action compatible avec celui du modèle GPR1 fondé sur la population. Il a été prouvé que cette boucle de rétroaction STN-GPe pouvait être « découplée » fonctionnellement par une hausse tonique du niveau de DA, alors qu’une baisse de ce niveau les « recouplerait ». Ce travail est de plus cohérent avec une partie des résultats expérimentaux obtenus *in vivo* (enregistrements électrophysiologiques) aussi bien en condition physiologique que dans les cas de maladie de Parkinson.

1.4.2.5 Modèle de Guthrie (2013)

Le modèle computationnel des BG de [Guthrie et al., 2013] se situe dans la filiation des précédents travaux de [Leblois et al., 2006a]. Il s’en distingue par sa complexité accrue et notamment par sa capacité de simuler une sélection de l’action selon plusieurs modalités. Comme son prédécesseur, il simule notamment les interactions entre les voies directe et hyperdirecte et permet de mettre en évidence leurs rôles respectifs dans la prise de décision et la sélection de l’action. Il est construit à l’aide d’un formalisme de population (« *rate model* ») afin de rester ciblé sur l’étude de la dynamique du réseau. Le modèle de Guthrie est capable de reproduire efficacement la tâche d’apprentissage chez le singe évoquée plus haut ainsi que plusieurs de ses

caractéristiques dynamiques [Pasquereau et al., 2007].

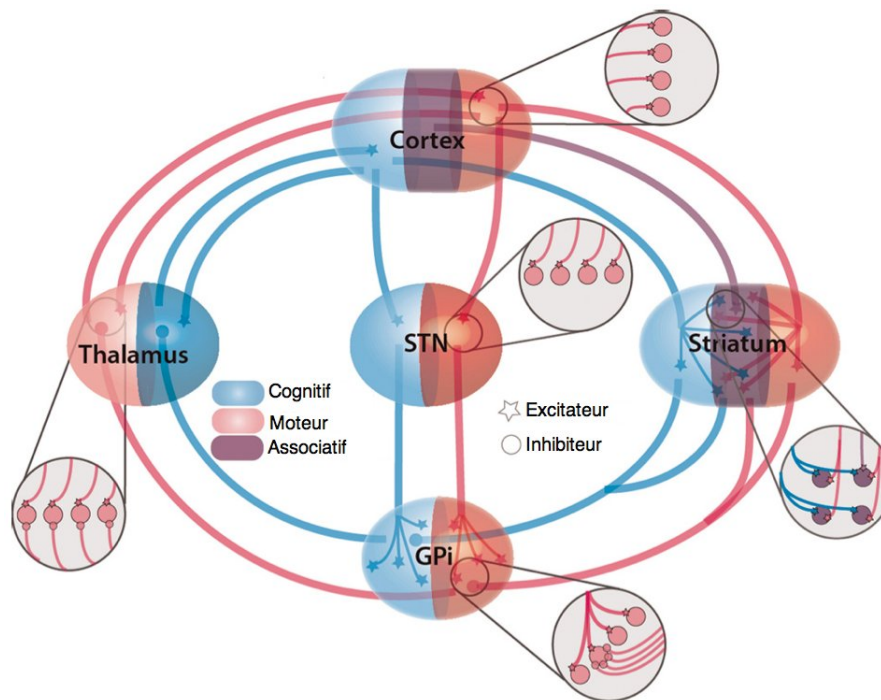


FIGURE 1.25 – Architecture du modèle des BG de Guthrie [Guthrie et al., 2013]. Deux boucles cortico-basales sont modélisées ici, représentant deux niveaux de prise de décision : une boucle cognitive en bleu et une boucle motrice en rouge. Chaque boucle comprend la voie directe de rétroaction positive et la voie hyperdirecte de rétroaction négative. Les expansions des entrées pour chaque structure sont montrées à l'intérieur des cercles. Dans le but d'éviter toute confusion, seules les entrées sont montrées ici, les sorties pouvant être déduites du schéma.

Ce modèle introduit un mécanisme de sélection de l'action fondé sur la compétition entre une rétroaction positive au travers de la voie directe et une rétroaction négative via la voie hyperdirecte. Il permet d'avoir une organisation parallèle des circuits entre les BG, le cortex et le thalamus, en utilisant des boucles séparées (Figure 1.25) : une boucle dite « cognitive » pour le choix entre les deux cibles présentées et une autre « motrice » pour la décision du mouvement final. En effet, pour que le modèle soit capable de réussir la tâche comportementale (voir partie 1.2.2.1 pour plus de détails), il convient tout d'abord de choisir la cible, et seulement après ce choix une décision de mouvement vers la cible sélectionnée est possible. Ce mécanisme de la sélection de l'action est fondé sur la rupture de la symétrie entre les différents canaux activés par la présentation des cibles en ajoutant du bruit aux structures (comme dans [Leblois et al., 2006a]) ce qui permet au système d'explorer les différentes actions possibles.

Enfin, ce modèle incorpore également un mécanisme d'apprentissage entre les cibles présentées et leurs conséquences (récompense ou non). Ce phénomène se produit au niveau cortico-striatal, dans la partie cognitive, en utilisant l'apprentissage par renforcement. Un signal récompense est utilisé afin de mimer les changements phasiques de la concentration de DA, nécessaires pour produire la LTP. Le fait qu'une récompense soit délivrée ou non dépend de la probabilité de récompense de la cible associée à la direction choisie.

Il faut noter qu'une fonction de transfert cortico-striatale sigmoïde est employée ici. Elle a pour objectif de mimer la nécessaire coordination des entrées corticales pouvant entraîner une

activité dans le striatum (notoirement silencieux au repos). Cet aspect du modèle de Guthrie a son importance et sera discutée ultérieurement quand nous présenterons notre approche en détail.

1.4.2.6 Travaux récents

Depuis 2013, nous avons relevé trois publications supplémentaires très ciblées qui nous ont guidés dans notre compréhension et dans la mise en perspective des processus de prise de décision dans les BG. L'étude de Chersi [Chersi et al., 2013] utilise un modèle de SNN avec des neurones de type LIF dont les paramètres (R_m et C_m) ont été fixés sur des bases d'études quantitatives précédentes. Le réseau ainsi constitué est de grande taille et comporte 14600 neurones distribués selon les différentes régions des BG, du thalamus et du cortex. Il est impliqué dans l'apprentissage et la réalisation d'une tâche comportementale dans laquelle un singe reçoit une récompense lorsqu'il choisit l'option adéquate. Cette récompense, modélisée par un pic phasique de DA, va modifier les synapses entre les neurones actifs durant l'essai suivant une règle de STDP et la présence de traces de sélection (« *eligibility traces* »). Le classique paradigme du « *Temporal-Difference learning* » (TD learning) y est donc remplacé par un simple principe mécanique de variation des poids ne faisant appel à aucun superviseur. Cette étude nous a été précieuse car nous y avons trouvé la confirmation du bien fondé de certains choix du modèle comme ceux cités précédemment (par exemple l'utilisation du modèle neuronal LIF). D'autre part les notions de trajets directs, indirects et hyperdirects sont rappelés mais mis de côté au bénéfice de données générales sur la connectivité entre différentes régions et de notre point de vue, cela permet un regard sinon neuf, du moins plus objectif sur la circuiterie si complexe des BG. D'une façon générale cette publication, malgré les limitations parfaitement décrites par les auteurs, a l'immense bénéfice de sortir de quelques impasses conceptuelles concernant l'anatomie fonctionnelle et la dynamique des BG lors des processus de prise de décision.

En 2015 Mandali [Mandali et al., 2015] a publié une étude sur ce même thème en employant un formalisme neuronal différent. A la place d'utiliser des neurones LIF les auteurs ont utilisé des SNN constitués de modèles d'Izhikevich et un paradigme d'apprentissage par renforcement. Cette étude se focalise notamment sur le rôle des connexions latérales dans le STN (et donc du trajet indirect même si le trajet hyperdirect n'est pas pris en compte) dans le comportement d'exploration (vs. exploitation). Ce choix prend tout son sens quand l'on sait que le STN est la cible privilégiée du traitement des syndromes parkinsoniens par la DBS. Bien que très focalisée dans son approche, cette étude donne des informations importantes concernant notamment l'équilibre exploration-exploitation qui est un élément clé des processus de prise de décision et dont une partie au moins se joue dans les BG.

Plus récemment enfin Berthet [Berthet et al., 2016] a publié une mise à jour de son étude de 2012 en utilisant lui aussi un modèle de SNN et un formalisme de neurones LIF. Le réseau de 725 neurones pour 70000 synapses (simulé sur un système de calcul CRAY XC30) a permis aux auteurs de montrer les contributions respectives des trajets directs et indirects sur les processus d'apprentissage avec récompense ainsi que sur des tests lésionnels et sur le syndrome parkinsonien. Dans ces dernières simulations ils ont notamment montré que la déplétion dopaminergique impliquant D2 jouait un rôle majeur par rapport à son effet sur D1. Un autre résultat remarquable de cette étude concerne la question controversée de l'inhibition latérale dans le striatum. Les

auteurs montrent grâce à leur modèle que celle-ci peut s'avérer nécessaire dans certaines situations et semble jouer un rôle graduel selon la complexité de la tâche conduite : « *We believe that the low number of states and actions along with the simple test setup might be particularly well-suited for the condition without lateral inhibition, but such a condition would fail to perform as well in more complex situations* ». Ce rôle nuancé (et controversé) de l'inhibition striatale latérale sera développé plus avant dans notre étude.

1.4.2.7 Synthèse et choix méthodologiques

Récapitulatif des modèles Tous les modèles passés en revue dans la partie précédente ont été utiles dans la compréhension des fonctions des BG, ainsi que dans l'élaboration du modèle dont il sera question dans la partie suivante. Dans le but de comparer les apports de ces différents modèles, la Table 1.2 propose un résumé des objectifs et paramètres d'intérêts de ceux-ci.

TABLE 1.2 – Résumé des modèles computationnels des BG. Un modèle de type « rate model » signifie qu’il prend en compte l’activité fréquentielle d’une population neuronale par opposition aux modèles de type « spiking neuron model » dans lesquels chaque potentiel d’action de chaque cellule est pris en compte. Le symbole « O » signifie que le protocole comprend au moins une boucle ouverte et le symbole « F » signifie au moins une boucle fermée (cf. 1.2.1).

Modèles		Albin et al. (1989)	Gurney et al. (2001)	Leblois et al. (2006)	Humphries et al. (2006)	Guthrie et al. (2013)	Chersi et al. (2013)	Mandali et al. (2015)	Berthet et al. (2016)
Parties des BG représentées	Striatum	Oui	Oui	Oui	Oui	Oui	Oui	Oui	X
	GPe	Oui	Oui	Non	Oui	Non	Oui	Oui	X
	GPI	Oui	Oui	Oui	Oui	Oui	Oui (SNr)	Oui	X
	STN	Oui	Oui	Oui	Oui	Oui	Oui	Oui	X
Type de modèle		Boîtes/Flèches	”Rate”	”Rate”	”Spiking”	”Rate”	”Spiking”	”Spiking”	”Spiking”
Fonctions comprises dans le modèle	Motrices	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui
	Sélection	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui
	Plasticité	Non	Non	Oui	Non	Oui	Oui	Oui	Oui
Boucles		/	O	F	O	F	F	F	F
Voies	Directe	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui
	Indirecte	Oui	Oui	Non	Non	Non	Oui	Oui	Oui
	Hyperdirecte	Non	Oui	Oui	Oui	Oui	Oui	Non	Non

Conclusion Comme nous venons de le voir, la modélisation de l'apprentissage et de la prise de décision dans les BG est une tâche complexe et peut être abordée à différents niveaux d'intégration. L'objectif de cette thèse est de mieux comprendre comment les BG en interaction avec le cortex et le thalamus, permettent l'apprentissage de processus de prise de décision et de sélection de l'action capables de reproduire le comportement de modèles animaux complexes. Ce modèle doit être également capable de réaliser avec succès différentes tâches comportementales sans adaptation et ce afin de s'assurer de sa généralité. Dans le cas présenté ici, les tâches comportementales initiales sont réalisées sur des singes (voir sections 1.2.2.1 et 1.2.2.2). Ce modèle doit enfin être capable dans une certaine mesure de simuler les effets des lésions localisées de manière cohérente en accord avec les résultats expérimentaux.

Afin de limiter le degré d'abstraction du modèle, le choix est fait dans cette étude de décrire dynamiquement le fonctionnement de ses structures au niveau du neurone et de la synapse. Il s'agit donc pour nous de trouver un compromis qui permette, au vu de la puissance de calcul dont nous disposons, de pouvoir faire fonctionner avec des délais raisonnables, un modèle de réseau détaillé mais capable de reproduire des protocoles de prise de décision en tenant compte des principales structures impliquées. Ce choix répond selon nous à deux besoins :

- la possibilité de faire émerger les fonctionnalités du réseau par une approche « *bottom-up* » et ainsi de mettre en évidence la relation structure-fonction dans la circuiterie neuronale étudiée.

- la possibilité à terme de tester des hypothèses mettant en jeu des mécanismes à l'échelle cellulaire sur la dynamique d'ensemble du réseau. Il peut s'agir par exemple du rôle des propriétés cellulaires intrinsèques, de l'action d'agents pharmacologiques sur certains récepteurs, de l'effet de la DBS sur une région donnée ou encore de l'effet de la plasticité synaptique dépendant du temps.

Les études expérimentales sur lesquelles nous nous sommes appuyés implique des primates. Bien que ce modèle soit le plus proche de l'espèce humaine et donc à ce titre le plus à même de donner des résultats transposables aussi bien sur le plan fondamental que sur celui des retombées thérapeutiques, il pose de nombreux problèmes à la fois sur le plan pratique et éthique. Pour ces deux raisons, une partie de notre travail de recherche a visé à développer un modèle comportemental de prise de décision sur un modèle animal moins controversé : l'urodèle¹⁸. Ce choix nous a semblé d'autant plus pertinent qu'il existe désormais un modèle Parkinsonien de ces animaux comme nous l'avons mentionné auparavant (cf. 1.2.2.3). Ce travail de développement qui a fait l'objet d'une collaboration avec une équipe de l'Institut Magendie possédant une longue expertise dans ce sujet sera décrit plus en détails dans les parties suivantes.

18. Ou pleurodèle de Waltl (*Pleurodeles waltlii*), appelé plus communément « salamandre ». Les deux termes seront utilisés dans ce manuscrit.

Développement et étude d'un modèle connexionniste des BG : application à la simulation d'une tâche d'apprentissage dans un protocole de prise de décision

*« It's not about making the right choice.
It's about making a choice and making it
right. »*

J. R. Rim

Sommaire

2.1	Introduction	55
2.2	Matériel & Méthodes	56
2.2.1	Tâches comportementales étudiées	56
2.2.2	Développement des outils de simulation et d'analyse	57
2.2.3	Implémentation du circuit	57
2.2.4	Implémentation de la plasticité synaptique	63
2.2.5	Développement et mise au point du réseau	64
2.3	Résultats	70
2.3.1	Étude de sensibilité	70
2.3.2	Tests léSIONNELS	71
2.3.3	Validation dans le protocole initial (Pasquereau et al., 2007)	73
2.3.4	Validation dans le protocole complémentaire (Piron et al., 2016)	81
2.4	Discussion	83
2.4.1	Apports et prédictions du modèle	83
2.4.2	Limites et perspectives de l'approche	89

FONDÉE sur l'hypothèse d'Alan Turing, l'approche computationnelle des neurosciences suppose que tout processus cognitif est le résultat d'un calcul. Les modèles développés en neurosciences computationnelles cherchent donc en général à reproduire par des simulations informatiques les types de calculs effectués par le système nerveux dans l'accomplissement de ses fonctions ¹.

2.1 Introduction

La prise de décision est un processus permettant à un organisme de choisir ses actions afin de maximiser les bénéfices résultants [Lee et al., 2012]. Ce processus se déroule au sein d'un réseau compliqué constitué de structures corticales et sous-corticales, parmi lesquelles les BG et le thalamus jouent un rôle majeur. Historiquement, les BG sont connus pour leur implication dans la sélection des actions motrices, mais ils sont désormais également connus pour leur action dans les processus limbique et cognitif [Bar-Gad et al., 2003]. Les structures des BG ont donc un rôle important dans le traitement de l'information au cœur du système nerveux. Cependant, même si leurs motifs de connectivité ont été décrits en grande partie [Gurney et al., 2001a, Hazrati et al., 1995], leur complexité est telle qu'il est difficile de les analyser formellement.

Beaucoup de modèles et d'hypothèses fonctionnelles ont été développés depuis que la description originale de la connectivité des BG a été développée il y a plus de 25 ans à l'aide d'un formalisme en boîtes et flèches (« box and arrows ») [Albin et al., 1989, Alexander et al., 1986]. Ces modèles ont tenté de comprendre les interactions dynamiques entre les différentes voies ainsi que leurs conséquences sur la fonctionnalité du système [Bar-Gad and Bergman, 2001, Mink, 1996, Redgrave et al., 1999, Suri and Schultz, 1999]. Dans ce type de formalisme, un compartiment représente à lui seul une sous-population neuronale décrivant une région donnée. Le transfert de l'information entre ces compartiments est représenté par des valeurs numériques scalaires d'activité. Ces méthodes présentent des vues mésoscopiques des échanges d'informations et reproduisent à cette échelle des dynamiques qui sont compatibles avec les résultats expérimentaux comportementaux et électrophysiologiques [Berthet et al., 2012, Frank et al., 2001, Leblois et al., 2006a, Sukumar et al., 2012]. Elles ont un avantage important qui est que leur implémentation informatique est relativement simple dans la mesure où, une seule équation différentielle permet à une population entière d'être simulée. De plus, l'espace des paramètres est en général assez restreint car ces derniers sont principalement utilisés pour contrôler les interactions entre les différents compartiments. Il est en revanche moins fréquent qu'une tentative soit faite pour décrire le fonctionnement global d'un circuit neuronal biologique à l'échelle microscopique ², particulièrement quand l'activité de celui-ci est étudiée également au niveau comportemental. Il est cependant nécessaire de tenter cette approche à un moment donné pour une meilleure compréhension des mécanismes complexes du système nerveux central. Par exemple, il peut être nécessaire de démontrer les relations entre un mécanisme physiopathologique ou pharmacologique décrit à l'échelle cellulaire et son impact global sur les propriétés du système nerveux et donc sur le fonctionnement et le comportement de l'organisme.

1. D'autres branches des neurosciences computationnelles appliquent la simulation informatique à l'étude de processus biophysiques à l'échelle cellulaire et sub-cellulaire mais ce n'est pas l'objectif poursuivi ici.

2. Dans ce cas, microscopique signifie « à l'échelle cellulaire ».

Il est donc clair que ces étapes ont besoin d'être étudiées en parallèle avec le développement de modèles à une échelle plus grossière. L'idée d'un modèle intégré doit donc selon nous être vue comme une garantie supplémentaire de validité et une condition pour l'amélioration des capacités prédictives des modèles computationnels.

Les difficultés inhérentes aux modèles de réseaux décrits à l'échelle neuronale sont, outre la nécessité d'une grande puissance de calcul, une explosion combinatoire du nombre de leurs paramètres et par voie de conséquence, la difficulté de leurs ajustements respectifs. Malgré ces obstacles, certaines études de réseaux neuronaux de petite taille et de taille moyenne ont montré leur faisabilité ainsi que la pertinence de leurs résultats [Sarvestani et al., 2013]. Il existe d'ailleurs plusieurs projets de simulation à grande échelle qui appliquent cette approche en vue de l'intégration et de la cohérence multi-échelles de modèles computationnels de régions entières du CNS [Hines et al., 2008, Migliore et al., 2006].

Dans cette étude, nous avons développé un modèle de réseau reposant sur trois sources. Premièrement les différentes voies et constituants du système formé par le cortex, les BG et le thalamus ont été instanciés sous la forme de populations neuronales interconnectées. À cette fin, nous avons réalisé une synthèse des données expérimentales de l'anatomie fonctionnelle et des connections somatotopiques communément admises [Mink, 1996, Nambu, 2011]. Ensuite, nous avons choisi des résultats expérimentaux comportementaux et électrophysiologiques obtenus sur des primates non-humains par notre équipe [Pasquereau et al., 2007, Piron et al., 2016] pour tester les caractéristiques du modèle. Enfin, nous nous sommes appuyés sur les apports de modèles computationnels antérieurs comme ceux décrits dans la section 1.4.2. Nous les avons étendus et nous avons comparé nos résultats avec les leurs quand cela semblait pertinent. Notre objectif final était, tout en restant cohérent avec les acquis précédents, de proposer une description dynamique du fonctionnement des BG à l'échelle microscopique.

2.2 Matériel & Méthodes

2.2.1 Tâches comportementales étudiées

La tâche principale a été précédemment décrite en détails dans la section 1.2.2.1. Les enregistrements électrophysiologiques et comportementaux des performances des singes réalisés en parallèle ont fourni une base de travail lors de la réalisation du modèle computationnel. Pour rappel, les singes étaient placés devant un écran sur lequel deux cibles parmi quatre possibles (chacune avec sa probabilité de récompense) étaient affichées à deux positions différentes (quatre fois deux positions possibles au total). Le but de la tâche était pour les singes d'apprendre quelle cible avait la meilleure probabilité de récompense afin de maximiser celle-ci. Dans ce contexte, la forme des cibles est en relation avec la partie cognitive des structures cérébrales impliquées, alors que leur position sur l'écran va impliquer la partie motrice (en rapport avec la direction du mouvement). Le modèle computationnel comprendra donc deux ensembles de boucles en parallèle pour la sélection de l'action.

2.2.2 Développement des outils de simulation et d'analyse

Comme nous l'avons mentionné précédemment dans la section 1.4.1.3 nous avons fait le choix de développer des outils de simulation dédiés pour notre étude. Pour la simulation le langage Java a été testé initialement puis après comparaison de la vitesse de calcul de nos algorithmes nous avons décidé de passer en C++14 et le modèle a donc été développé avec la librairie GNU C++ 4.9.

Le moteur de simulation est très simple et comprend 3 classes : une classe pour les neurones (LIF1DNrn), une classe pour les synapses (VJSyn) et une classe pour le réseau qui permet de les organiser et de les simuler de manière cohérente, de les stimuler et de les enregistrer (LIF1DNetwork). À chaque pas de temps, l'état de chaque objet neurone ou synapse contenu dans la classe réseau est mis à jour, ainsi que la position de chaque signal sur les axones. Le système est donc simulé en parallèle et il est possible d'en enregistrer chaque élément à chaque instant à l'aide de méthodes appropriées pour une analyse « *offline* ».

L'analyse et la visualisation des résultats de simulation ont été obtenus grâce à des scripts R [R-Team, 2015] et Python utilisant respectivement les librairies cowplot [Wilke, 2016], ggplot2 [Wickham, 2009] et matplotlib [Hunter, 2007]. Les analyses statistiques sur les résultats de nos simulations sont parfois réalisées avec des tests non paramétriques pour des raisons de distribution, d'hétéroscédasticité ou de taille d'échantillon. Dans ce cas nous nous sommes appuyés sur les paquetages R suivants : PMCMR [Pohlert, 2014] et pgirmess [Giraudoux, 2016] qui permettent de réaliser des comparaisons multiples à l'aide de tests non-paramétriques. Pour les tests non-paramétriques de comparaison de deux échantillons le test de Mann-Whitney a été appliqué dans le cas où ils étaient indépendants et le test de Wilcoxon dans le cas où ils étaient appariés. Les tests non-paramétriques de comparaison à une moyenne utilisent également un test de Wilcoxon. Les courbes d'activité tracées en fonction du temps (AFR) ont été lissées à l'aide d'un algorithme de convolution reposant sur un calcul de moyenne mobile de la bibliothèque scipy [Jones et al., 2001] et ce, afin d'améliorer leur lisibilité. L'intégralité des informations se rapportant à l'environnement de développement est disponible en annexe (cf. Table A.5) et le code du moteur de simulation et des outils d'analyse est disponible sur demande en attendant d'être mis en ligne sur un dépôt git.

2.2.3 Implémentation du circuit

2.2.3.1 Architecture du modèle

L'architecture du modèle computationnel que nous avons développé s'inspire en partie de celle du modèle de [Guthrie et al., 2013] qui est elle-même dérivée du modèle de Leblois [Leblois et al., 2006a] (cf. sections 1.4.2.5 et 1.4.2.3). Le modèle comprend les cinq structures cérébrales suivantes : le cortex, les BG (striatum, GPi, STN) et le thalamus (Figure 2.1).

Voies de connectivité Les boucles cortico-corticales comprennent trois voies principales, à savoir directe, indirecte et hyperdirecte. La voie directe (cortex → striatum → GPi → thalamus) exerce une rétroaction positive et excitatrice et est convergente et focalisée [Percheron et al., 1984]. La voie hyperdirecte (cortex → striatum → STN → GPi → thalamus) exerce au contraire

une rétroaction négative et inhibitrice et est plus diffuse à cause de la divergence des connexions partant du GPi et du STN. La voie indirecte se comporte plus comme un sous-réseau comprenant des boucles de rétroaction positive et négative au cœur même des BG. Cette complexité anatomique et son ambivalence fonctionnelle qui a été discutée dans la section 1.1.2.2 font que cette voie ne sera pas prise en compte dans l'architecture actuelle du modèle.

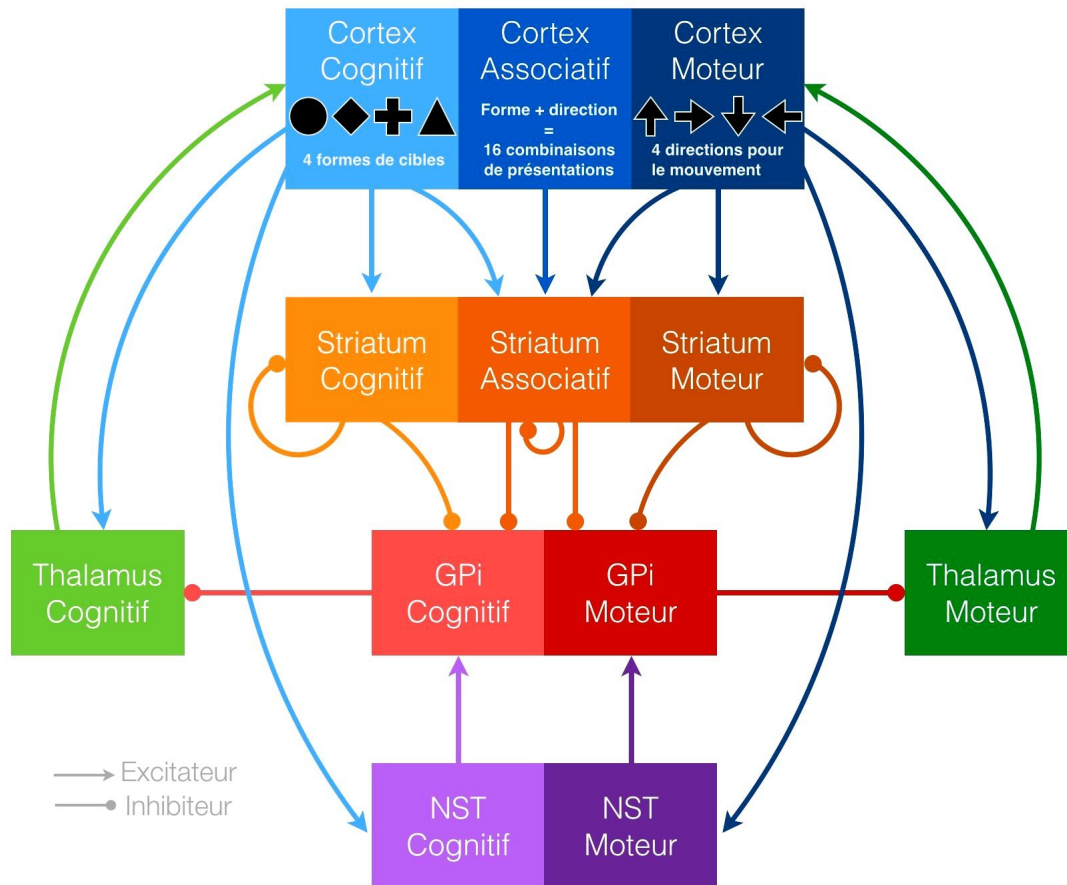


FIGURE 2.1 – Architecture du modèle computationnel des BG et structures associées, d'après les travaux de [Leblois et al., 2006a, Guthrie et al., 2013, Jason T. Moyer and Wolf, 2014]. Le cortex est en bleu, le striatum en orange, le GPi en rouge, le STN en violet et le thalamus en vert. Les dégradés de couleurs servent à illustrer les différentes parties de ces structures : la couleur la moins intense (sur la gauche) représente la boucle cognitive, la couleur la plus intense (sur la droite) représente la boucle motrice. Une partie associative a été ajoutée pour le cortex et le striatum, entre les parties motrices et cognitives. Les flèches « pointues » représentent les projections glutamatergiques excitatrices, alors que les flèches « arrondies » représentent les projections gabaergiques inhibitrices.

Inhibition striatale latérale Une des différences majeures apportée au modèle par rapport aux versions précédentes est la prise en compte de l'inhibition latérale existant au niveau du striatum. En effet, le striatum possède un certain nombre d'interconnexions inhibitrices dont le rôle joué dans son fonctionnement demeure controversé. Dans son étude de l'inhibition intra-striatale latérale, [Jason T. Moyer and Wolf, 2014] et ses collègues travaillent sur le rôle fonctionnel de l'inhibition latérale dans un modèle computationnel. Pour les auteurs, celle-ci améliorerait la synchronisation des décharges des MSN. Elle aurait effectivement tendance à améliorer la précision temporelle d'un ensemble de réponses de manière à coordonner les entrées

en supprimant l'activité des MSN qui ne seraient pas précisément synchronisés avec les autres. L'inhibition latérale supprimerait donc l'asynchronie dans la décharge des MSN.

L'instant précis de décharge des MSN modulé par l'inhibition latérale serait également important pour la régulation de la plasticité synaptique à long terme et surtout pour celle qui serait modulée par la DA. Dans ce cas il s'agirait de la plasticité dépendante du temps de décharge (« *Spike-Timing Dependent Plasticity* », STDP) [Lindskog et al., 2006]. La plasticité cortico-striatale serait en effet entraînée par un type de STDP dans laquelle la présence ou l'absence de DA déterminerait le renforcement ou l'affaiblissement d'une connection synaptique. Ce type de plasticité étant dépendant de la synchronie entre l'émission d'un spike par une cellule par rapport aux instants de ses signaux présynaptiques, l'inhibition latérale aiderait à définir précisément la fenêtre temporelle durant laquelle les MSN répondraient à l'entrée corticale, ce qui contribuerait à construire des circuits d'apprentissage plus efficacement [Jason T. Moyer and Wolf, 2014].

La prise en compte de l'inhibition latérale au niveau du striatum dans le modèle permettrait donc théoriquement de favoriser la population cellulaire la plus active lors du processus de prise de décision ainsi que lors de l'apprentissage.

Parties cognitives, motrices et associatives Chaque région cérébrale du modèle va comprendre au moins deux parties : une région cognitive (pour encoder la reconnaissance de la forme de la cible présentée) et une région motrice (pour encoder la sélection de la direction du mouvement). Le cortex et le striatum vont également inclure une région associative en plus afin de permettre les interactions des parties motrices et cognitives [Guthrie et al., 2013]. Le cortex associatif et ses projections vers le striatum représentent les régions corticales pariétales qui codent pour la combinaison de la forme et de la localisation de la cible. En d'autres termes, la partie associative sert à répondre à la question « Quelle cible est à quel endroit ? ».

2.2.3.2 Organisation du réseau

Pour cette partie, nous allons prendre l'exemple du cortex pour une meilleure compréhension de l'organisation détaillée du modèle. Le cortex est donc divisé en trois parties (cognitive, motrice et associative). La tâche comportementale comportant quatre formes de cibles possibles et quatre positions, il y aura quatre sous-régions cognitives (nommées C0, C1, C2 et C3, où le C signifie « cognitif ») et quatre sous-régions motrices (nommées M0, M1, M2 et M3, où le M signifie « moteur »). La partie associative quant à elle possédera 16 sous-régions (de A0 jusqu'à A15) afin de combiner toutes les formes avec toutes les positions possibles, comme illustré dans la Figure 2.2. Le cortex comportera donc 24 sous-populations neuronales, tout comme le striatum, alors que le GPi, STN et thalamus en auront 8.

Le nombre de neurones par sous-population prend en compte la taille approximative des structures anatomiques pour garder un ratio cohérent dans le modèle. Il est évident que le nombre de neurones dans le modèle n'est ni identique, ni exactement proportionnel à chaque structure mais il a fallu faire un compromis afin d'avoir un réseau d'une taille suffisante pour réaliser des interactions synaptiques complètes tout en restant utilisable sur un ordinateur de bureau standard. Il est bien connu [Schroll et al., 2015] qu'une analyse complète du cerveau à l'échelle cellulaire n'est pas encore possible à cause du nombre beaucoup trop important de

neurones et surtout de synapses. C'est pourquoi la plupart des auteurs utilisant un formalisme de ce type choisissent des tailles de population d'un ordre de grandeur raisonnable et décrivant en général un échantillonnage réaliste des sous-populations (bien que partiellement arbitraire) pour implémenter leur modèle (cf section 1.4.2.6). La Table 2.1 indique le nombre de neurones par sous-population dans notre modèle. La taille du réseau au total est de 3 680 neurones et l'ensemble de la connectivité entre les régions représente 723 200 synapses.

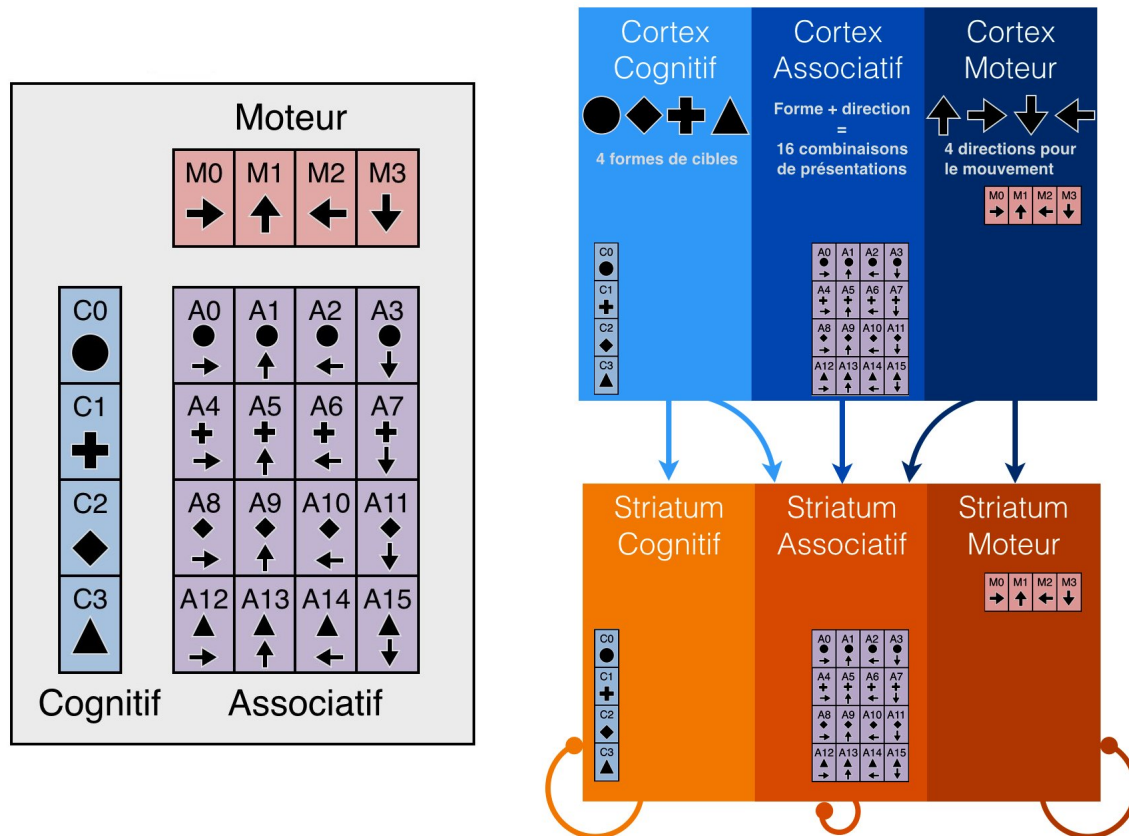


FIGURE 2.2 – Exemple de l'architecture détaillée du modèle pour le cortex et le striatum. La partie cognitive du cortex va comprendre quatre sous-régions (une pour chaque forme de cible), comme la partie motrice (une pour chaque direction) et la partie associative en comprendra seize (quatre formes combinées avec quatre directions). Les autres structures du modèle seront construites sur la même organisation pour les sous-régions motrices et cognitives, mais ne posséderont pas de sous-région associative. La connectivité complète du réseau sous forme de matrice est montrée dans la Figure A.1.

2.2.3.3 Paramètres du modèle

Pour cette étude, nous avons besoin d'un modèle neuronal suffisamment simple pour être implémenté dans un réseau à grande échelle, mais également suffisamment plausible pour ne pas avoir trop d'abstraction biologique. Comme nous l'avons vu précédemment (section 1.4.1.2), les modèles dérivés d'Hodgkin-Huxley par exemple demanderaient des capacités de calcul trop importantes [Hines et al., 2008] pour être utilisés dans le genre de réseau qui nous intéresse ici. Ils peuvent cependant fournir un outil intéressant dans l'exploration des propriétés fonctionnelles des circuits à une échelle plus réduite [Terman et al., 2002]. Afin de restreindre la complexité de calcul et de limiter le nombre de paramètres dans le modèle, nous avons choisi d'utiliser

TABLE 2.1 – Nombre de neurones et de synapses par sous-populations dans le modèle.

Région	Nombre de sous-populations	x	Nombre de neurones par sous-population	=	Sous-Total par région
Cortex	24		100		2 400
Striatum	24		40		960
STN	8		10		80
GPi	8		10		80
Thalamus	8		20		160
Total neurones			3 680		
Nombre de synapses			723 200		

le modèle neuronal issu des travaux de Lapicque dit « Intégrateur à Fuite » à une dimension (« *Leaky Integrate-and-Fire* », LIF), représenté par l'équation 2.1.

$$C \cdot \frac{dV_m(t)}{dt} + \frac{V_m(t) - V_{rest}}{R} = I \quad (2.1)$$

Dans ce modèle, la réinitialisation du potentiel d'action se fait avec l'équation 2.2.

$$\text{Si } V_m(t) > V_{thres} : V_m(t) \leftarrow V_{peak} \quad (2.2)$$

La variable $V_m(t)$ représente le potentiel membranaire et V_{rest} celle du potentiel de repos. Le temps est exprimé en s , les voltages en V ou en mV , et les courants en A . Pour les autres paramètres du modèle, les valeurs par défaut sont : $C = 10^{-10}F$ pour la capacité, $R = 10^8 ohms$ pour la résistance, $V_{peak} = 40mV$, $V_{rest} = -65mV$ pour le potentiel de repos et $V_{thres} = -55mV$ pour le seuil. Nous avons introduit dans notre implémentation un élément supplémentaire qui est que, suite à un potentiel d'action, le neurone se retrouve à une valeur de potentiel légèrement hyperpolarisée et non pas à sa valeur de potentiel de repos comme c'est le cas dans les LIF classiques. Cette possibilité qui a été conçue pour prendre en compte un élément factuel de la dynamique neuronale n'a pas été explicitement exploitée dans nos simulations et est sans incidence sur le fonctionnement et les conclusions du modèle. En effet dans ce cas c'est surtout la constante de temps membranaire ($\tau = R.C$) et la différence $V_{rest} - V_{thres}$ qui sont les principaux facteurs de modulation de l'émission des potentiels d'action.

Les synapses utilisées sont des synapses à saut de potentiel dont le poids augmente (ou diminue suivant les cas) sous le contrôle des règles de plasticité synaptique. Les valeurs des poids synaptiques sont exprimées en mV afin que leur effet (saut de potentiel) puisse être directement appliqué et sommé au niveau post-synaptique. La figure (Figure 2.3) montre un exemple de fonctionnement du modèle pour un réseau minimaliste de deux neurones et une connexion.

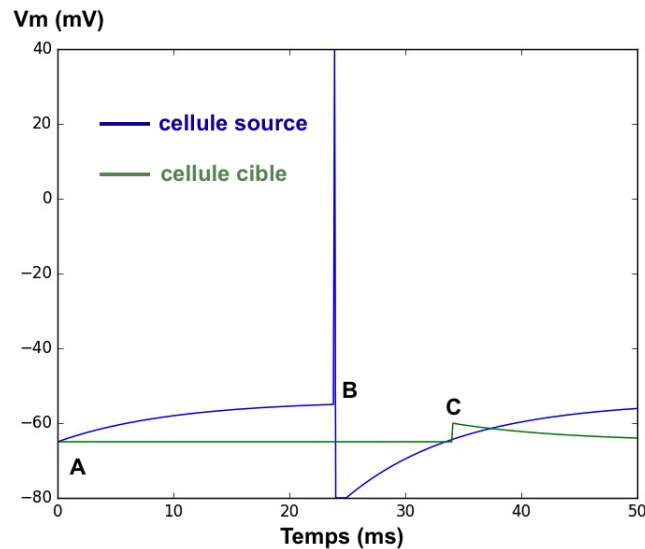


FIGURE 2.3 – Exemple de simulation de deux cellules interconnectées Une cellule source est connectée à une cellule cible via une synapse de poids 5 mV et de délai 10 ms. La cellule source est stimulée en continu avec un courant de 110 pA. (A) $t=0$, le potentiel de membrane (V_m) des 2 cellules est fixé à -65 mV. (B) un potentiel d'action est émis dans la cellule source quand le potentiel seuil est atteint. Le potentiel de membrane de la cellule source est remis à -80 mV (post-hyperpolarisation) et maintenu à cette valeur durant 1 ms (période réfractaire absolue). (C) 10 ms après l'émission du potentiel d'action de la cellule source, un potentiel post-synaptique excitateur est observé sur la cellule cible d'amplitude correspondant au poids : 5 mV.

La figure 2.3 est obtenue grâce au code C++ suivant :

```
#include "nsource.h" // librairie de simulation

int main(int argc, char **argv){
    LIF1DNetwork net; // construction d'un réseau
    LIF1DNetwork::dt=0.1; // fixe le pas de simulation à 0.1 ms
    net.addCell("pop0"); // crée une cellule de la population "pop0"
    net.addCell("pop1"); // crée une cellule de la population "pop1"
    net.connect(0,1,5,10,"pop0pop1"); // connecte les 2 cellules avec un
        poids de 5 mV et un délai de 10 ms
    net.setPopI("pop0",110); // ajoute une stimulation en courant imposé
        sur les cellules de pop0
    net.addVRec(0); // ajoute un enregistreur sur le cellule numéro 0
    net.addVRec(1); // ajoute un enregistreur sur le cellule numéro 1
    net.run(50); // exécute une simulation de 50 ms
    net.writeData("./DATA/");
    return 0;
}
```

Les délais de transmission interneuronale (déterminés d'après [Humphries et al., 2006]) et les poids nécessaires à la cohérence de la dynamique du réseau sont résumés dans la Table 2.2.

Pour réaliser nos simulations, nous utilisons une approche de type Monte Carlo, c'est à dire que nous produisons un échantillonnage aléatoire des objets³ réseaux à chaque simulation.

3. Le terme « objet » fait référence ici à la notion informatique de l'instance de classe.

TABLE 2.2 – Poids synaptiques et délais utilisés pour la connectivité entre les sous-populations du réseau. *M* : région motrice ; *C* : région cognitive ; *A* : région associative.

Connection		Poids (mV)	Délais (ms)
Cortex (M/C)	→ Striatum (M/C)	0.1	11
Cortex (M/C)	→ Striatum (A)	0.035	11
Cortex (A)	→ Striatum (A)	0.02	11
Cortex (M/C)	→ STN	0.125	6
Cortex (M/C)	→ Thalamus	0.125	5
STN	→ GPi	2	2
Thalamus	→ Cortex	10	5
Striatum	→ GPi	-5	6
Striatum	→ Striatum	-0.25	2
GPi	→ Thalamus	-10	5

Chaque nouvel objet réseau est donc différent des précédents tout en étant statistiquement cohérent avec les autres. Cela nous permet de déterminer les propriétés génériques du modèle et de tester la robustesse de l’approche. En effet, chaque réseau est construit suivant les mêmes règles, mais nous évitons les biais déterministiques à chaque simulation en introduisant de la stochasticité. Pour chaque réseau nouvellement généré, les poids synaptiques ainsi que les délais de propagation entre les neurones sont initialisés suivant une distribution normale, centrée sur les valeurs de la Table 2.2. Par ailleurs un bruit blanc de courant est ajouté aux neurones pour modéliser l’effet de leur environnement comme les effets thermiques ou le bombardement synaptique continu et aléatoire [Neymotin et al., 2011].

2.2.4 Implémentation de la plasticité synaptique

Il a été démontré que la STDP se produisait (notamment) au niveau des synapses cortico-striatales [Pawlak and Kerr, 2008]. Pour utiliser la STDP dans un modèle de SNN de la taille de celui développé ici, cela nécessite l’ajustement et l’exploration d’un très grand nombre de paramètres qui viennent s’ajouter à ceux qui contrôlent la dynamique du réseau. Le but de cette étude étant de montrer comment se déroulent les mécanismes de prise de décision et d’apprentissage, nous avons choisi d’implémenter une simple règle de LTP/LTD pour la mise à jour des poids. Ce phénomène est contrôlé par la présence ou l’absence de récompense suite au mouvement effectué par le modèle [Cools et al., 2009]. Afin de modéliser l’augmentation ou la diminution des poids synaptiques, nous avons appliqué le mécanisme suivant : l’apprentissage dans le modèle se produit au niveau des synapses cortico-striatales cognitives, où des changements phasiques de la concentration en DA sont nécessaires pour la production des variations des poids synaptiques à long terme [Pawlak and Kerr, 2008, Kerr and Wickens, 2001, Reynolds et al., 2001]. Lorsque le choix d’une cible aboutit à une récompense, tous les poids des projections cortico-striatales cognitives impliquées dans le choix de cette cible sont légèrement augmentés (LTP). Si le choix n’aboutit pas à une récompense, ces poids seront alors diminués selon le même principe (LTD). Pour maintenir les modifications des poids synaptiques dans un certain intervalle, nous avons

implémenté la correction de Klopff qui applique aux variations de poids la classique courbe sigmoïde qui est également observable dans les cinétiques d'apprentissage des animaux [Klopff, 1988]. Appliqué ici cette règle modélise mathématiquement le fait que plus une synapse tend vers sa force maximale, plus elle augmente lentement car la variation s'opère en fonction de ce qui lui reste comme marge de progression par rapport à sa force maximale.

L'apprentissage peut se dérouler dans d'autres structures et voies à l'intérieur des BG, mais nous avons choisi de nous concentrer initialement sur le rôle spécifique de la voie directe cortico-striatale dans l'apprentissage et pour produire le comportement qui nous intéresse ici.

En nous inspirant du protocole appliqué dans le modèle de [Guthrie et al., 2013], lui-même construit sur les résultats expérimentaux de [Pasquereau et al., 2007] nos sessions d'apprentissage consiste en des séries de 120 essais durant lesquels sont enregistrés les instants des potentiels d'action pour chaque neurone de chaque sous-population (chaque neurone est numéroté suivant la sous-population à laquelle il appartient, les détails sont dans les Tables A.2 et A.3), les cibles présentées (forme et direction), la prise de décision du modèle (SB dans les activités corticales motrices et cognitives), l'occurrence de la récompense ou non, les bons choix et les bonnes décisions. Les « bonnes décisions » (« *Good Decision* », GD) sont réalisées lorsque la cible avec la plus haute probabilité de récompense a été choisie dans le cortex cognitif. Les « bons choix » (« *Good Choice* », GC) sont réalisés lorsque la direction sélectionnée dans le cortex moteur correspond également à celle qui présente la plus haute probabilité de récompense. Cette nomenclature est inspirée des données issues des expériences de [Pasquereau et al., 2007] à la fois sur le plan comportemental et électrophysiologique.

2.2.5 Développement et mise au point du réseau

2.2.5.1 Analyse de l'activité corticale

L'activité corticale des sous-régions de notre modèle sert à représenter les activités des entrées sensorielles, cognitives et motrices mais elle sert également à identifier le choix effectué, et donc à déterminer le comportement de l'animal. Les enregistrements de l'activité électrophysiologiques [Pasquereau et al., 2007] montrent que le réseau possède une activité basale en l'absence de stimulation. Dans notre modèle celle-ci sera obtenue par l'injection d'un courant de base (« *Basal Current* », BC) tout au long des simulations. En parallèle, chaque neurone est soumis à un bombardement synaptique aléatoire modélisé par l'application d'un bruit blanc de courant à chaque neurone du réseau comme cela a été mentionné dans la section 2.2.3.3.

La présentation des cibles est symbolisée par l'injection d'un courant supplémentaire au niveau des sous-populations corticales correspondantes (pour rappel, une sous-population correspond à une cible ou une direction comme c'est indiqué dans la Figure 2.2). Lors d'une simulation, chaque essai effectué dure une seconde à partir de la présentation du stimuli. Durant cette période, tous les paramètres mentionnés plus haut (cf. section 2.2.4) sont enregistrés dans des fichiers de données pour être analysés « *offline* ». Nous avons utilisé les instants de chaque potentiel d'action pour chaque neurone, que nous avons ensuite regroupé en fonction des sous-populations neuronales, comme principale métrique électrophysiologique d'analyse des résultats fournis par le modèle. Ces potentiels d'action ont donc été additionnés pour chaque pas de temps (1 ms) dans chaque sous-population et divisés par le nombre de neurones par sous-population. Cette

manipulation nous a permis d'obtenir les fréquences moyennes de décharge neuronale (« *Average Firing rate* », AFR) exprimés en Hz, par sous-population et par session de 120 essais. Les paramètres du modèle ont été ajustés manuellement afin d'obtenir pour chaque région, des AFR qui soient en accord avec les données électrophysiologiques généralement observées. Des techniques d'optimisation des paramètres automatiques ou semi-automatiques existent déjà [Grassia et al., 2011] et auraient pu être appliquées à notre modèle, mais il aurait été laborieux de les mettre en place à ce stade. De plus, il est tout à fait incertains que ces méthodes auraient pu fonctionner pour un modèle comme le nôtre compte tenu de la complexité de la tâche comportementale et de la variabilité des résultats possibles.

Pour associer l'activité de décharge des sous-régions corticales du modèle avec un comportement et donc pour définir le moment où la décision de choisir une cible est prise, nous avons calculé la différence des taux de décharge moyens entre deux sous-populations. Quand cette différence devient supérieure à un seuil fixé à 40 Hz (cf. [Guthrie et al., 2013]), il se produit ce que nous avons appelé une « rupture de la symétrie » (« *Symmetry Breaking* », SB). En effet lors de cette SB, le taux de décharge moyen d'une des sous-population corticale va augmenter au détriment de ceux des autres populations par des mécanismes d'inhibition « *feed-forward* » dus à la structure des boucles mais aussi de façon indirecte en raison de l'inhibition latérale présente dans le striatum.

Quand cette SB se produit dans le cortex cognitif, elle est interprétée comme une prise de décision. Quand elle est observée au niveau moteur, elle est interprétée comme la sélection et l'initiation du mouvement qui sera alors déclenché par la sous-population corticale motrice ayant le plus haut taux moyen de décharge. Les instants de la prise de décision et de l'initiation du mouvement (ou sélection de l'action) sont donc les instants où cette SB est effectivement observée respectivement au niveau cortical cognitif et moteur. Dans le cortex moteur, le mouvement produit à l'issue d'un essai sera codifié par M0, M1, M2 ou M3 selon la direction choisie. Dans le cortex cognitif la codification du choix de la cible sera donnée par C0, C1, C2 ou C3.

Le fait qu'une récompense soit délivrée ou non repose sur la probabilité de récompense $P(R)$ de la cible choisie par la le mouvement. Si le mouvement est en direction de la cible avec la plus haute probabilité de récompense, il s'agit d'un GC même si la récompense n'est pas donnée (elle peut ne pas l'être car elle déterminée par un tirage aléatoire). De la même manière, une GD se produit quand l'activité d'une sous-population corticale cognitive associée à la cible de plus haute probabilité de récompense est sélectionnée. Par exemple après un essai, si l'activité de la région corticale cognitive C3 ($P(R) = 0.66$) est supérieure à celle de C3 ($P(R) = 0.33$), alors nous considérons qu'il s'agit d'une GD.

Une fois une cible choisie par la direction du mouvement, un tirage uniforme sous sa probabilité de récompense associée est réalisé et s'il est réussi la récompense est donnée, même si le choix a amené à sélectionner celle des deux cibles qui a la $P(R)$ la plus basse. La Figure 2.4 résume et clarifie ces notions. Il peut parfois arriver que, après un apprentissage réussi, le mouvement sélectionné ne soit pas en accord lors d'un essai avec le choix de la cible mais c'est un phénomène qui ne se produit que de façon marginale dans nos simulations et qui est imputable à la nature stochastique des simulations.

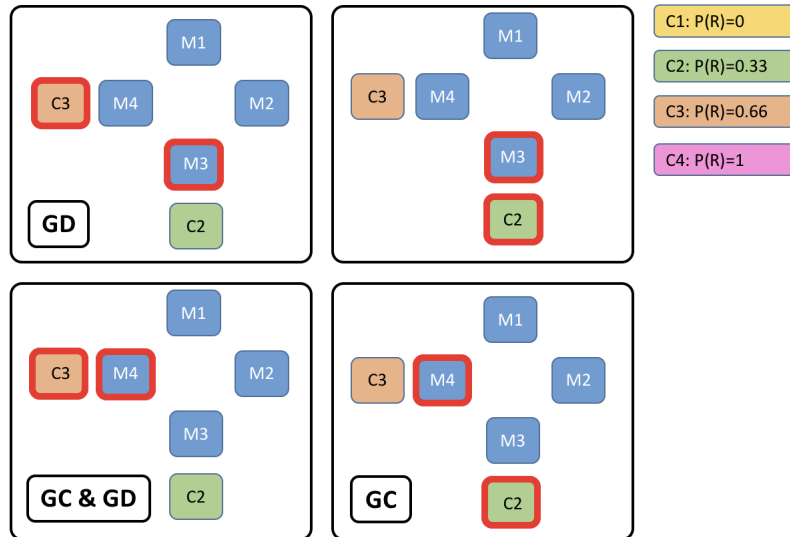


FIGURE 2.4 – Illustration des notions de GC et GD à travers un exemple. Dans cet exemple, la cible C3 est présentée en position M4 et la cible C2 en position M3. C3 est un meilleur choix que C2 comme l’indique le panneau de droite ($P(R)_{C3} > P(R)_{C2}$). Dans la combinaison présentée ici (M4-C3 et M3-C2) une GD correspond donc à la sélection de C3 et un GC à la sélection de M4 puisque c’est le mouvement qui amène à sélectionner C3.

2.2.5.2 Paramétrage et comportement du modèle

La première étape dans la mise au point du modèle a été l’ajustement de ses paramètres pour que le réseau atteigne un état d’activité basale qualitativement et quantitativement cohérent durant la tâche. Même si notre modèle neuronal est relativement simple et homogène, de nombreux paramètres contrôlent en effet cette activité basale. Certains sont impliqués dans la dynamique synaptique et cellulaire alors que d’autres le sont dans la connectivité entre les différentes régions modélisées. Tous ces paramètres ont été ajustés manuellement afin de produire des activités de taux de décharge moyens dans les différentes régions qui soient en accords avec ce qui est observé *in vivo*. Par exemple, l’activité striatale est généralement plus basse comparée à celle du cortex ou du GPi [Pasquereau et al., 2007, Garenne et al., 2011]. De la même manière, les activités des sous-régions thalamiques du modèle montrent une grande variation, mais la gamme des fréquences observées reste, là encore, cohérente avec les données expérimentales [Chen et al., 2010].

Comme cela a été mentionné dans la section 2.2.3.3, nous avons utilisé une approche de type Monte Carlo pour exécuter nos simulations. En raison de la répartition aléatoire des poids synaptiques initiaux et des délais de conduction, chaque instance du modèle est donc unique. Dans le but d’exclure toute possibilité d’un effet d’échantillonnage sur les propriétés fonctionnelles du réseau, 30 simulations pour chacune des six combinaisons de cibles possibles ont été exécutées.

La Figure 2.5 montre un exemple d’enregistrement des activités de chaque sous-région du modèle faisant suite à la présentation des cibles M3-C3 et M2-C4. Les taux de décharge restent globalement cohérents avec le motif d’entrée en prenant en compte de la connectivité. Par exemple une activité élevée dans la région corticale M2 a comme conséquence une activité élevée dans les régions striatales et subthalamiques où elle se projette. À l’inverse, une activité élevée dans une sous-région striatale implique une activité correspondante plus faible dans la

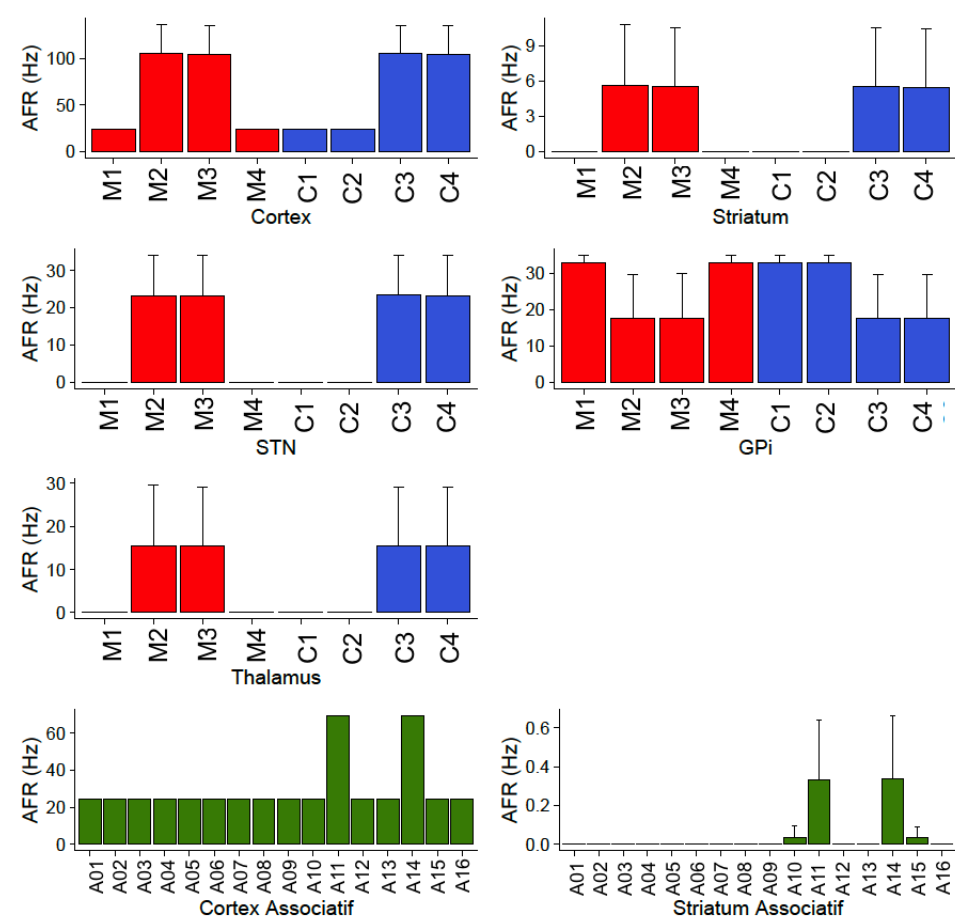


FIGURE 2.5 – Fréquences moyennes d’activité en fonction des sous-régions. Deux cibles sont présentées à un modèle naïf (combinaisons M3-C3 et M2-C4). Les AFR sont enregistrés dans chaque sous-région et regroupés à partir de 30 réseaux différents, enregistrés chacun pendant 30 essais d’une seconde chacun. Les barres d’erreur représentent l’écart-type des données regroupées. Les activités des régions motrices sont représentées en rouge, en bleu pour les cognitives et en vert pour les régions associatives.

région pallidale où elle se projette ce qui est cohérent avec ses effet inhibiteurs.

Cette Figure 2.5 illustre également le fait que les régions corticales stimulées par la présentation d’une cible ont une fréquence moyenne plus élevée, comme nous pouvions logiquement nous y attendre. De plus, nous pouvons y visualiser facilement le résultat de la propagation de l’information à travers le réseau. En nous référant à la Figure 2.1 pour la connectivité du modèle, nous pouvons voir que les sous-populations striatales cognitives et motrices possèdent le même motif d’activation que leurs homologues au niveau cortical. Les sous-populations au niveau du striatum associatif reçoivent la somme des entrées corticales cognitives, motrices et associatives. Dans la Figure 2.5, les pics d’activité au niveau de A11 et A14 peuvent être vus au niveau des sous-populations striatales. Il en est de même pour les pics au niveau de A10 et A15, mais dans une moindre mesure en raison du manque d’entrée corticale associative supplémentaire vers ces sous-régions.

L’activité des sous-régions dans le STN est alignée (par numéro) avec celles observées dans le cortex. Celles qui sont observées dans le GPi sont en revanche l’inverse de celles observées dans les sous-populations cognitives et motrices du striatum. Tout cela est cohérent avec la connectivité et le rôle inhibiteur du striatum.

Enfin, l'activité des sous-régions thalamiques semble également bien résulter de façon conjointe de l'excitation du cortex et de l'inhibition du GPi. Les différences dans l'activité du thalamus seront donc amplifiées. Ces amplifications sont un élément important des mécanismes conduisant à la SB, où l'activité d'un canal augmente alors que celle des autres diminue.

A l'issue de ces simulation, aucune différence significative portant sur l'AFR n'a été retrouvée entre les 30 réseaux générés ($N = 30, p < 0.05, ANOVA$ à une voie). On n'observe donc pas d'effet lié à « l'instance d'un réseau » dans la dynamique du modèle et tous les réseaux générés se comportent sensiblement de la même façon entre eux et de façon cohérente en interne.

2.2.5.3 Sélection de l'action dynamique rupture de la symétrie

Une des propriétés émergentes attendues du réseau est la divergence des activations corticales pour les sous-populations cognitives et motrices, c'est à dire que, même en l'absence d'apprentissage, le réseau est tout de même capable de prendre une décision (de générer une SB). Ce phénomène est l'équivalent d'une prise de décision durant la phase d'exploration de l'apprentissage par renforcement de la part d'un réseau naïf. Au début d'un essai, deux cibles sont présentées. Ce phénomène est simulé par l'injection d'un courant d'intensité équivalente dans chaque neurone concerné (cf. Tables A.2-A.4 pour le détail des numéros de neurone par sous-région) permettant d'activer les représentations des deux formes dans les deux sous-populations corticales cognitives correspondantes et de leurs deux directions dans les deux sous-populations corticales motrices correspondantes.

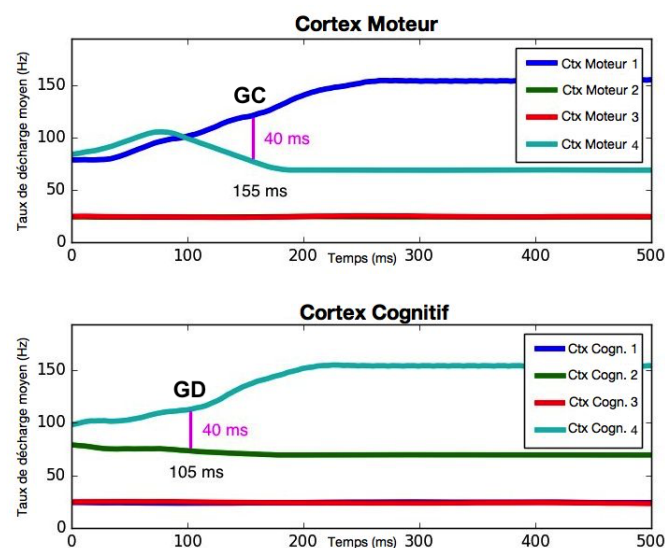


FIGURE 2.6 – Exemples de taux de décharge moyens enregistrés dans des sous-régions du cortex. Les sous-régions enregistrées ici sont M1, M4, C2 et C4 : la cible C3 a été présentée à la position M1 et la cible C2 a été présentée à la position M4. Dans la légende des figures, Ctx Cogn._i correspond à la sous-région *i* du cortex cognitif et Ctx Moteur_i à la sous-région *i* du cortex moteur. Cet essai illustre un exemple dit de « Bon Choix » (« Good Choice », GC) du fait que l'activité motrice sélectionnée soit M1, correspondant à la direction de la présentation de la cible C4 (cible avec la plus grande probabilité de récompense). Les instants $t=155$ ms et $t=105$ ms marqués respectivement sur les enregistrements des sous-régions motrices et cognitives correspondent aux instant des SB. Un GC est donc réalisé à $t=155$ ms et une GD à $t=105$ ms.

Ceci va conduire à l'activation des deux canaux dans la boucle de rétroaction positive de

la voie directe. La SB entre les deux canaux va se produire de la façon suivante : le gain d'une boucle de la voie directe étant plus élevé que celui de l'autre, l'activation différentielle d'un des canaux va être amplifiée sous l'effet conjugué du bruit et de l'état présent des connexions du réseau. L'augmentation de l'activité corticale résultante pour ce canal et le fait que les canaux de la voie directe soient séparés vont initier une rétroaction positive pour ce canal spécifiquement. Cependant, l'activation corticale accrue d'un canal provoque également une augmentation de l'entrée de la voie hyperdirecte divergente à rétroaction négative. Ceci va causer une repression de tous les canaux et notamment de ceux qui sont en compétition. Il en résulte une différence dans l'activation des canaux qui va s'amplifier au fur et à mesure que les signaux parcourent les boucles du réseau.

La connectivité du réseau étant aléatoire mais homogène, en l'absence d'apprentissage la SB va permettre de sélectionner une des deux sous-populations avec des probabilités équivalentes. Par exemple dans la Figure 2.6 où le seuil de 40 Hz pour la SB est représenté, si les cibles C2 et C4 sont présentées respectivement aux positions M4 et M1, la SB devrait se produire la moitié du temps en faveur de C2 et en faveur de C4 l'autre moitié pour ce qui est des décisions cognitives. Il en est de même pour la sélection motrice entre M4 et M1.

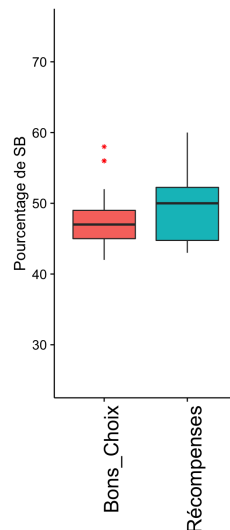


FIGURE 2.7 – Taux moyen de récompense et de GC (« Bon Choix »). Les fréquences moyennes du taux de récompense délivrées et de GC ont été obtenus sur 16 sessions de 100 essais chacune. Les deux scores sont conformes avec le fait d'avoir 50% de chances de réussite ($N=16$, $p<0.05$, test de Wilcoxon).

Le taux moyen de GC est donc de 0.5 dans un modèle naïf, c'est à dire non entraîné. C'est également le cas pour le taux de récompense moyen correspondant à la valeur moyenne de la probabilité : $(1 + 0.66 + 0.33 + 0) / 4 = 0.5$. Ces éléments sont cohérents avec les résultats de nos simulations illustrés dans la Figure 2.7 et qui concernent les taux de de GC et de récompense obtenus. Dans l'essai montré dans la Figure 2.6, le niveau de décision cognitif précède la sélection de l'action de 50 ms, et le délai de la sélection de l'action est de 155 ms après présentation de la cible. Les délais de sélection de l'action et de prise de décision sont estimés grâce à l'utilisation de courbes lissées comme cela a été expliqué dans la section 2.2.2.

2.3 Résultats

Dans cette partie nous allons voir en détails les résultats fournis par le modèle lors de l'étude de sensibilité, étape nécessaire de son développement. Nous aborderons ensuite les effets de lésions localisées du réseau sur l'apprentissage et le comportement du modèle. Nous étudierons ensuite son comportement dans les différentes tâches comportementales envisagées puis nous conclurons cette partie par une discussion sur la validation et les limites du modèle computationnel dans les protocoles comportementaux envisagés.

2.3.1 Étude de sensibilité

2.3.1.1 Effet du bruit synaptique

Comme nous l'avons mentionné précédemment dans la partie 2.2.5.1, chaque neurone d'un réseau biologique est soumis à un bombardement synaptique en conditions physiologiques. Afin de mimer cet effet sur un modèle de neurone LIF, nous avons appliqué un bruit blanc gaussien supplémentaire sous la forme d'un courant appliqué à chaque instant et à chaque cellule indépendamment [Tuckwell, 1989]. La Figure 2.8-A montre que les effets de différents niveaux de bruit modifient qualitativement et quantitativement l'activité du réseau, et en conséquence le taux rupture de symétrie. En effet quand la sous-région corticale la plus active n'a pas au moins une différence d'AFR supérieure à 40 Hz avec la seconde (cf. section 2.2.5.1), nous considérons que la SB n'est pas effective et donc qu'aucune décision n'a été prise ou aucun choix réalisé à l'issue de l'essai.

Avec un niveau de bruit trop faible ou au contraire trop fort, la capacité du réseau à générer une SB est altérée (Figure 2.8-A). Le niveau de bruit ne doit pas dépasser pas un certain seuil car sinon il risque de contrôler l'activité du réseau en la submergeant. S'il est trop faible, le modèle semble incapable d'émettre un choix. Il se comporte de façon moins « exploratoire » ce qui est cohérent dans la mesure où il est effectivement plus déterministe. Un niveau de bruit intermédiaire est donc nécessaire pour permettre la SB dans le réseau, et par conséquent la prise de décision et la sélection d'une action. D'après les résultats de ces simulations de test, nous avons donc fixé le niveau de bruit à 50 pA pour la suite des travaux.

2.3.1.2 Amplitude de l'activité basale

Toutes les structures cérébrales étudiées dans notre modèle sont spontanément actives physiologiquement. Afin de mimer cette activité basale, nous avons donc ajouté un courant de base cortical (BC) en continu. Tout comme pour le bruit de fond, nous avons fait plusieurs tests sur la capacité du modèle à générer une SB en fonction de la quantité de BC. De la même façon que ce que nous avons observé avec le bruit, la Figure 2.8-B montre qu'un BC trop faible ou trop fort empêche le processus spontané de SB et par conséquent de sélection de l'action dans le cortex moteur. En effet, quand ce courant est trop faible, le gain dans les boucles de la voie directe des BG n'est pas suffisant pour déstabiliser leurs activités relatives et de ce fait il n'y a aucune prépondérance entre les sous-populations corticales motrices. À l'inverse lorsque le BC est trop fort, il surpasse la valeur des entrées sensorielles et motrices et contrôle l'activité de tout

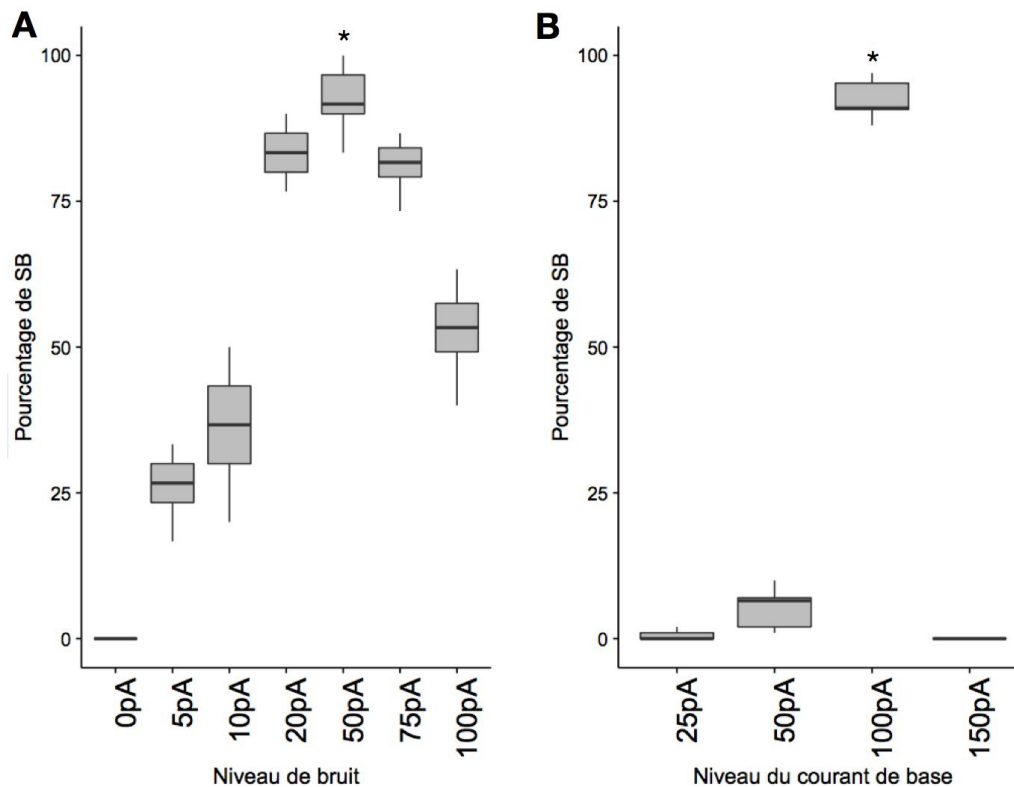


FIGURE 2.8 – Effet du bruit (A) et du courant de base (B) sur les fréquences de prise de décision (SB). L'axe des ordonnées représente le pourcentage de rupture de la symétrie (c'est-à-dire de prise de décision) faisant suite à la présentation de deux cibles choisies et positionnées aléatoirement à deux positions différentes. Les valeurs de la figure ont été obtenues après 16 sessions de 30 essais chacune ($N=16$, $p < 0.05$, test de comparaison multiple de Kruskal-Nemenyi).

le réseau en l'amenant à un état d'équilibre où la compétition entre les canaux est abolie. Nous avons donc fixé le BC à 100 pA afin de permettre au modèle de favoriser l'apparition des SB. Cette valeur permet également au réseau de fournir un niveau d'activité en accord avec les taux de décharge enregistrés *in vivo*. Couplée avec le niveau de bruit approprié, cette valeur de BC nous est donc apparue comme une combinaison optimale à appliquer afin d'obtenir un modèle computationnel efficace pour notre étude.

2.3.2 Tests lésionnels

Afin de vérifier la cohérence de notre modèle, nous avons exécuté une série de simulations en effectuant des lésions en plusieurs points du circuit et observé les effets sur son activité. Pour les mesurer, nous avons lancé 30 sessions d'une seconde chacune, sur huit réseaux différents. Les résultats ne sont pas intégralement présentés ici mais ils montrent que le modèle fonctionne comme nous l'avions prévu en cas de lésions. Pour chaque condition, nous avons interrompu une interconnexion spécifique dans le réseau et enregistré et analysé les activités des différentes régions. Les résultats ont été regroupés dans la Figure 2.9 sous forme d'histogrammes de taux de décharge moyens, en fonction de plusieurs conditions : modèle intact (« *Contrôle* »), lésions striato-pallidales (« *LSP* »), lésions cortico-thalamiques (« *LCT* ») et lésions pallido-thalamiques (« *LPT* »).

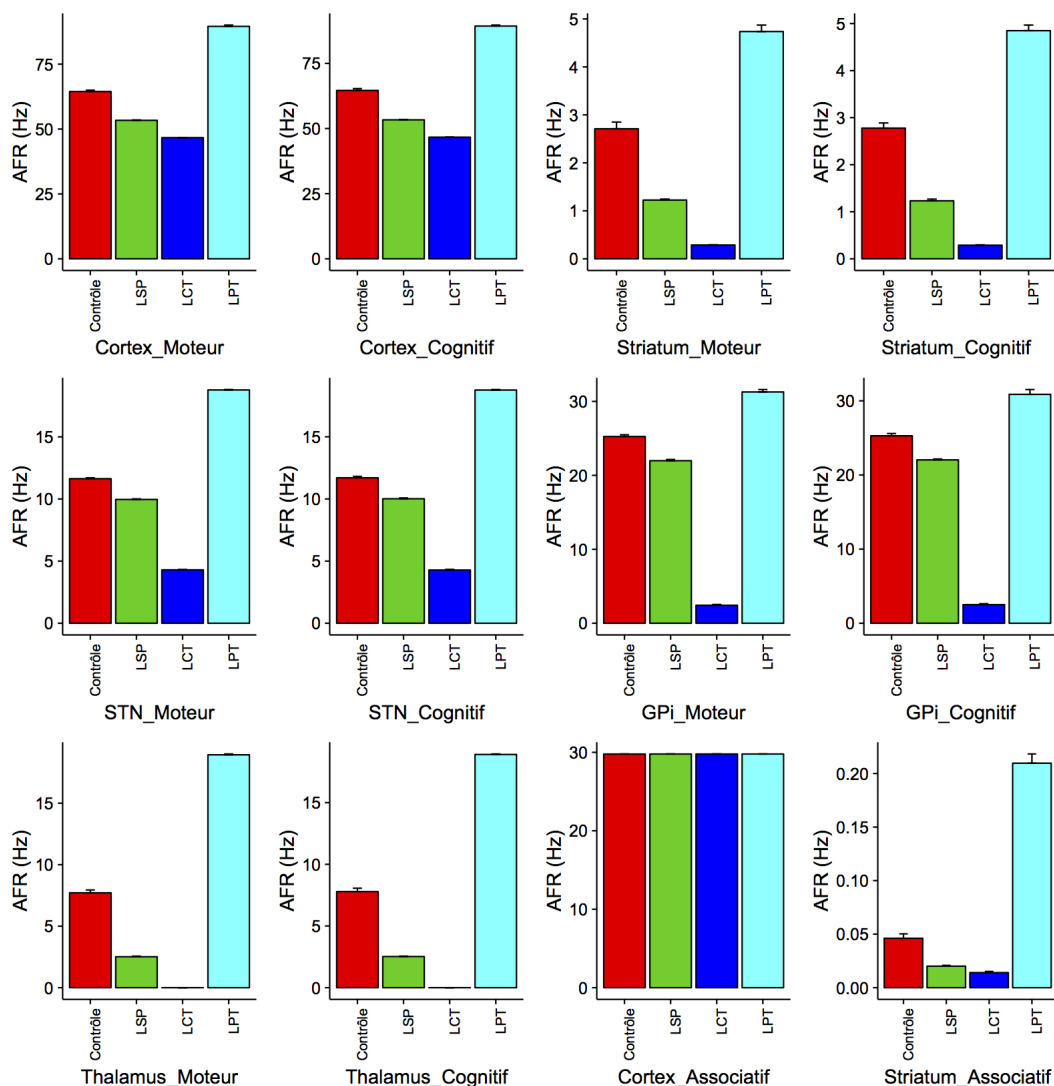


FIGURE 2.9 – *Effet de lésions spécifiques sur l'activité de chaque population du réseau.* L'axe des ordonnées représente les taux de décharge moyens des neurones (AFR), en Hz. L'axe des abscisses représente les différents types de lésions appliquées au modèle : « Contrôle » : Conditions de contrôle (modèle intact) ; « LSP » : Lésions Striato-Pallidales ; « LCT » : Lésions Cortico-Thalamiques ; « LPT » : Lésions Pallido-Thalamiques.

Lésions striato-pallidales Ces lésions vont diminuer d'une manière générale le taux moyen de décharge des neurones dans toutes les régions du réseau. À première vue, une lésion à ce niveau de la voie directe devrait augmenter l'activité du GPI. C'est en effet ce qui se passe quand un stimulus est présenté initialement, mais après quelques dizaines de ms, l'activité du GPI augmente de façon drastique, inhibant l'activité thalamique et donc l'activité corticale. La diminution résultante de l'activité corticale compensant et dépassant la stimulation provoquée par la présentation des cibles, il y a donc une diminution générale dans l'activité des régions.

Lésions cortico-thalamiques Les LCT provoquent également une atténuation globale du taux moyen de décharge dans le réseau. La voie cortico-thalamique est une boucle de rétroaction de gain positif. Quand cette connexion est supprimée, les entrées corticales (principales forces motrices de l'activité du réseau) le sont également et la seule entrée thalamique reste celle provenant du GPI (inhibitrice). L'activité thalamique spontanée en est donc drastiquement

diminuée et par voie de conséquence l'activité corticale. Le taux de décharge global du réseau diminue également, tout comme lors des LSP mais par un mécanisme différent.

Lésions pallido-thalamiques Ce type de lésion permet également un contrôle qualitatif du comportement de notre modèle. En effet, quand les projections inhibitrices du GPi vers le thalamus sont supprimées, les structures thalamo-corticales sont désinhibées. L'activation mutuelle entre le thalamus et le cortex devient donc plus efficace, ce qui conduit à une augmentation des taux de décharge dans les régions corticales et donc secondairement dans toutes les sous-populations du réseau sous son influence.

Conclusion Ces tests lésionnels nous permettent d'avoir une vue globale de la dynamique interne et des interactions au sein de notre réseau. Nous avons donc pu vérifier la cohérence interne du modèle avant de le soumettre aux tâches comportementales et aux protocoles d'apprentissage que nous voulons lui voir réaliser.

2.3.3 Validation dans le protocole initial (Pasquereau et al., 2007)

Nous allons maintenant étudier la concordance des résultats fournis par notre modèle computationnel avec ceux des protocoles comportementaux dont il est issu. Nous verrons tout d'abord ses performances dans le protocole initial à quatre cibles de [Pasquereau et al., 2007] puis dans le protocole complémentaire à deux cibles de [Piron et al., 2016].

2.3.3.1 Propriétés d'apprentissage

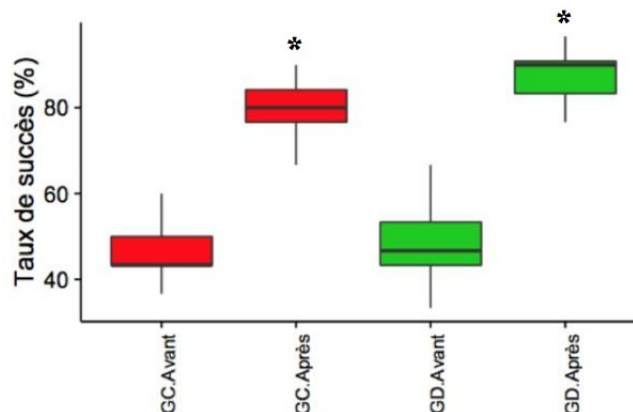


FIGURE 2.10 – *Effet de l'apprentissage sur les capacités d'apprentissage du modèle lors du protocole à quatre cibles. Les taux de GC et GD « Avant » se réfère aux 30 premiers essais de chaque session et « Après » aux 30 derniers. Les résultats sont obtenus sur 16 sessions de 120 essais au total (N=16, $p < 0.05$, test de Wilcoxon).*

Sélection préférentielle de la cible optimale Lors d'une session d'apprentissage standard du protocole à quatre cibles de [Pasquereau et al., 2007], le nombre de GC et de GD pris par le modèle est significativement augmenté (Figure 2.10). En effet, avant ces sessions, la

probabilité que le modèle choisisse la cible « optimale » (celle avec la plus haute probabilité de récompense) est proche de 50%. Après l'apprentissage, cette probabilité passe à près de 90% aussi bien pour les GC que pour les GD. Mis dans des conditions similaires à celles de la tâche, notre modèle est donc capable d'apprendre à choisir efficacement et à sélectionner préférentiellement la cible optimale en vue de maximiser ses récompenses.

Ces sessions d'entraînement montrent que le modèle a donc appris à créer un lien dynamique entre les composantes sensorielles cognitives et motrices d'une cible, sans avoir à modifier les projections corticales motrices. En effet, la Figure 2.11-B montre que la moyenne des récompenses délivrées ainsi que celle des GC augmentent progressivement au cours d'une session.

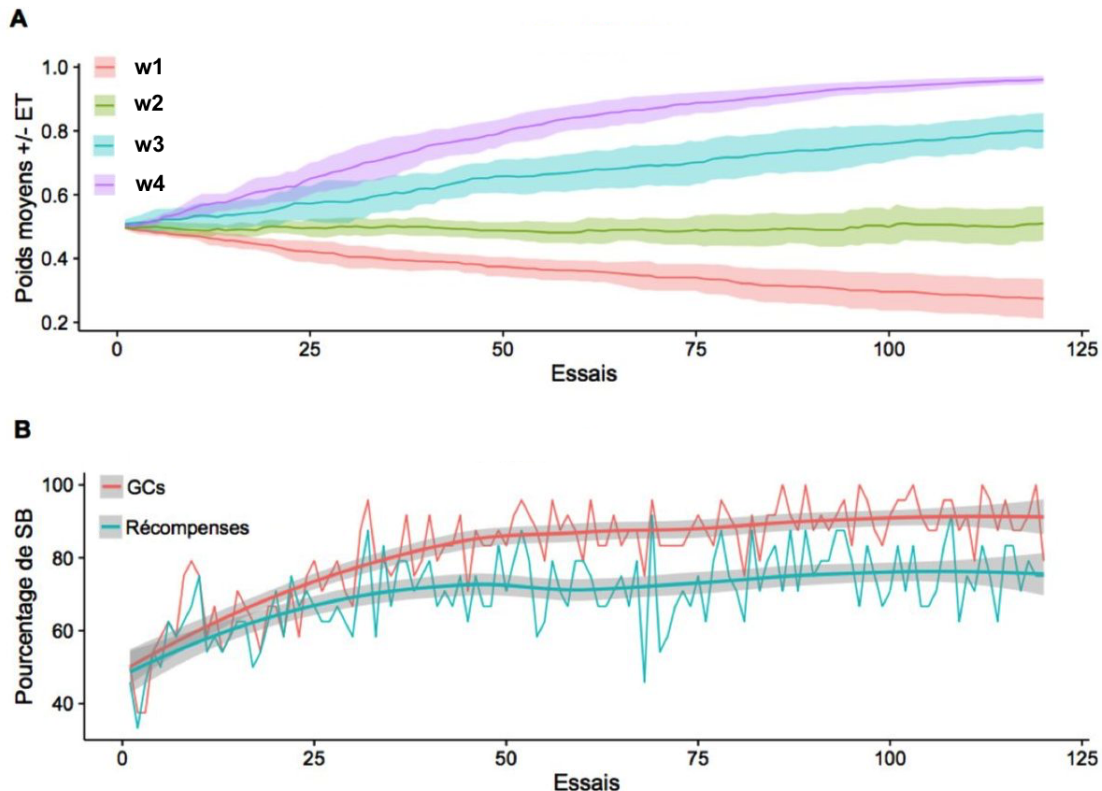


FIGURE 2.11 – Effet de l'apprentissage sur les poids et le comportement du modèle. (A) Variation de la moyenne des poids synaptiques cortico-striataux cognitifs normalisés ($w_1 \rightarrow w_4$) pendant une session d'entraînement. La dispersion est représentée par l'écart-type (E.T). w_i représente la valeur moyenne des poids entre une sous-population corticale cognitive d'indice i codant pour la cible d'indice i et la sous-population striatale cognitive correspondante (cf. Figure A.1-B). Dans ce cas précis et pour simplifier la lisibilité du graphe, les valeurs de récompense associée à chaque cible d'indice i suivent également un ordre croissant ($P_1(R) = 0, P_2(R) = 0.33, P_3(R) = 0.66, P_4(R) = 1$). (B) Normalisation et lissage du taux de GC et de récompenses sur 16 sessions du protocole à quatre cibles. La courbe lissée est obtenue par régression locale polynomiale avec un intervalle de confiance à 95%.

Plus le poids synaptique cortico-striatal va être élevé et plus le gain du canal correspondant sera grand. L'effet est le même que celui décrit pour le bruit lors de la rupture de la symétrie (voir sections 2.2.5.3 et 2.3.1.1), faisant en sorte que le canal avec le poids cortico-striatal le plus élevé soit en général sélectionné (Figure 2.11-A). En effet, la voie hyperdirecte va exciter simultanément les différentes régions du GPi, qui vont inhiber sélectivement les entrées thalamiques projetant sur les régions corticales dont les projections striatales ont les poids les

plus bas. Comme nous l'avons vu dans la section 2.3.2, le fait de couper les projections thalamo-corticales dans le réseau permet de bloquer le processus de SB, montrant que la prise de décision est bien due aux propriétés de bouclage du circuit.

Diminution du délai d'initiation du mouvement avec l'apprentissage Comme nous l'avons vu dans la section 2.2.5.3, il y a en général un délai entre le moment auquel la décision est prise au niveau cognitif et celui auquel le mouvement est initié au niveau moteur (respectivement les GD et les GC). Nous allons ici nous intéresser au temps de réaction du modèle (ou délai d'initiation du mouvement) qui est déterminé par la différence de temps entre le moment de la présentation de deux cibles à l'écran ($t=0$) et celui où le seuil des 40 Hz est atteint (cf. la Figure 2.6-haut pour un exemple). La Figure 2.12 montre que ce temps de réaction est diminué avec l'apprentissage. Ces résultats sont en accord avec de précédentes études de modélisation [Guthrie et al., 2013] mais également avec les résultats expérimentaux. En effet les temps de réaction dans ce genre de tâche sont plus longs avant un apprentissage que ceux observés en conditions de routine ce qui est logique dans la mesure où pour un réseau naïf, la connaissance de la valeur de chaque cible n'est pas encore renforcée [Piron et al., 2016].

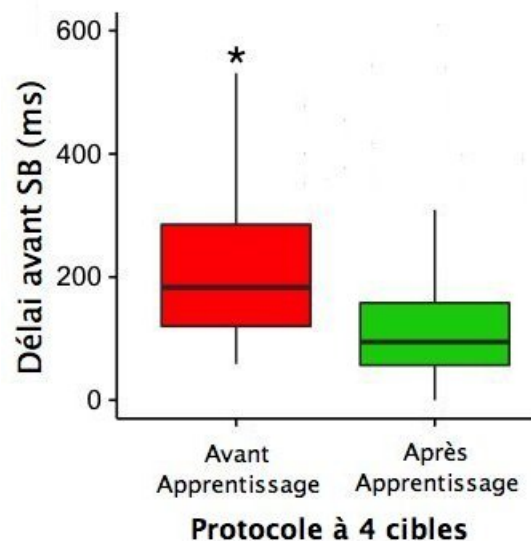


FIGURE 2.12 – Délai moyen de déclenchement du mouvement. La figure montre l'évolution des temps moyens de réaction mesurés en ms avant et après un apprentissage réalisé lors de 16 sessions du protocole à quatre cibles ($N=16$, $p<0.05$, test de Wilcoxon).

Les projections cortico-striatales codant pour les cibles avec les probabilités de récompense les plus élevées sont renforcées par l'apprentissage, conduisant mécaniquement à une augmentation de leurs poids synaptiques. Cette augmentation a deux conséquences pouvant expliquer un temps de réaction moyen plus court :

- plus un poids synaptique sera important et plus court sera le temps nécessaire à la génération d'un spike post-synaptique (notamment à cause du phénomène de sommation temporelle mais aussi parce que les courants entrants excitateurs seront plus importants et que le seuil sera atteint plus vite)
- l'apprentissage fait que le modèle va sélectionner préférentiellement la cible avec la plus grande probabilité de récompense, et donc les entrées striatales avec les poids les plus

importants, ce qui contribue également à réduire les retards synaptiques pour des raisons similaires à celles mentionnées au dessus.

Ces deux phénomènes amènent logiquement à la réduction du temps moyen de réaction dans notre modèle au cours de l'apprentissage.

Rôle de l'inhibition striatale latérale sur les capacités d'apprentissage du modèle

L'inhibition striatale latérale (SLI) joue un rôle important dans la dynamique interne du réseau et donc dans le traitement de l'information. Notre modèle nous offre la possibilité de supprimer ce mécanisme de façon spécifique pour étudier son effet sur ses capacités fonctionnelles. La Figure 2.13 montre que, quand ce phénomène est absent, les taux de GC et de GD sont altérés.

Avant l'apprentissage, l'absence de la SLI a tendance à diminuer le taux de GC comme nous pouvons le voir sur la Figure 2.13-A-s1-s2. Après l'apprentissage le taux de GC s'effondre en l'absence de SLI (Figure 2.13-A-s3-s4). Ce premier résultat confirme le rôle que peut jouer la SLI dans le contrôle du gain au sein du réseau [Jason T. Moyer and Wolf, 2014]. Nous nous sommes néanmoins interrogés sur le score très bas obtenu par le modèle en l'absence de SLI qui semblait paradoxal dans la mesure où il était inférieur à ce qui pouvait être obtenu pour un tirage au hasard à 50% (Figure 2.13-A-s4).

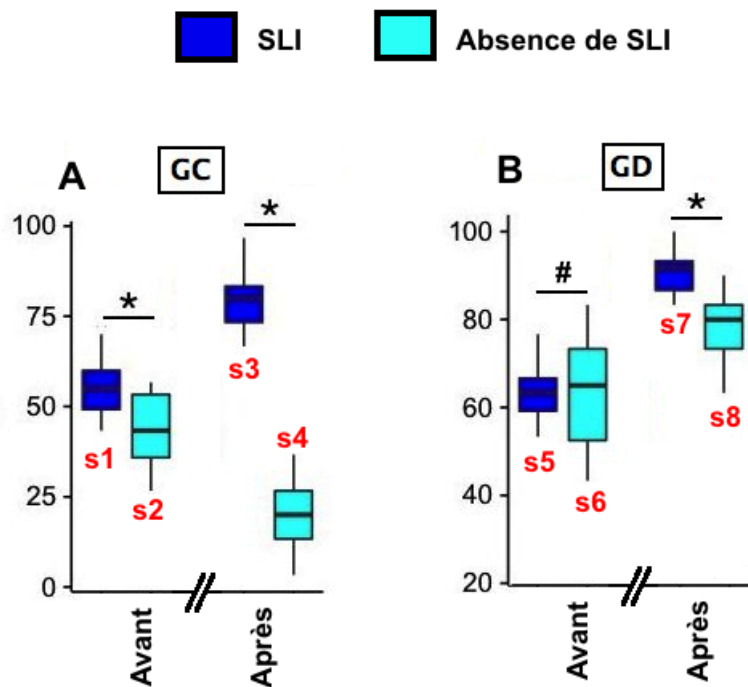


FIGURE 2.13 – Effets de l'absence de SLI sur les taux de réussite et sur la performance d'apprentissage du modèle. « Avant » signifie que le réseau est naïf (avant apprentissage) et « Après » signifie que le réseau est entraîné (après apprentissage). (A) Le pourcentage de GC diffère significativement de celui observé avant et après l'apprentissage quand il n'y a pas d'inhibition latérale. (B) Après l'apprentissage, le pourcentage de GD est diminué en cas d'absence de SLI ($N=16$, $p<0.05$, tests de Mann-Whitney-Wilcoxon).

Nous avons alors constaté alors que le pourcentage de ruptures de symétries (SB) au niveau cortical moteur s'était également effondré. Comme la survenue d'un GC (ou d'un mauvais choix)

nécessite au minimum qu'un choix soit effectué⁴ il est logique de noter un taux de réussite qui soit paradoxalement pire que pour un réseau naïf. L'absence de SLI biaise donc considérablement les capacités d'apprentissage du réseau.

Au niveau cognitif, il apparaît que l'absence d'inhibition n'a pas d'effet sur le score de GD d'un réseau naïf (Figure 2.13-B-s5-s6) en dehors d'une augmentation visible de la dispersion. Cependant après l'apprentissage, l'absence de SLI provoque une diminution du score de GD (Figure 2.13-B-s7-s8). Ce phénomène peut s'expliquer au moins en partie par une déficience du processus de sélection de l'action. En effet la sélection de l'action étant déficiente comme nous l'avons vu sur la Figure 2.13-A-s2-s4, elle empêche le modèle d'améliorer ses performances dans l'apprentissage des valeurs associées aux cibles en ne favorisant pas suffisamment la sélection de la bonne cible par le mouvement approprié et donc l'obtention de récompenses.

L'ensemble de ces résultats montrent le rôle joué par la SLI dans le traitement de l'information par notre modèle de BG et par conséquent dans le processus d'apprentissage.

2.3.3.2 Protocole d'inversion de l'apprentissage

L'apprentissage inversé (ou RevL pour « *Reversal Learning* ») consiste en une procédure d'entraînement de plusieurs étapes au cours desquelles le sujet va devoir modifier ce qu'il a appris précédemment. Dans ce protocole en effet, les associations stimulus-récompense vont être elles-mêmes modifiées et ce, que ce soit explicitement signalé ou non en cours de tâche. En pratique si une cible A apporte un bénéfice 1 et une cible B un bénéfice 2, à l'étape suivante, les deux bénéfices associés à chaque cible sont échangés. Sachant que les animaux sont capables d'adapter leurs facultés d'apprentissage expérimentalement, nous avons donc soumis le modèle à un protocole de ce type. Lors de la première étape, le sujet (modèle) doit apprendre à effectuer les choix adéquats et à prendre les bonnes décisions dans le but de maximiser sa récompense comme cela a été réalisé précédemment. Une fois un certain niveau de performance atteint, une autre étape « *reversal* » est mise en place.

Cette dernière consiste à modifier les critères de récompense. L'individu va ainsi devoir adapter son comportement le plus vite possible pour maintenir un niveau de récompense optimal. Cette méthode est souvent utilisée dans l'étude des mécanismes neurophysiologiques [Dombrowski et al., 2015, Cools et al., 2002, Costa et al., 2015] et également mais de façon plus marginale dans le développement de méthodes de diagnostic innovantes [Buelow et al., 2015]. Dans le but d'évaluer les caractéristiques d'apprentissage de notre modèle, nous avons implémenté un protocole de RevL qui a été conçu sans chevauchement des conditions de récompenses [Berthet et al., 2012]. Ses caractéristiques sont résumées dans la Table 2.3.

Dans notre protocole, la phase initiale d'entraînement (étape 1) comprend 120 essais. Lors du passage à l'étape 2, les probabilités de récompense de toutes les cibles ont été changées de manière à maintenir une variation de la probabilité de récompense homogène pour toutes les cibles de 2/3. Dans cette seconde période, 240 essais supplémentaires ont été exécutés avec cette configuration. Enfin, dans l'étape 3, les probabilités de récompense initiales (celles de l'étape 1) ont été remises en place et 240 nouveaux essais ont été réalisés ce qui génère des sessions de

4. Pour rappel un choix est réalisé dans notre modèle quand une SB est réalisée au niveau moteur ce qui implique qu'une sous-région présente une activité supérieure de 40 Hz aux autres sous-régions.

TABLE 2.3 – Paramètres du protocole d'inversion de l'apprentissage.

Nombre de cibles	C0	C1	C2	C3
Étape 1	P(R)=0	P(R)=1/3	P(R)=2/3	P(R)=1
Étape 2	P(R)=2/3	P(R)=1	P(R)=0	P(R)=1/3
Étape 3	P(R)=0	P(R)=1/3	P(R)=2/3	P(R)=1

600 essais au total.

Le modèle se comporte de façon cohérente avec ce qui a été déjà observé expérimentalement, notamment en ce qui concerne la récupération des taux de GC [Xue et al., 2013, Morita et al., 2016]. Chaque nouvelle distribution des probabilités de récompenses a été suivie par une adaptation du modèle qui a spontanément réussi à améliorer son score, et ce après quelques dizaines d'essais (Figure 2.14-A-C). Nous avons également réalisé le suivi du processus d'apprentissage au cours des essais à travers l'évolution des poids des projections cortico-striatales cognitives (Figure 2.14-B).

Comme cela a été observé lors d'études précédentes, le modèle a besoin de plus d'essais après une étape d'inversion pour atteindre un taux de GC équivalent par rapport à ce qu'il obtenait juste avant (Figure 2.14-A). Ce phénomène pourrait s'expliquer par le fait que l'inversion se produisant après la première étape de l'entraînement, le modèle va spontanément sélectionner ce qu'il considère comme un GC, mais avec des probabilités de récompenses inversées. Le modèle va donc porter ses choix vers les plus mauvaises cibles durant une courte période d'adaptation aux nouvelles conditions. C'est ce mécanisme qui explique la diminution temporaire de son score de GC sous le taux des 50% initiaux obtenu à l'état naïf [Morita et al., 2016].

Bien qu'intéressants, ces résultats doivent être nuancés au regard des résultats expérimentaux obtenus *in vivo* quand les valeurs des probabilités de récompenses changent. En effet, durant un protocole d'apprentissage inversé comprenant de nombreux changements de probabilités, les singes ont tendance à s'adapter plus rapidement à une nouvelle configuration au cours des étapes successives. S'ils se contentaient de mettre en œuvre un apprentissage par renforcement classique pour la mise à jour des valeurs de cible, nous devrions observer des temps de ré-apprentissage identiques [Costa et al., 2015]. Or en réalité, ils sont capables d'apprendre plus rapidement les nouvelles configurations au cours des différentes étapes de RevL afin d'optimiser leur taux de récompense.

Cette diminution du temps d'apprentissage n'est pas incompatible avec nos résultats, mais elle souligne le fait qu'il manque à notre modèle un processus d'apprentissage d'ordre supérieur ou bien un mécanisme de modulation pour que celui-ci soit plus en adéquation avec les observations expérimentales. Cette capacité d'adaptation plus rapide aux changements des probabilités de récompense en fonction des cibles pourrait être attribuée, du moins en partie, à l'hippocampe. Cette structure est en effet connue pour jouer un rôle dans l'apprentissage par configuration (« *configural learning* ») [Shohamy et al., 2009]. Notre réseau n'est pour le moment pas capable d'améliorer son adaptation au cours de temps, et le fait d'ajouter ce « contrôleur » supplémentaire pourrait sans doute y contribuer.

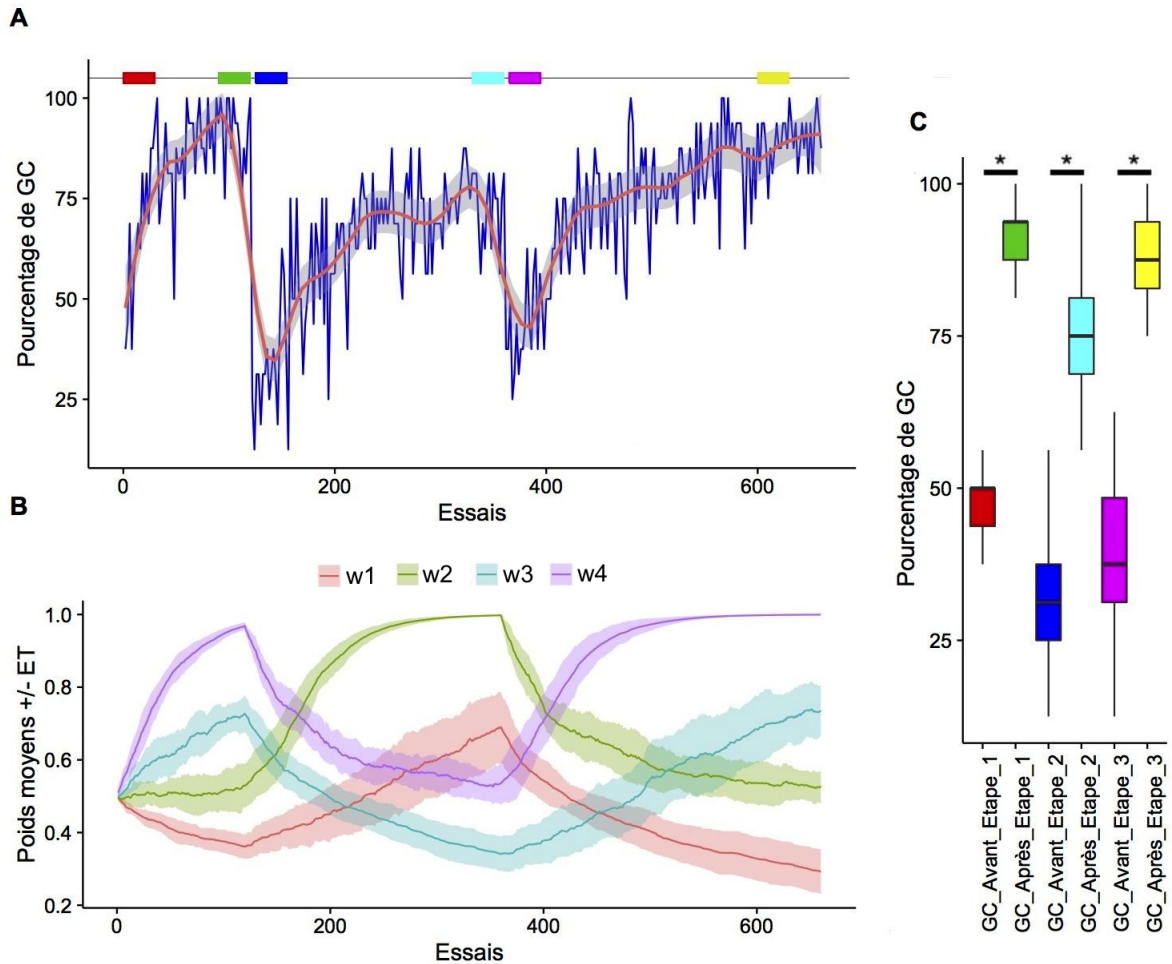


FIGURE 2.14 – Résultats des simulations pour les trois étapes du protocole d’inversion de l’apprentissage (A) Pourcentage de GC pendant les sessions d’entraînement (en rouge la courbe lissée est obtenue par régression locale polynomiale avec un intervalle de confiance à 95%). Des échantillons de 30 essais (boîtes colorées en haut de la figure) ont été analysés au début et à la fin des trois étapes afin de comparer l’effet de l’apprentissage inversé. Les couleurs des boîtes sont reportées à l’identique dans le « whisker-box » de la partie C de la Figure. (B) Variation des moyennes des poids synaptiques cortico-striataux cognitifs normalisés pendant une session d’apprentissage. Les w_i de la légende représentent les valeurs moyennes des poids pour chacune des quatre cibles de même index i . Les associations entre numéro de cible et probabilité de récompense sont similaires à celles ayant été choisies dans la Figure 2.11. La dispersion est représentée par l’écart-type (E.T). (C) Comparaison du taux de GC avant et après l’entraînement en fonction des modifications des probabilités de récompense des cibles ($N=16$, $p<0.05$, test de comparaison multiple de Kruskal-Nemenyi). « Avant » signifie avant apprentissage et « Après » signifie après l’apprentissage.

2.3.3.3 Protocole d’extinction

L’apprentissage par renforcement partage d’importantes caractéristiques avec les travaux de Pavlov sur le conditionnement classique [Pavlov, 1927] et avec celles de Skinner sur le conditionnement opérant [Skinner, 1979]. Dans l’expérience de Pavlov par exemple, lorsque le stimulus conditionné n’est plus présenté avec le stimulus non conditionné, la réponse conditionnée diminue progressivement. Plus généralement dans les tâches d’apprentissage, un phénomène similaire se produit quand un stimulus appétitif (ou au contraire aversif) est supprimé après l’apprentissage par renforcement. Notre modèle devrait être capable de prédire, du moins partiellement, ce

phénomène d’extinction car il s’agit d’une capacité assez commune dans la plupart des processus d’apprentissage associatif [Barker et al., 2014]. Nous avons donc soumis notre modèle à un protocole d’extinction simple, c’est à dire que nous avons supprimé les récompenses après une période d’apprentissage et ce, quelques soient le comportement et les réponses du modèle [Goodman and Packard, 2016, Papachristou et al., 2013]. Les résultats sont présentés dans la Figure 2.15.

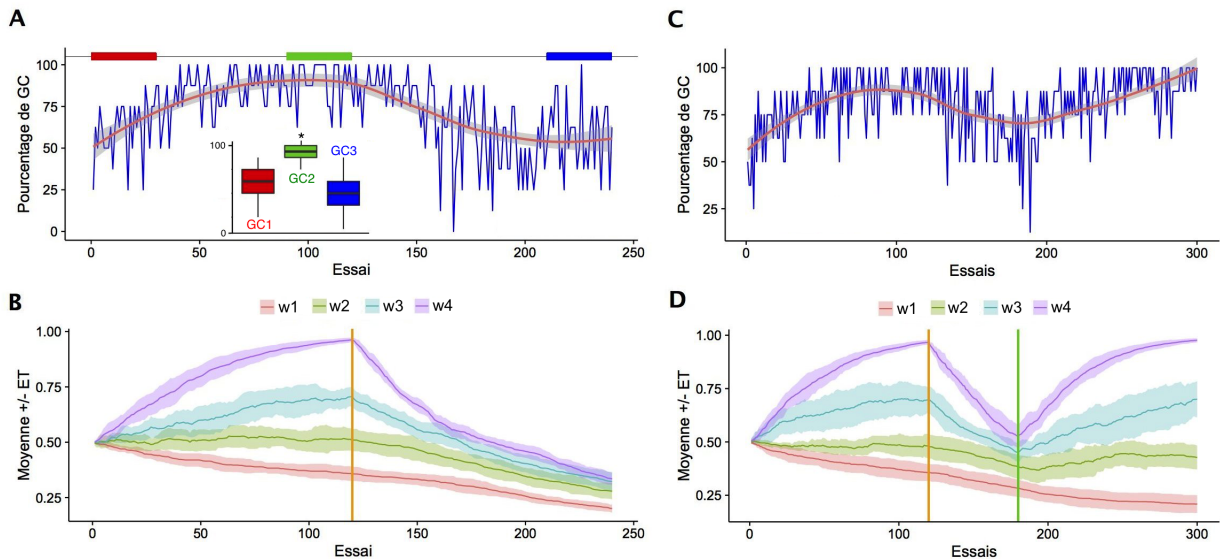


FIGURE 2.15 – Résultats des simulations du protocole d’extinction (A) Taux de GC au cours d’un protocole d’extinction simple. Afin de tester l’effet de ce protocole, nous avons analysé des échantillons de 30 essais (boîtes colorées) au début de la session (GC1), avant (GC2) et après (GC3) l’extinction. Les mesures des taux de GC (insert) suivent les étapes de l’extinction ($N=16$, $p<0.05$, test de comparaison multiple de Kruskal-Nemenyi). La courbe lissée est obtenue par régression locale polynomiale avec un intervalle de confiance à 95%. (B) Variation des moyennes des poids synaptiques cortico-striataux cognitifs normalisés au cours des essais. Les w_i de la légende représentent les valeurs moyennes des poids pour chacune des quatre cibles de même index i . L’extinction commence à l’essai numéro 120 (marqueur orange). Les associations entre numéro de cible et probabilité de récompense sont similaires à celles ayant été choisies dans la Figure 2.11. La dispersion est représentée par l’écart-type (E.T). (C) Pourcentage de GC et (D) variation des moyennes des poids synaptiques cortico-striataux cognitifs normalisés au cours d’un protocole d’extinction simple de 60 essais, suivi par un rappel de 120 essais (300 essais en tout). L’extinction commence à l’essai numéro 120 (marqueur orange) et fini au numéro 180 (marqueur vert). Les autres conditions sont similaires à celles appliquées en A et B.

La même procédure d’entraînement que précédemment est appliquée initialement (120 essais) puis les récompenses ne sont plus délivrées pendant les 120 essais supplémentaires. Enfin durant un nouveau cycle de 120 essais les récompenses sont redonnées sur les mêmes critères que durant la phase initiale. Au cours de la phase d’extinction les performances de GC vont progressivement diminuer, comme le montre la Figure 2.15-A. Le score moyen de ces performances présente également une différence significative avant et après l’extinction. Ceci se produit en parallèle avec la variation des poids synaptiques cortico-striataux cognitifs au cours des essais (Figure 2.15-B). Ces deux résultats sont globalement cohérents avec le processus d’extinction tel qu’il est généralement décrit dans un contexte de conditionnement classique ou opérant. En effet, nous montrons ici que, quand le modèle est privé de récompense, il va perdre spontanément sa capacité d’optimisation de ses choix et « oublier » progressivement ce qu’il a appris (Figure

2.15-A-insert).

Il a par ailleurs été démontré que le temps de rappel d'une probabilité de récompense associée à une cible suite à une première extinction était plus court que celui mesuré lors de la phase initiale d'apprentissage [Bouton, 2004]. Ce phénomène assez général est apparu au cours de l'évolution chez les invertébrés [Sandoz et al., 2016]. Dans sa forme actuelle, notre modèle ne présente pas un tel comportement. Quand la phase d'extinction est raccourcie nous pouvons observer un léger effet de mémoire mais en pratique le temps de rappel n'est pas significativement plus court que lors de la phase initiale d'apprentissage (Figure 2.15-C-D).

Tous ces résultats montrent que les propriétés d'extinctions de notre modèle ne traitent que partiellement du processus complexe de l'extinction. Ceci est assez simple à expliquer dans la mesure où toutes les structures cérébrales impliquées dans ce phénomène ne sont pas incluses pour le moment. Par exemple, les interactions entre le striatum dorso-latéral et le noyau central de l'amygdale jouent un rôle majeur dans ce mécanisme [Lingawi and Balleine, 2012] et ces structures ne sont pas prises en compte dans le modèle pour le moment. Cependant, malgré ces restrictions, notre modèle adopte des comportements en accord, au moins qualitativement, avec les résultats des données expérimentales actuelles et est en mesure de les expliquer partiellement.

2.3.4 Validation dans le protocole complémentaire (Piron et al., 2016)

Pour tester de l'adaptabilité de notre modèle « en l'état » nous l'avons soumis à un contexte de tâche différent sans le modifier. Nous nous sommes donc inspirés de la tâche d'apprentissage à deux cibles utilisée dans [Piron et al., 2016] sans toutefois y inclure la complexité des alternances nouveauté/routine. Dans ce protocole les probabilités de récompense des deux cibles sont respectivement de 0,25 et 0,75. Comme nous allons le constater, l'ensemble des résultats obtenus est assez similaire à ce qui est observé dans le protocole à quatre cibles. C'est également cohérent avec les résultats expérimentaux du papier princeps et notre modèle des BG rend compte ici encore de ses capacités d'apprentissage et de son adaptabilité.

2.3.4.1 Sélection préférentielle de la cible optimale

Le protocole à deux cibles produit des résultats similaires à celui à quatre cibles, à savoir que le taux de succès (GC et GD) est significativement supérieur après l'apprentissage, comme le montre la Figure 2.16. Ce taux est même supérieur à celui du protocole à quatre cibles (environ 90% de GD pour le protocole à quatre cibles, et près de 100% pour celui à deux cibles). Cette différence peut s'expliquer par le niveau d'incertitude plus faible dans le cas du deuxième protocole, car seulement deux formes différentes sont alors en compétition, avec une différence entre les probabilités de récompense sensiblement plus importante (0,5 au lieu de 0,33).

2.3.4.2 Diminution du délai d'initiation du mouvement avec d'apprentissage

Comme nous l'avons constaté pour le protocole à 4 cibles dans la partie précédente, le temps de réaction pour la sélection d'une cible est diminué avec l'apprentissage dans le cas du protocole à 2 cibles (Figure 2.17). Ces résultats sont en accord avec les résultats précédents aussi

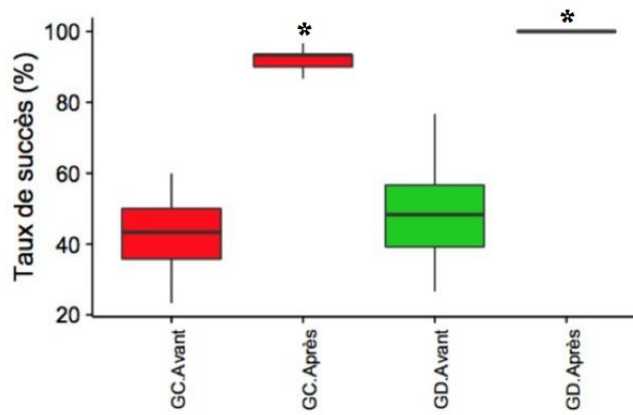


FIGURE 2.16 – Etude des capacités d’apprentissage du modèle sur un protocole à deux cibles. « Avant » se réfère aux 30 premiers essais et « Après » aux 30 derniers, sur une sessions de 120 essais au total ($N=16$, $p<0.05$, test de Wilcoxon).

bien sur le plan expérimental [Bloch et al., 2016] que théorique [Guthrie et al., 2013, Piron et al., 2016]. Sur un plan plus intuitif et avec toutes les réserves que cela implique, nous sommes en général conscient du fait que la pratique d’un geste ou d’une tâche nous permet de le réaliser progressivement de façon plus efficace et au besoin aussi, plus rapidement.

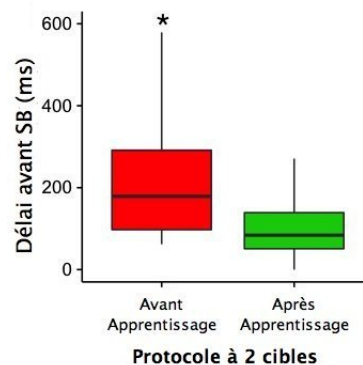


FIGURE 2.17 – Délai moyen de déclenchement du mouvement après apprentissage lors du protocole à deux cibles. La figure montre l’évolution des temps moyens de réaction mesurés en ms avant et après un apprentissage réalisé lors de 16 sessions du protocole à deux cibles. ($N=16$, $p<0.05$, test de Wilcoxon).

2.3.4.3 Effet de la lésion du GPi sur les capacités d’apprentissage du modèle

Pour compléter les tests du modèle en comparant ses capacités et limites avec celles de contrôles expérimentaux, nous avons répliqué *in silico* l’effet d’une lésion du GPi sur le processus d’apprentissage comme cela a été pratiqué dans l’étude de [Piron et al., 2016]. Dans leur protocole, les auteurs pratiquent des injections de muscimol dans le GPi ce qui inactive son activité. Ils constatent que si cette lésion n’a pas d’effet sur les conditionnements préalablement effectués, en revanche un apprentissage *de novo* est empêché et les animaux restent à des taux de succès proches de 50% ce que nous observons également chez un modèle naïf.

Cet effet est visible sur nos simulations. Après avoir inactivé le GPi et donc rendues inopérantes ses projections nous constatons qu’il devient impossible pour le réseau d’apprendre à

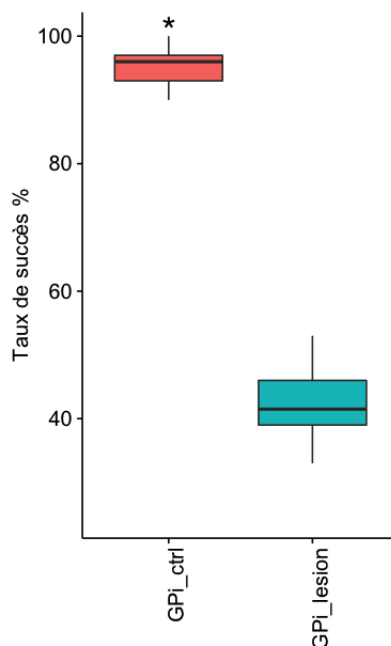


FIGURE 2.18 – *Effet d’une lésion du GPi sur les performances d’apprentissage du modèle.* L’axe des ordonnées montre les taux de réussite des GC en condition contrôle (GPI_ctrl) et après une lésion du GPi (GPI_lesion) calculés après apprentissage sur les 30 derniers essais de 16 sessions. ($N=16$, $p<0.05$, test de Wilcoxon).

réaliser la tâche à 2 cibles comme le montre la Figure 2.18. Ce phénomène était assez prévisible dans la mesure où en inactivant le GPi, nous bloquons du même coup les boucles cortico-corticales.

2.4 Discussion

2.4.1 Apports et prédictions du modèle

L’approche connexionniste que nous avons appliquée à l’étude de la prise de décision dans les BG poursuivait un double objectif, à la fois scientifique et méthodologique.

Objectif méthodologique Notre méthodologie n’était pas novatrice en soi, que ce soit en terme de matériel utilisé comme le type des neurones, les méthodes de connectivité ou encore d’activation synaptique. Elle nous a cependant permis de démontrer qu’à grande échelle, un modèle computationnel de la circuiterie des BG au niveau neuronal pouvait être implémenté, développé puis paramétré et exécuté dans un temps raisonnable et ce, avec une puissance de calcul et une mémoire standard (cf. Table A.5). Ce travail nous a fourni le cadre ainsi que les outils pour une exploration plus poussée des processus de prise de décision, en conditions normales et pathologiques et en pouvant tenir compte des mécanismes à l’échelle cellulaire.

Objectif scientifique Le volet scientifique de notre étude a porté sur la description des mécanismes de la prise de décision au niveau synaptique et neuronal. Il était important de montrer qu’un modèle connexionniste de la circuiterie des BG à l’échelle cellulaire était capable

de réaliser le processus de sélection de l'action lors d'une tâche comportementale complexe et donc d'étendre des résultats obtenus à l'aide d'autres formalismes plus généraux. Dans le but d'identifier clairement les flux d'information dans le modèle ainsi que le rôle des interactions entre ses composants, nous avons considéré qu'il était nécessaire d'amener le niveau de description du modèle en le ramenant à ses unités de traitement élémentaires, à savoir les neurones et synapses. En effet, une question se pose habituellement à ce point de la discussion : pourquoi utiliser une approche fondée sur un formalisme neuronal (« *spiking-model* ») au lieu d'une approche utilisant un formalisme de population (« *rate-model* »)?

2.4.1.1 Existe-t-il une échelle idéale pour suivre les flux d'information ?

La question reste ouverte de décider si oui ou non les modèles de populations peuvent imiter de façon fiable les dynamiques émergentes des interactions entre les unités neuronales. Une autre façon de présenter la question serait : est-il nécessaire de descendre à l'échelle du spike et du neurone pour expliquer clairement les mécanismes computationnels à l'œuvre dans les circuits neuronaux ?

Les modèles décrits à l'échelle de la population sont indéniablement utiles pour pouvoir gérer de grands ensembles de neurones avec des ressources calculatoires limitées ou en présence d'une quantité de paramètres les rendant difficiles à ajuster. Ils peuvent à ce titre être utiles en tant qu'étape intermédiaire lors d'une approche de modélisation « *top-down* »⁵ d'un phénomène biologique complexe. Dans ce type d'approche, l'on cherche en général à établir rapidement une relation entre les observations comportementales, les enregistrements électrophysiologiques et la structure générale de la région du système nerveux étudié, compartimenté en sous-populations. Les compartiments (ou sous-populations) constituent supposément des éléments anatomo-fonctionnels distincts et autonomes mais dont la dynamique globale n'est pas considérée en délaissant les propriétés particulières de leurs composants. Cette approche réductrice (le terme est à prendre ici au sens propre et non pas au sens péjoratif) est susceptible de générer au moins deux types de biais.

(i) Le premier consiste à accorder trop d'importance aux délimitations anatomiques et anatomo-fonctionnelles telles qu'elles existent ou ont été décrites. En effet, en compartimentant certaines régions en raison de leur supposée spécialisation, nous risquons de passer artificiellement à côté d'effets synergiques ou plus simplement de créer des limites là où la nature les ignore. Il est tentant de distinguer les régions en fonction de leur types cellulaires, de leur origine embryologique ou encore de la présence ou de l'absence de tel ou tel type de récepteur. C'est d'ailleurs une étape nécessaire dans la compréhension du système considéré dans son intégralité. Toutefois, il apparaît de plus en plus évident que les régions du CNS sont redondantes, qu'elles fonctionnent en parallèle et même qu'elles sont capables de se compenser mutuellement jusqu'à un certain point. Il apparaît également évident que des structures auparavant décrites comme « motrices » ou « sensorielles » ont des capacités génériques de traitement de l'information

5. Une approche « *top-down* » ou « descendante » consiste à partir d'une vue d'ensemble d'un phénomène, puis à la subdiviser en éléments fonctionnels nécessaires et suffisants pour le reproduire *in silico*. Elle s'oppose à l'approche « *bottom-up* » ou « ascendante » où le point de départ du formalisme est d'emblée le plus détaillé possible sur un plan structural et fonctionnel et à partir duquel il faut remonter les niveaux de description et d'intégration pour en faire émerger les propriétés.

quelle que soit sa nature. C'est notamment le cas du cervelet et des BG.

(ii) Le second biais est tout simplement du à la pertinence du « *population rate-model* » en tant que système capable de restituer de façon fiable l'activité d'un ensemble de neurones. Si nous acceptons comme vrai le fait que le comportement d'un système émerge des interactions de ses unités de traitement fondamentales, l'échelle du modèle doit inévitablement être amenée à l'échelle cellulaire (voire plus loin) afin de fournir une description plus fidèle de ce qui se passe réellement dans un système nerveux biologique⁶.

Ce dernier point a été fortement discuté et demeure encore actuellement très controversé. Une étude récente de [Brette, 2015] traite de ce sujet. L'auteur y explique que la modélisation à l'échelle de la population étant devenu un postulat méthodologique, et même un article de foi de nos jours, les partisans de celui-ci devraient apporter la preuve de son bien-fondé. La stratégie consisterait alors à montrer premièrement sous quelles conditions il serait possible de réduire l'échelle d'un modèle fondé sur les spikes vers un modèle fondé sur les populations (tâche essentiellement théorique), et ensuite à déterminer dans quelle mesure ces conditions sont rencontrées dans le CNS. Dans la même étude, l'auteur ajoute que si ce postulat méthodologique est pratique voire utile, il ne possède pratiquement aucun soutien empirique ou théorique venant l'étayer.

L'importance des instants des spikes dans l'encodage de l'information neuronale (et donc la nécessité d'un modèle décrit à l'échelle de la cellule) est avérée dans d'autres publications récentes. Dans [Portelli et al., 2016], les auteurs démontrent que de simples stratégies de codage comme celles qui s'appuient sur le décompte brut du nombre de spikes (« *spike count-based coding* ») sont insuffisantes pour coder l'information provenant des cellules ganglionnaires de la rétine par rapport aux stratégies qui prennent en compte les instants des spikes (« *spike-timing* »). Ce dernier mode d'encodage s'avère ici indispensable pour faire le lien entre l'information véhiculée et la performance comportementale observée. Dans [Jacobs et al., 2009] les auteurs ont appliqué un mode de décodage bayésien de l'information aux trains de spikes enregistrés. Ils ont montré à cette occasion qu'un codage reposant sur le décompte du nombre total de spikes était là encore insuffisant pour tenir compte de la performance comportementale et que les instants des spikes de cellules individuelles devaient être pris en compte afin d'enrichir l'information produite. Récemment [Saal et al., 2015] a mis en évidence une relation entre les informations véhiculées par des neurones individuels et les instants de leurs spikes. Les auteurs y montrent notamment l'importance de l'échelle de la ms comme base d'encodage temporel de l'information dans le cortex somato-sensoriel : « *precise spike timing conveys information about tactile events and shapes perception* ». Dans une autre étude qui s'intéresse à des structures que nous avons évoquées et donc plus proche encore de nos centres d'intérêt immédiats, [Jason T. Moyer and Wolf, 2014] aborde le rôle des instants des spikes des MSN et des interneurons à décharge rapide (« *Fast-Spiking Interneurons* », FSI). Les auteurs y démontrent leur rôle nécessaire sur la synchronie neuronale et donc sur les propriétés de la dynamique globale du réseau. Dans cette même étude, les auteurs mentionnent également la contribution cruciale des instants des spikes sur les propriétés d'apprentissage. En effet, les instants précis de décharge des MSN véhiculés par l'inhibition latérale jouent un rôle important dans la régulation de la plasticité synaptique,

6. C'est d'ailleurs l'une des raisons qui devrait nous amener autant que possible à nous tourner vers des modèles animaux simples pour en développer des versions *in silico*

particulièrement dans le cas où elle est modulée par la DA et les instants relatifs des spikes pré et post-synaptiques. Même si cette dernière fonctionnalité n'a pas encore été explicitement implémentée dans notre modèle, il est en effet avéré que la STDP modulée par la DA joue un rôle majeur dans la corrélation neuronale de l'apprentissage par renforcement [Chou et al., 2015] et sa prise en compte dans de prochaines études fait partie des perspectives de ce travail.

Les modèles reposant sur les neurones impulsifs (« *spiking-neurons* ») et dont la construction est contrainte par les données anatomiques et physiologiques de la connectivité des structures étudiées peuvent donc apporter des résultats fructueux et particulièrement riches d'enseignement. Ceci est rendu possible grâce au lien direct qui s'établit alors entre la dynamique de la circuiterie détaillée et la fonction qui en émerge [Chersi et al., 2013, Mandali et al., 2015, Medina et al., 2001].

2.4.1.2 Comparaison des mécanismes de sélection de l'action

La sélection d'une action nécessite que, parmi plusieurs alternatives possibles, une seule soit sélectionnée ce qui implique *a priori* l'utilisation d'un réseau de type « *winner-takes-all* ». Il existe au moins deux grandes familles de mécanismes évoqués dans de précédents modèles et qui permettent de produire un tel réseau.

Inhibition latérale Le premier mécanisme permettant cette fonction dans une structure neuronale est l'inhibition latérale. Dans cette situation, tous les neurones sont interconnectés par des collatérales inhibitrices et quand un neurone est activé, il aura alors tendance à inactiver tous les autres. Le striatum est d'un des sites candidats pour un tel mécanisme. Il possède en effet un vaste champ dendritique récurrent associé aux neurones de projections inhibiteurs du striatum, les MSN, qui, potentiellement, innervent des milliers d'autres MSN [Oorschot, 1996]. Cela constituerait un réseau d'inhibition latérale où l'activité d'un neurone dominant émergerait localement au détriment de celle de tous les autres. Beaucoup de modèles computationnels antérieurs ont utilisé explicitement l'inhibition latérale pour expliquer le fonctionnement du striatum ou pour produire une sélection de l'action de type « *winner-takes-all* » [Bar-Gad and Bergman, 2001, Suri and Schultz, 1999, Alexander and Wickens, 1993, Groves, 1983, Kotter and Wickens, 1998, Houk et al., 1994], tandis que d'autres font appel à des algorithmes de type « *winner-takes-all* » mais indépendamment des mécanismes d'inhibition latérale [Frank et al., 2001, Frank, 2005, O'Reilly and Frank, 2006]. Quelqu'en soient les modalités et même s'il est démontré expérimentalement que l'inhibition récurrente est assez faible comparée à l'inhibition « *feedforward* » [Jaeger et al., 1994, Koos et al., 2004, Tunstall et al., 2002], ce phénomène doit donc être pris en considération lorsque l'on étudie la sélection de l'action comme cela a encore récemment été démontré par [Jason T. Moyer and Wolf, 2014] et c'est pour cette raison que nous l'avons inclus dans notre travail.

Propriétés de rétroaction Une alternative au mécanisme précédent est l'utilisation des propriétés de rétroaction du réseau et ce mécanisme a également été pris en compte dans notre étude. Ce mécanisme s'appuie sur la présence de deux voies de rétroaction, passant du cortex aux BG et revenant au cortex [Albin et al., 1989, Leblois et al., 2006b]. L'une de ces voies

procurant une rétroaction positive et l'autre une négative. En effet, d'après cette anatomie, [Mink, 1996] propose un modèle de sélection de l'action schématique dans les BG. En utilisant une telle architecture, Mink a postulé que des réseaux corticaux séparés sont activés pour chacune des actions possibles. La voie hyperdirecte (divergente, de rétroaction négative) procure dans un premier temps une inhibition corticale rapide et généralisée afin de retenir la libération de toutes les actions proposées. La voie directe (rétroaction positive) applique en parallèle une désinhibition plus localisée, amplifiant toute différence dans l'activation entre les réseaux corticaux en compétition afin de sélectionner une action spécifique.

2.4.1.3 Deux niveaux pour la prise de décision et la sélection de l'action

Des travaux antérieurs ont proposé des alternatives à la traditionnelle dichotomie voie directe/-voie indirecte. Dans [Gurney et al., 2001a] les auteurs proposent une alternative sous la forme de deux mécanismes en interaction synergiques modulés par la DA et permettant la prise de décision : l'un de contrôle et l'autre de sélection. Dans l'étude de [Leblois et al., 2006b] les auteurs montrent que l'interaction entre les voies directes et hyper-directes permet naturellement la sélection de l'action grâce à l'action compétitive de boucles cortico-corticales. C'est également dans cette étude qu'il est proposé d'appliquer un seuil à l'activité corticale comme critère permettant d'interpréter une décision et une action. Mais c'est dans le travail de [Guthrie et al., 2013] que cette notion de deux niveaux de prise de décision et de sélection de l'action est formalisée et décrite. Dans notre étude, initialement fondée sur les données expérimentales de [Pasquereau et al., 2007] chez des primates non humains, nous avons séparé la présentation des cibles en deux types de stimuli sensoriels : (i) la forme de la cible et (ii) la direction du mouvement pour la sélection de la cible choisie. Selon cette interprétation, le mécanisme de sélection de l'action nécessite donc deux choix : un pour la forme de la cible et un autre pour sa direction. Quand ces choix sont réussis, nous les avons désigné respectivement dans notre étude par GD et GC. Il est communément admis que les BG contiennent de multiples boucles en parallèle [Alexander et al., 1986]. Nous avons donc implémenté deux modules de sélection de l'action afin de traiter en parallèle les deux niveaux de prise de décision dans notre modèle : le premier servant à sélectionner la forme de la cible (composante cognitive) et le second à sélectionner la direction du mouvement (composante motrice). Il est clair que, pour qu'une telle architecture soit fonctionnelle, une communication doit se mettre en place en quelque point entre ces deux niveaux de décision. En effet tant que les deux boucles de sélection sont simulées strictement en parallèle, il est évident qu'il n'y a pas de convergence ni de divergence (et donc pas de diffusion) possible de l'information dans les BG. Ceci implique d'emblée que ces boucles ne peuvent pas être entièrement distinctes [Alexander and Crutcher., 1990, Graybiel et al., 1994, Parent et al., 2000]. Les interactions entre les boucles cognitives et motrices ont été incluses au niveau striatal dans le modèle en raison des données anatomo-fonctionnelles de la divergence dans les connexions cortico-striatales. Le travail sur cette architecture constitue d'ailleurs l'un des apports du modèle de [Guthrie et al., 2013].

2.4.1.4 Convergence et divergence dans l'architecture des BG

Les neurones corticaux pyramidaux possèdent des projections somatotopiques vers le striatum [Webster, 1961], mais s'étendent également à l'intérieur de ce dernier [Cowan and Wilson, 1994, Parthasarathy et al., 1992, Wilson, 1987]. Cette arborisation striatale se fait à la fois en suivant des aires localisées (spécifiques de la formation du bouton ⁷) et en même temps avec un motif discontinu vers des aires d'innervation plus denses, elles-mêmes alternées avec d'autres aires où l'innervation est plus dispersée [Brown et al., 1998, Flaherty and Graybiel, 1991]. Ce schéma de connectivité produit *de facto* un chevauchement entre les multiples entrées des aires corticales [Takada et al., 2001]. Cette architecture particulière de la voie directe à l'intérieur d'une boucle fait qu'un canal cortical cognitif (ou moteur) présent dans une boucle corticale fermée des BG (rétroaction positive) va être en interaction avec son homologue cognitif (ou moteur) dans une autre structure, dans la même boucle cognitive (ou motrice).

Par ailleurs, la divergence cortico-striatale est suivie par une convergence du striatum associatif vers le pallidum dans ses composantes motrice et cognitive ce qui permet une synthèse et un transfert de l'information d'une boucle à l'autre. Ainsi, la décision prise à un niveau influence graduellement la décision émergente dans l'autre niveau. En pratique dans le cas qui nous intéresse, quand la forme d'une cible est privilégiée, la direction correspondante l'est également grâce aux données anatomiques que nous venons de décrire.

2.4.1.5 Apprentissage dans le réseau

Même avant un apprentissage, notre réseau est capable d'effectuer spontanément la sélection d'une action. Ce processus est dû à une rupture de symétrie (SB), elle-même consécutive au motif de connectivité précédemment décrit et au bruit synaptique appliqué aux cellules. Comme aucun pré-câblage initial n'est favorisé pour aucune des quatre (ou deux) cibles, même avec un taux de SB proche de 90% dans un réseau naïf, les moyennes de GC et GD sont proches de 50% (Figures 2.10 et 2.16). Le modèle permet donc un choix « à l'aveugle » et se trouve dans une phase que l'on pourrait décrire comme purement exploratoire. Le processus d'apprentissage par renforcement prenant place durant une session d'entraînement induit des changements des poids synaptiques des projections du cortex cognitif vers le striatum, et est modulé par les variations phasiques des concentrations en DA, nécessaires pour les mécanismes de LTP et LTD [Pawlak and Kerr, 2008, Kerr and Wickens, 2001, Reynolds et al., 2001]. Les valeurs des paramètres que nous avons ajustées pour les règles de plasticité synaptique permettent de retrouver qualitativement et quantitativement les modalités d'apprentissages observées expérimentalement, bien entendu au prix d'une certaine simplification. En effet il peut y avoir (et il y a sans doute) des possibilités d'apprentissage et de plasticité synaptique dans d'autres structures et dans d'autres voies à l'intérieur des BG. Toutefois l'un des objectifs initiaux de cette étude était de confirmer que la plasticité des projections du cortex cognitif vers le striatum était suffisante pour produire elle-même les comportements observés ce qui a pu être montré ici.

7. Un bouton synaptique représente l'extrémité terminale de l'arborisation de l'axone d'un neurone (Figure 1.16).

2.4.2 Limites et perspectives de l'approche

Notre modèle computationnel des BG possède un formalisme décrit à l'échelle neuronale et est capable d'exécuter une sélection de l'action reposant sur les niveaux de description cognitif et moteur. Il est étroitement lié à l'anatomie et la physiologie décrites pour ces régions et permet de montrer un mécanisme plausible de sélection de l'action dans le réseau en étendant les travaux de [Leblois et al., 2006a] et [Guthrie et al., 2013]. De plus, ce modèle confirme que si la sélection de l'action est permise par l'effet du bruit à l'intérieur du réseau avant apprentissage (phase d'exploration), après apprentissage en revanche l'effet ce bruit est en grande partie atténué et le modèle devient capable de produire des performances de sélection de l'action à un niveau optimal (phase d'exploitation).

Cependant, l'utilisation de l'échelle neuronale elle-même impose des limites à notre approche. En effet, lors de la construction d'un modèle de réseau, les limites de ce dernier sont imposées par le modèle neuronal choisi qui est dans ce cas le facteur limitant. Notre choix d'un modèle LIF à une dimension le rend inapte à reproduire toutes les propriétés cellulaires intrinsèques (ICP) des neurones biologiques. Par exemple des activités en « *burst* »⁸, les effets de rebond ou d'adaptation sont impossibles à reproduire avec ce modèle simplifié [Izhikevich, 2000]. Très récemment, l'utilisation de modèles de « *spiking neurons* » présentant des « *patterns* » de décharge plus complexes que les LIF a été mise à profit dans une étude portant sur les structures des BG mais abordant une problématique différente de la nôtre [Wei et al., 2015]. Nous avons pleinement conscience qu'une grande partie de l'information sur les propriétés dynamiques des neurones réellement actifs dans ces processus nous est pour le moment inaccessible. C'est pourquoi nous avons fait le choix volontaire de travailler sur des populations de neurones homogènes afin de nous focaliser sur les effets de réseau dans un premier temps mais notre implémentation permet également d'inclure des modèles neuronaux à deux dimensions dans notre réseau et donc par la suite de tester le rôle des ICP sur sa dynamique et ses capacités.

L'apprentissage dans notre modèle se situe uniquement au niveau cortico-striatal cognitif et dans de futurs travaux il serait pertinent d'ajouter ce mécanisme dans d'autres connexions au sein des BG. De plus, la STDP n'est pas encore explicitement implémentée dans notre modèle. Un de nos objectifs proches serait donc d'ajouter le mécanisme de STDP modulée par la DA pour les propriétés d'apprentissage et de régulation de la plasticité synaptique dans notre modèle.

La voie indirecte n'a pas été prise en compte dans notre modèle et nous avons déjà argumenté ce choix (cf. 1.1.2.2), par ailleurs discutable. C'est souvent dans le parti-pris d'inclure ou non certains éléments expérimentaux que réside l'art du modélisateur. Dans l'étude de [Schroll et al., 2012] par exemple, seule la voie directe est prise en compte mais c'est parce que dans le contexte de l'étude elle suffit à rendre compte par elle-même des mécanismes étudiés. Les auteurs y ont fait le choix, comme nous, de suivre le principe de parcimonie⁹ et donc de limiter leur étude aux régions impliquées qui à leurs yeux étaient nécessaires et suffisantes pour décrire les phénomènes observés. De plus il n'est pas impossible que dans la mise au point de notre modèle, le trajet hyperdirect prenne en charge partiellement l'effet supposé du trajet indirect dans la mesure où les deux boucles ont un effet globalement inhibiteur. Ce phénomène de « masquage » de l'effet

8. Alternance entre un état de repos et une série de spikes à haute fréquence

9. Egalement connu sous le nom de rasoir d'Ockham

d'une structure par une autre ne remet pas nécessairement en question la pertinence globale de notre étude mais nous rappelle que tant que toutes les pièces du puzzle ne sont pas assemblées, nous n'avons accès qu'à une partie du tableau que nous cherchons à découvrir et qu'une part de spéculation est indissociable de nos conclusions indépendamment de la rigueur de notre travail.

Nous n'avons hélas pas encore eu le temps de nous pencher sur les effets de synergie produits par les interactions entre les structures connectées aux BG. Le fait est que la tâche est rendue d'autant plus ardue que leur situation centrale dans le CNS les positionne au carrefour de presque toutes les autres régions notables (hippocampe, cervelet ...) et par ailleurs nous avons attribué un rôle minime au thalamus et au cortex dans le traitement de l'information ce qui constitue bien évidemment une abstraction importante. Notre approche « BGcentrée » était inévitable mais nous n'avons pas perdu de vue qu'elle demeurerait intrinsèquement déséquilibrée. Nous avons mentionné à plusieurs reprises ces omissions (cf. sections 2.3.3.2 et 2.3.3.3) mais en l'état elles ne pourront être comblées que progressivement à l'aide d'études complémentaires.

Nous clôturerons ce chapitre en rappelant que le niveau de description que nous avons choisi permet néanmoins de combler en partie le fossé entre les approches « *top-down* » (niveau de description mésoscopique [Guthrie et al., 2013]) et « *bottom-up* » en nous appuyant sur les propriétés émergentes de la dynamique des SNN. Nous avons en effet en tête que les moyens mis en œuvre pour acquérir la connaissance contribuent à la connaissance elle-même. Notre modèle s'est montré capable de prédire d'importantes caractéristiques comportementales comme la conséquence de lésions localisées sur la déficience de l'apprentissage, son inversion et sur les protocoles d'extinction. À ce titre il nous semble qu'il constitue une étape aboutie dans la compréhension du fonctionnement des ganglions de la base mais qu'il demeure encore une importante marge de progression pour qu'un portrait global réellement fidèle puisse en être esquissé.

Mise au point d'un protocole de prise de décision chez *Pleurodeles waltlii* en condition physiologique et à l'aide d'un modèle parkinsonien

« Je croyais être indécis mais je n'en suis pas certain. »

Robert Bourassa

Sommaire

3.1	Introduction	92
3.2	Choix d'un modèle animal : l'urodèle	92
3.3	Matériel et méthodes	93
3.3.1	Dispositif d'étude comportementale	93
3.3.2	Métriques et études des effets des lésions dopaminergiques	97
3.3.3	Immunohistochimie	100
3.4	Résultats	102
3.4.1	Protocole de prise de décision	102
3.4.2	Lésions dopaminergiques	105
3.5	Discussion	109
3.5.1	Synthèse des travaux	109
3.5.2	Points sensibles des méthodes d'analyse	109
3.5.3	Perspectives	110

3.1 Introduction

Ce travail s'est effectué en parallèle du développement du modèle computationnel. D'une part, nous avons contribué au développement d'un protocole novateur de prise de décision chez l'urodèle, notamment en apportant des éléments d'analyse. D'autre part nous avons mis en place, au moins partiellement, un modèle parkinsonien chez ce même animal. Dans cette partie, nous verrons en détails les raisons pour lesquelles nous avons choisi cet animal et nous présenterons nos résultats préliminaires.

Les parties faisant référence au conditionnement et au test comportemental (sections 3.3.1.2, 3.3.1.3 et 3.3.1.4) ainsi que les résultats associés font référence au travail réalisé par Manon BONNET-SAVE lors de ses stage de Master 1¹ et 2². La présentation de ces travaux initiés par le Dr. Thomas BORAUD et le Pr. Jean-Marie CABELGUEN sont nécessaires à l'introduction des contributions qui ont ensuite été réalisées dans le cadre de cette thèse.

3.2 Choix d'un modèle animal : l'urodèle

Le point de départ de cette étude a été l'exploration des mécanismes élémentaires de la prise de décision par le biais d'une approche comportementale chez un vertébré inférieur. En effet, chez les premiers vertébrés, nous pouvons déjà observer les structures élémentaires qui constituent le substrat neurobiologique impliqué dans la prise de décision et dont les BG font partie [Grillner et al., 2013, Reiner et al., 1998]. L'hypothèse de travail est que ces structures qui se sont développées au cours de l'évolution par des processus de sélection, fonctionnent selon des principes similaires (avec un degré de complexité supplémentaire) à ceux des vertébrés dit supérieurs tels que les primates. Le choix du modèle animal s'est donc porté sur l'urodèles car celui-ci se prête bien à l'évaluation de cette hypothèse [Stephenson-Jone et al., 2013]. Les urodèles ont la particularité d'avoir une organisation des BG assez similaire à celle des mammifères mais avec un nombre de neurones nettement inférieur. Cette particularité réduit sensiblement la complexité de l'étude des processus de prise de décision dans leur structure cérébrale. En effet, le cortex n'existe pas encore à proprement parler chez ces amphibiens, mais ils possèdent un pallium (cf. section 1.1.3.3) dont l'organisation neuronale et l'anatomie sont analogues bien que plus sommaires [Nieuwenhuys et al., 1998]. Pour des raisons similaires, les agnates comme la lamproie ont été le sujet de nombreuses études concernant la prise de décision et et même de modèles computationnels [Sarvestani et al., 2013]. Toutefois les tests comportementaux disponibles sur cet animal sont marginaux ou inexistant à ce jour, contrairement à ce que l'on peut envisager chez les amphibiens. Bien que les urodèles soient des animaux assez primitifs, ils semblent tout de même capables d'apprentissages exploitables pour notre thématique [Douglas and Adler, 1973].

Nous avons donc choisi de travailler avec le Pleurodèle de Walzl pour sa situation phy-

1. BONNET-SAVE Manon. Développement d'un conditionnement opérant chez le pleurodèle de Walzl. *Master 1 Biologie Santé, Spécialité Neurosciences et Neurophychopharmacologie. Université de Bordeaux.* Mai 2014

2. BONNET-SAVE Manon. Implication des Ganglions de la Base dans un processus d'apprentissage : Mise en place d'un modèle chez l'urodèle. *Master 2 Biologie Santé, Spécialité Neurosciences et Neurophychopharmacologie. Université de Bordeaux.* Juin 2015

logénétique au sein des vertébrés inférieurs, faisant ainsi le choix d'un niveau de complexité anatomique moindre mais suffisant. Encore une fois le principe de parcimonie nous a guidé. De plus, ces animaux possèdent de fortes capacités de régénération, capacités utiles lorsque l'on envisage des études de plasticité post-lésionnelle.

3.3 Matériel et méthodes

Nous voulions que le conditionnement des animaux soit le plus simple possible, avec un minimum de degrés de liberté afin de restreindre les choix envisageables et donc de rester le plus proche possible de la tâche comportementale de [Pasquereau et al., 2007]. Une première adaptation d'un protocole habituellement réalisé chez le rongeur (« Labyrinthe en Y », [Retailleau et al., 2013]) a été testée, sans succès. En effet, les différences de comportement et de conditionnement sont trop importantes entre ces espèces. Nous avons alors simplifié l'environnement d'entraînement et donc les modalités du processus de prise de décision afin de mettre en place un protocole d'évaluation du choix avec ce modèle animal.

À l'état naturel, les urodèles peuvent s'orienter grâce au changement circadien de la polarisation de la lumière solaire [Phillips, 1977, Freake and Phillips, 2005, Schlegel et al., 2006]. Un conditionnement opérant a donc été mis en place en utilisant les capacités naturelles de ces animaux à s'orienter en fonction de la lumière polarisée. Le conditionnement opérant permet d'associer un stimulus neutre à une valeur spécifique par le biais de l'apprentissage faisant suite à l'exploration déclenchée par l'animal. Lorsque ce stimulus va être présenté au sujet, ce dernier aura le choix entre plusieurs actions possibles, dont une sera spécifiquement renforcée par rapport aux autres. Ceci rend possible l'étude des structures impliquées dans la sélection de l'action chez ces animaux. Ce mécanisme a été simplifié au maximum pour aboutir à un choix entre deux directions dont les valeurs respectives seront apprises grâce à deux conditionnements opérants différents. Les mécanismes de sélection de l'action en renforcement négatif étant moins étudiés pour des raisons éthiques, le protocole présenté ici se limite uniquement au conditionnement positif pour la tâche de prise de décision. C'est du reste ce type de conditionnement qui avait été mis en place chez les primates [Pasquereau et al., 2007, Piron et al., 2016]. La mise en place de cet apprentissage chez les urodèles permettrait donc théoriquement d'étudier aussi l'implication des BG dans ce type de tâche.

Enfin, une approche lésionnelle a été utilisée afin de déterminer le degré d'implication de chaque structure d'intérêt des BG dans le processus de prise de décision. Pour cela, nous nous sommes inspirés d'une précédente étude [Parish et al., 2007] portant sur le même thème et utilisant sur un modèle animal très proche (voir la section 3.3.2.3 pour plus de détails).

3.3.1 Dispositif d'étude comportementale

3.3.1.1 Animaux

Pour le test comportemental, 20 urodèles (*Pleurodeles waltlii*) juvéniles (12 à 18 mois post-éclosion) provenant du fournisseur « Reptile Home » ont été utilisés. Le groupe contrôle dit « Naïfs » est constitué des animaux avant entraînement et le groupe « Entraînés » est constitué

des mêmes animaux, après entraînement. Pour les lésions dopaminergiques, 6 autres urodèles juvéniles provenant du même fournisseur ont été utilisés.

Conditions d'hébergement dans l'animalerie Les animaux ont été hébergés dans des aquariums/bacs remplis d'eau, complétée par quelques gouttes d'antifongique (Bleu de Méthylène 2.5 %), dans une salle maintenue à 18°C avec un cycle de 12h de lumière par jour. Les bacs étaient nettoyés une fois par semaine et les pleurodèles nourris deux fois par semaine (chironomes et/ou cœur de veau). Dans les bacs, les sexes étaient mélangés mais des fourchettes d'âge étaient respectées.

Conditions d'hébergement spécifiques lors de l'expérimentation Pour le test comportemental, les conditions de rétention ont été les mêmes pendant la phase d'entraînement et la phase de test. Les sujets ont été maintenus dans le bassin d'entraînement (voir section 3.3.1.2) en chambre noire et selon les conditions décrites précédemment. Ils ont été nourris deux fois par semaine après leur second entraînement du jour et la nourriture était déposée dans la zone de la cachette pour renforcer son attrait.

Conditions d'hébergement spécifiques post-lésionnelles Après chirurgie, les animaux lésés étaient placés dans des bacs isolés, avec un fond d'eau (complété par du Bleu de Méthylène) permettant également de vérifier que celle-ci ne pénétrait pas dans le crâne le temps de la cicatrisation.

3.3.1.2 Bassin d'entraînement

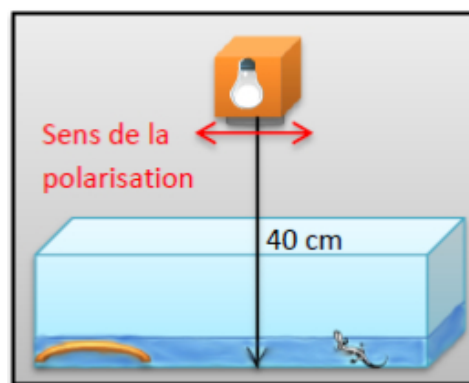


FIGURE 3.1 – Schéma du bassin d'entraînement. La partie bleu foncé en bas de la figure représente le niveau d'eau et la forme orange en bas à gauche la tuile servant d'abri à la salamandre.

Les animaux ont été entraînés dans l'aquarium qui leur a servi de lieu de vie durant toute la période d'expérimentation. Il s'agissait d'un bac rectangulaire en plastique gris (60x40 cm) avec une hauteur d'eau de 10 cm. Dans cet aquarium, un abri était présent sous la forme d'une tuile (16x19 cm) dans un coin, soit à 45° du vecteur de la lumière polarisée. L'installation était maintenue dans une chambre noire avec un éclairage de 12h par lumière polarisée, dont le vecteur était orienté dans la direction de la longueur de l'aquarium (Figure 3.1). La lumière polarisée était obtenue en illuminant un filtre polarisant (72 mm de diamètre) par une ampoule (20W, 235

lm) contenue dans un cube en carton de 12,5 cm de côté afin que tous les rayons sortants passent obligatoirement par le filtre.

3.3.1.3 Arène

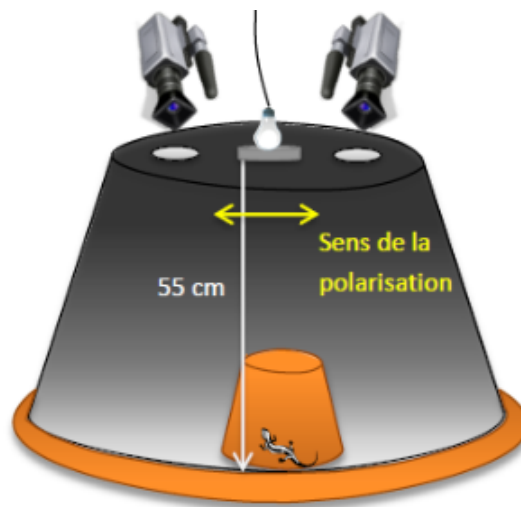


FIGURE 3.2 – Schéma de l'arène.

Une arène circulaire en plastique couleur terre de 90 cm de diamètre (75 cm à sa base), remplie avec 5 cm d'eau et recouverte par une chambre noire de 75 cm de diamètre à sa base ont été utilisées pour cette installation (Figure 3.2). Le dispositif lumineux utilisé était pratiquement identique à celui employé lors de l'entraînement et était disposé sur le plafond de la chambre noire, soit à 50 cm de la surface de l'eau. Le filtre étant disposé à même le toit, tous les rayons entrant dans l'arène étaient polarisés. Deux caméras disposées à l'opposé, chacune au dessus d'un orifice ont permis de visualiser ce qui se passait à l'intérieur de l'arène et d'enregistrer les déplacements des pleurodèles.

3.3.1.4 Test comportemental

Tous les tests ont été réalisés dans une salle spécifique, dans l'obscurité et à une température de 18°C. Ce test a été inspiré par les travaux de Douglas & Adler [Douglas and Adler, 1973] afin de démontrer que les salamandres utilisaient la polarisation de la lumière pour s'orienter.

Conditions d'entraînement Les animaux ont été retirés de leur aquarium et placés provisoirement dans un bac contenant 3 à 5 cm d'eau, le tout effectué à l'obscurité. Ils ont ensuite été ajoutés un par un (délai de 1 minute entre chaque animal) au centre du bassin d'entraînement, sous la lumière polarisée, afin de pouvoir retourner à leur abri. Les animaux ayant été entraînés en premier ont été maintenus dans l'aquarium pour l'ajout des suivants (augmentant ainsi la valeur positive de l'abri en favorisant le grégarisme). Cette phase a duré 21 jours, avec 4 entraînements quotidiens étalés dans la journée pour chaque animal. Lors de l'entraînement, le vecteur de polarisation de la lumière était perpendiculaire à celui du champ magnétique terrestre (axe Nord-Sud) et repéré au moyen d'une boussole. Après chaque session et plusieurs fois par jour,

les nombres d'animaux sous, sur et à côté de l'abri ont été dénombrés afin d'obtenir un suivi du comportement des pleurodèles dans l'aquarium.

Test en arène À la fin de la période d'entraînement, la phase de test des salamandres a débuté en maintenant leur entraînement. Les animaux ont été testés un par un dans une salle adjacente à celle de leur lieu de vie. Le transfert entre les deux salles s'est fait individuellement dans un bac et à l'obscurité. Les animaux ont été placés au centre de l'arène circulaire sous un pot opaque de 16 cm de diamètre durant 30 secondes, puis le pot a été enlevé afin de laisser l'animal libre de ses mouvements. Le test était terminé dès que le pleurodèle touchait la paroi de la chambre noire ou au bout de 15 minutes passées dans l'arène. La position atteinte (angle en degrés) ainsi que le temps mis pour l'atteindre ont été relevés. Une réussite de l'animal dans ce test consistait à toucher la paroi de l'un des deux cadrans se trouvant dans la direction hypothétique de la cachette indiquée par le vecteur de polarisation de la lumière (Figure 3.3).

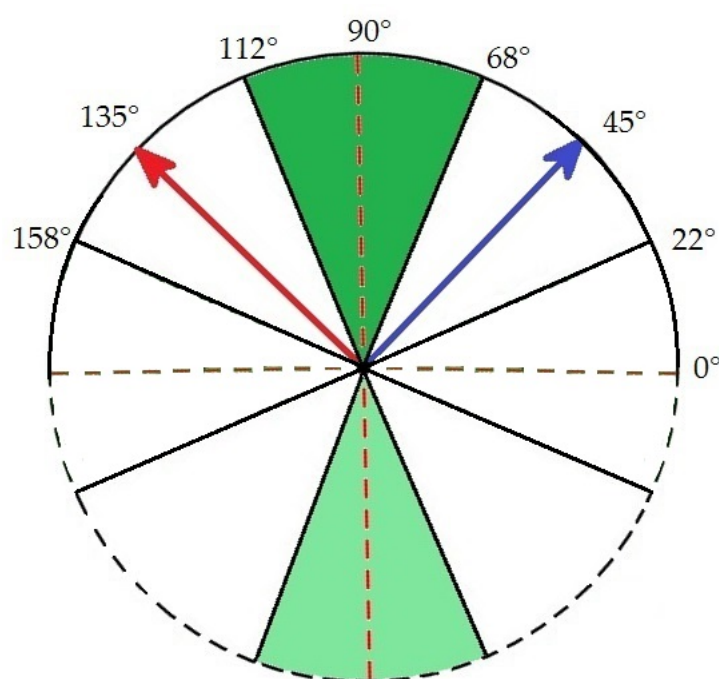


FIGURE 3.3 – Arène vue du dessus. Les deux flèches représentent les deux vecteurs, bleu pour la lumière polarisée (ici à 45°) et rouge pour le champ magnétique terrestre (axe Nord/Sud, ici à 135°). La zone considérée comme « positive » est représentée par les cadrans verts (de 68° à 112°). Ces cadrans sont superposés par symétrie centrale pour obtenir un seul jeu de résultats. La lumière polarisée indiquant une direction et non un sens, le fait de rassembler les cadrans de la sorte permet de n'avoir aucune perte de données.

Une fois le test terminé, l'animal était remplacé dans son lieu de vie (l'aquarium). L'eau de l'arène était agitée entre chaque animal afin d'éviter qu'une piste olfactive laissée par un animal donné puisse influencer le comportement des suivants.

Conditions testées Les effets de la lumière polarisée, de l'entraînement et du champ magnétique terrestre ont été évalués dans les différentes conditions présentées dans la Table 3.1. Chaque condition, de même que chaque contrôle, a fait l'objet de trois sessions individuelles.

TABLE 3.1 – *Présentation des conditions testées*

	Vecteurs Parallèles		Vecteurs Perpendiculaires	
Lumière Non Polarisée	Entraînés	Naïfs	Entraînés	Naïfs
Lumière Polarisée	Entraînés	Naïfs	Entraînés	Naïfs

3.3.2 Métriques et études des effets des lésions dopaminergiques

3.3.2.1 Enregistrements vidéo

Tests en arène Les résultats des tests comportementaux en arène ont été collectés par enregistrement vidéo afin que le déplacement des animaux soit suivi en temps réel. Les caméras utilisées sont des caméras CCD avec un zoom digital 27x. Elles étaient reliées à un connecteur USB à 4 canaux, permettant de numériser l'image analogique. Le logiciel de capture vidéo SuperViewer (compatible avec Windows XP) a été utilisé pour traiter les vidéos. Les données ont été regroupées dans des fichiers Microsoft Excel et des fichiers texte.

Tests post-lésionnels La caméra utilisée pour filmer le comportement des animaux pendant leur récupération post-opératoire était la même que lors des tests en arène. Les urodèles ont été filmés individuellement, dans un bac identique à ceux qui leurs servaient de lieu de vie à la différence près qu'un quadrillage avait été tracé au fond du bac (5x5 cm) avec un marqueur indélébile afin de pouvoir éventuellement mesurer la distance parcourue. La caméra était placée au dessus du bac grâce à un trépied de façon à le filmer en entier sur les vidéos.

3.3.2.2 Traitement des données

Tests en arène : mesure des angles Chaque position était relevée et indiquée sur une représentation schématique de l'arène en vue aérienne (voir Figure 3.2). La zone considérée comme positive (signifiant que les salamandres allaient dans la « bonne » direction et donc que l'apprentissage avait fonctionné) était comprise entre 68° et 112°, la direction du vecteur de polarisation de la lumière se situant à 45°. Chaque moitié de l'arène a fait l'objet de cette mesure afin d'obtenir des valeurs comprises uniquement entre 0° et 180°, de façon à faciliter les analyses ultérieurement.

Tests en arène : analyse statistiques L'analyse statistique a été réalisée sous Prism 6 pour les délais et la phase d'entraînement, et sous MATLAB pour les angles. Les figures ont été réalisées sous Microsoft Excel, Paint, MATLAB et Oriana.

Pour l'analyse des délais de réponse lors de la phase d'entraînement, les différences ont été considérées comme significatives lorsque la *p-value* était inférieure à 5% ($p < 0.05$). L'influence du jour et de l'heure lors de l'entraînement ont été testées à l'aide d'une ANOVA à une voie pour mesures répétées (jour x résultat et heure x résultat) suivie d'un test de Student. Les mêmes tests statistiques ont été utilisés pour comparer les délais des différents groupes.

Afin de définir le nombre de pics dans la distribution des positions des animaux, une analyse mathématique en deux étapes sous MATLAB [Ryczko et al., 2015] a été réalisée. Dans un premier temps, des modèles des distributions possibles ont été générés depuis les données brutes, comprenant un, deux ou trois pics. Pour cela, des modèles de distribution gaussienne comprenant k pics ont été construits en utilisant un algorithme de maximisation d'attente de résultats (*gmdistribution.fit* dans MATLAB). Dans un second temps, le modèle le plus proche des données brutes parmi les modèles de distribution possibles a été identifié en utilisant le critère d'information Akaike (*aic* dans MATLAB). Ce critère est une mesure de la qualité du modèle permettant la quantification de la perte d'information lors d'une comparaison entre le modèle et les données brutes réelles : meilleur est le modèle, plus faible sera la perte d'information et plus petite sera la valeur de l'*aic*. Pour illustrer les distributions d'angles correspondant aux positions des animaux dans l'arène, une estimation de la densité de probabilité a été calculée à partir des données en utilisant une fonction de lissage kernel (*ksdensity* dans MATLAB).

Les pics ont ensuite été comparés entre eux au moyen d'un test de Kruskal-Wallis (équivalent non-paramétrique d'une ANOVA à une voie pour échantillons indépendants) suivi d'une méthode de Dunn pour le post-hoc. Une différence était déclarée significative si $p < 0.05$. Pour démontrer l'effet de l'entraînement, les effectifs d'animaux par déciles des valeurs d'angle mesurées ont été calculés. Il a ensuite été possible de construire un diagramme d'effectifs par position. Après avoir déterminé la moyenne et l'écart-type, un postulat a été établi : si des valeurs présentaient un effectif supérieur à la moyenne + 3 écart-types, alors ces positions sortaient de la norme avec un risque d'erreur de 5% de chances de se tromper et les différences observées n'étaient donc pas considérées comme étant dues au hasard.

3.3.2.3 Lésions dopaminergiques

Dans cette partie nous allons présenter le protocole lésionnel mis en place, inspiré des travaux de [Parish et al., 2007].

Anesthésie et chirurgie Les animaux ont été profondément anesthésiés par immersion dans un bain à 0.1% de MS-222 (Méthyl-Sulfonate de Tricaine) durant une heure. Ces urodèles étant des animaux semi-aquatiques, nous avons maintenu les corps des animaux humides en les recouvrant d'un papier absorbant imbibé par une solution de Ringer (pH 7.4, 4°C). Ensuite, une craniectomie a été effectuée à la jonction du pariétal et du frontal sur la ligne médiane du crâne afin de pouvoir accéder au troisième ventricule lors de l'injection.

Lésions Une solution de 6-OHDA à 6 $\mu\text{g}/\mu\text{l}$ avec de l'acide ascorbique (0.2mg/ml) a été préparée et réservée dans un réfrigérateur à 4°C en attente du moment de l'injection. Ensuite, nous avons utilisé une micropipette en verre couplée par un tube à une seringue Hamilton de 2 μl , le tout monté sur un cadre stéréotaxique. La micropipette contenant la 6-OHDA a été insérée directement à l'intérieur du troisième ventricule afin que la toxine (400 nl) puisse atteindre les neurones dopaminergiques diencéphaliques/mésencéphaliques (voir Figure 3.4).

Afin de compléter l'injection, l'aiguille était laissée telle quelle *in situ* pendant une minute avant d'être lentement retirée. Enfin, l'ouverture du crâne était recouverte d'un coton imbibé

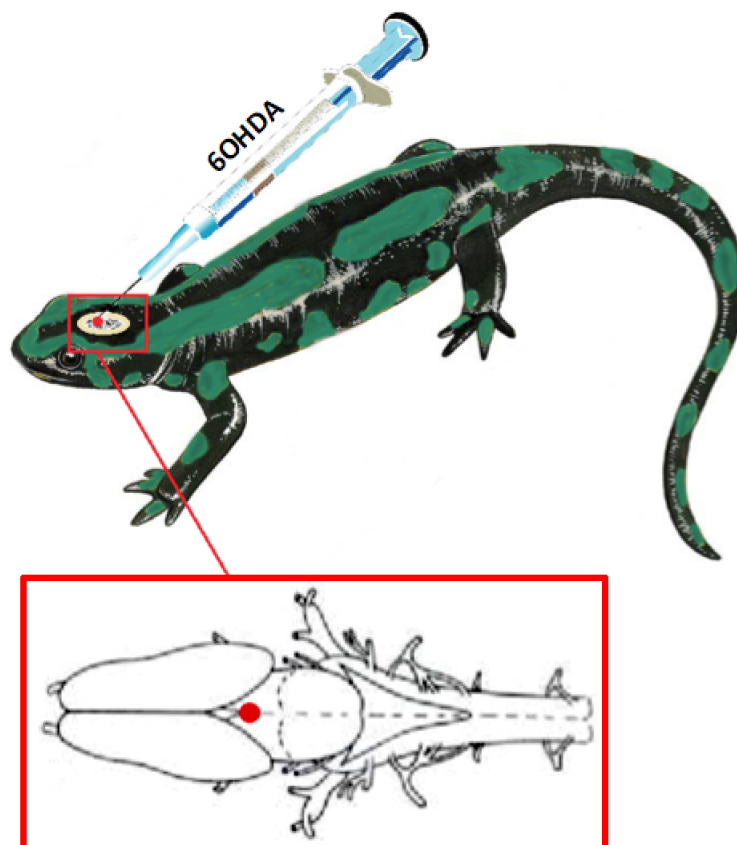


FIGURE 3.4 – Lésions à la 6-OHDA au niveau des neurones dopaminergiques chez la salamandre, d'après [Parish et al., 2007, Marín et al., 1997a, Marín et al., 1997b]. Le point rouge indique le site d'injection de la 6-OHDA, directement à l'intérieur du troisième ventricule au niveau des neurones dopaminergiques diencéphaliques/mésencéphaliques.

d'huile de paraffine afin d'éviter tout dessèchement et l'animal était placé dans un bac isolé avec un fond d'eau le temps de sa récupération.

Tests post-lésionnels Les urodèles ont été filmés en train de marcher et/ou de nager pendant les jours suivant l'opération afin de constater les effets moteurs de la lésion dopaminergique ainsi que l'évolution de la récupération de leurs capacités. Les résultats présentés pour ce travail sont préliminaires et ne concernent que deux animaux. La suite et la validation complète des travaux n'ayant malheureusement pas pu être achevée pour des raisons logistiques.

Pour étudier la marche, les animaux ont été filmés le jour précédent la lésion (J-1) et à J+1, J+3, J+7 et J+10 post-lésionnel. Ensuite, le nombre moyen de pas ainsi que leur durée moyenne ont été analysés. Un « pas » commence au moment où une patte de l'animal quitte le sol et s'achève lorsque cette même patte touche le sol à nouveau, et cela dans le cadre d'une marche continue de la part de l'animal (au moins 3 pas consécutifs). Pour quantifier cette métrique, les vidéos ont été découpées en images en utilisant l'application de lecture et d'encodage de vidéos FFmpeg³. Nous avons ensuite compté manuellement le nombre d'images pas par pas et, avec l'aide d'un script R réalisé par nos soins, nous avons déterminé la durée moyenne d'un pas ainsi que le nombre de pas moyens pour un animal donné.

3. <https://www.ffmpeg.org>

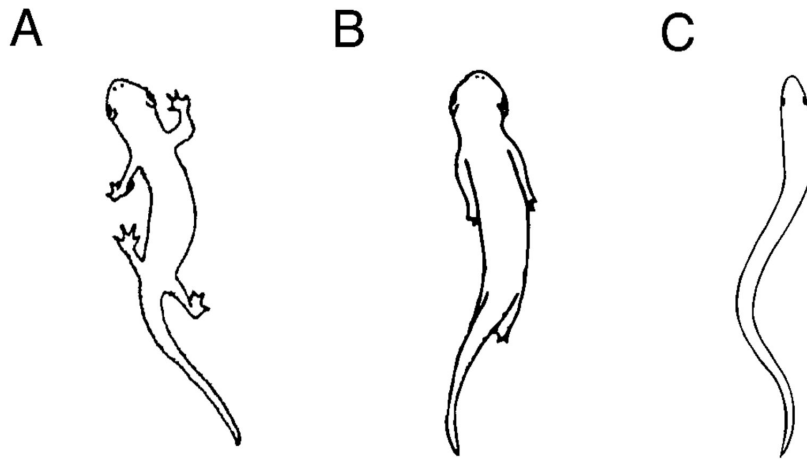


FIGURE 3.5 – Modes de locomotion comparés du pleurodèle et de la lamproie, d’après [Delvolvé et al., 1997]. (A) Position lors de la marche ; (B) Position lors de la nage où les pattes sont repliées le long du corps ; (C) Schéma d’une lamproie observée lors de la nage pour comparer avec B.

Pour la nage, les mêmes animaux ont été filmés à J-1, J+7 et J+13. Les tests pour la nage n’ont été réalisés qu’une semaine après l’opération car il était risqué de les immerger dans l’eau à cause de la cicatrice de la craniectomie. Cependant, ces urodèles étaient en mesure de nager au bout d’une semaine grâce à leurs grandes capacités de régénération. Lors de la nage, les salamandres replient leurs pattes et nagent en ondulant. Leurs cycles de nage ressemblent à ceux de la lamproie (comme illustré sur la Figure 3.5). Nous avons donc analysé, de la même façon que pour la marche, le nombre moyen de cycles de nage ainsi que leur durée moyenne. Notons ici que pour qu’un cycle de nage soit pris en compte, il fallait que l’animal en réalise au moins trois consécutifs lors de la nage.

3.3.3 Immunohistochimie

Durant les enregistrements vidéo, nous avons remarqué que les animaux récupéraient assez bien, voire très bien, leurs capacités motrices. Nous avons donc décidé d’effectuer des marquages par immunohistochimie grâce aux protocoles décrits dans les travaux de [Parish et al., 2007] pour les neurones concernés. Nous avons effectué ces marquages sur des animaux intacts servant de témoins, et sur les animaux ayant été lésés (1 et 2 mois après l’opération).

Prélèvement des cerveaux et fixation Les urodèles ont été anesthésiés de la même façon que pour les lésions. Ils ont ensuite été perfusés directement à travers l’oreillette gauche du cœur par une solution de Ringer, suivie d’une solution fixative (ParaFormaldéhyde - PAF - à 4% dans une solution de « Phosphate Buffered Saline » - PBS - à 1M) pour une meilleure qualité des prélèvements. La fixation a servi à protéger les tissus prélevés de toute hydrolyse due à la libération des enzymes contenus dans les lysosomes cellulaires. Une fois prélevés, les cerveaux ont été placés dans la solution fixative durant trois heures à température ambiante, puis transférés dans une solution tampon PBS 0.1M / Saccharose 15% / Azide de sodium 0.01% à 4°C pour la nuit. À cause de la toxicité du PAF, cette étape de fixation s’est effectuée dans un pièce ventilée dédiée.

Congélation Cette technique nous a permis de solidifier nos tissus afin de réaliser par la suite des coupes très fines au cryostat. Les cerveaux des urodèles étant de très petite taille comparée à ceux des rongeurs et primates, il a fallu adapter les protocoles généralement utilisés au laboratoire. Nous avons donc commencé par préparer des moules en silicone adaptés à nos échantillons. Nous les avons ensuite rempli de moitié par du Tissue-Tek (colle OCT⁴) avant de placer un cerveau dans chacun, en veillant à garder la même orientation pour la reproductibilité du sens des coupes. Le tout a ensuite été recouvert de Tissue-Tek en évitant la formation de bulles d'air pouvant causer des artéfacts lors de l'analyse des images au microscope. Enfin, nous avons placé nos blocs remplis dans de l'isopentane à -45°C pendant 1h, puis nous les avons démoulés, emballés dans une feuille d'aluminium et identifiés avant de les placer dans un congélateur à -80°C en attente du moment de la découpe.

Coupes au cryostat Nous avons effectué des coupes coronales sériées (séries 1 : 5), d'une épaisseur de 20 μm avec un cryostat à -20°C (Leica CM 3050S n°1163710, série : 5043). Les coupes ont ensuite été récupérées sur des lames Super-Frost gélatinées et conservées à -80°C jusqu'à utilisation pour le marquage.

Immunofluorescence sur coupes Pour identifier les neurones dopaminergiques, nous avons choisi de réaliser un marquage contre la tyrosine hydroxylase (TH - un des enzymes de synthèse de la dopamine).

Pour effectuer ce marquage, le premier jour, les coupes ont tout d'abord été rincées avec du PBS puis incubées avec le tampon de saturation pendant une demi-heure. Les lames ont ensuite été incubées avec la solution contenant l'anticorps primaire monoclonal de souris (MAB318, Millipore) à une dilution de 1/5000 sur la nuit à 4°C dans une chambre humide.

Le second jour, les lames ont été rincées abondamment au PBS (deux fois cinq minutes, deux fois huit minutes et une fois dix minutes) pour éliminer l'anticorps primaire non fixé. Ensuite, les coupes ont été incubées avec l'anticorps secondaire anti-espèce (chèvre anti-souris) couplé à une fluorophore (Alexa568, A11004, Invitrogen) à une dilution de 1/2000 pendant une heure à température ambiante, dans la solution d'incubation. Les coupes ont ensuite été rincées deux fois cinq minutes au PBS. Un marquage nucléaire au 4',6-diamidino-2-phénylindole (DAPI, Hoechst33258, Invitrogen) a ensuite été réalisé (dilution : 1/2000, incubation : deux minutes à température ambiante). Enfin, trois lavages au PBS ont été faits avant de mettre la lame sous lamelle (milieu de montage : DAKO fluorescent mounting medium S3023, DAKO).

Révélation au microscope à fluorescence Les observations et acquisitions ont été réalisées au microscope confocal à fluorescence (SP8, Leica) au sein de la plateforme d'imagerie de l'Université de Bordeaux (BIC). Pour les étapes de post-traitement et de création de la planche présentée dans les résultats, nous avons utilisé ImageJ (NIH, Bethesda), et le plugin FigureJ [Mutterer and Zinck, 2013].

4. « *Optimum Cutting Temperature* »

3.4 Résultats

Les résultats préliminaires présentés ici ont fait l'objet d'un poster présenté lors du « *Fifth Symposium on Biology of Decision Making* » en mai 2015 (voir Annexe A.3).

3.4.1 Protocole de prise de décision

3.4.1.1 Entraînement

L'analyse statistique des résultats obtenus n'a malheureusement pas pu mettre en évidence de corrélation entre l'heure de la journée et le nombre de pleurodèles à chaque emplacement c'est-à-dire sur, sous ou à côté de l'abri (ANOVA à une voie, $p > 0,05$). Cependant, les observations tendent à montrer, notamment au début de la phase d'entraînement, que le nombre d'animaux à proximité directe de l'abri (dessous et dessus) augmenterait durant la première semaine d'entraînement. Bien que ce résultat ne ressortent pas lors de l'analyse statistique, cela signifierait que les urodèles ayant simplement appris l'indication du stimulus pour trouver l'abri vont préférentiellement se diriger vers ce dernier par rapport au reste du bac lors d'une adaptation à un nouvel environnement.

3.4.1.2 Test en arène

Dans un premier temps, l'effet naturel de la lumière polarisée sur le comportement des pleurodèles a été examiné. Les animaux ont été testés sous lumière normale et polarisée, puis les réponses obtenues ont été comparées (Figure 3.6). Des courbes gaussiennes comme celles des Figures 3.6 et 3.7 ont été ajustées aux données recueillies afin de modéliser les résultats le plus fidèlement possible. Les courbes présentées ici sont celles qui montrent le pourcentage de robustesse de représentation des données le plus élevé.

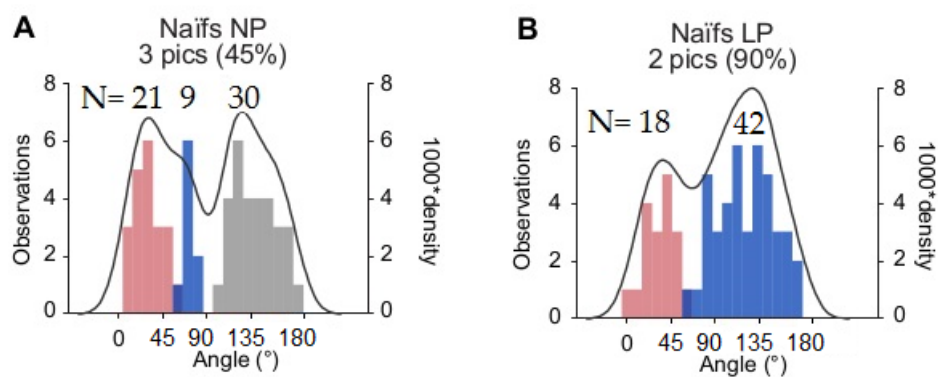


FIGURE 3.6 – Groupe d'animaux naïfs testés (A) sous lumière non polarisée (NP) et (B) sous lumière polarisée (LP). *N* représente l'effectif d'animaux pour une fourchette d'angles donnée (représentée par les différentes couleurs). Les couleurs des figures représentent les clusters de répartition des animaux suivant leurs positions d'arrivée dans l'arène. Les pourcentages au dessus des figures donnent le nombre des cas pour lesquels la distribution à deux ou trois pics sort dans un set de 20 analyses en prenant en compte le critère d'information d'Akaike.

Les tests statistiques ont révélé des résultats significativement différents (ANOVA à une voie, $p < 0,05$) entre les deux conditions, montrant ainsi un effet avant entraînement de la lumière

polarisée. Elle a tendance à mieux grouper les réponses obtenues et à diminuer la dispersion de la courbe, en passant d'une représentation gaussienne à trois modes à une distribution qui semble bimodale. Cet effet semble ici être concentré sur le second pic de réponse, son regroupement étant augmenté en présence de lumière polarisée. Ainsi, les animaux présentent des réponses plus homogènes en présence de lumière polarisée. La tendance de ces résultats nous conforte dans l'idée que les pleurodèles utilisent bien ce type de stimulus de façon innée pour s'orienter, leurs réponses étant moins aléatoires lorsqu'elle est présente.

Afin de vérifier dans quelle mesure cet effet apparemment naturel de la lumière polarisée était affecté par l'entraînement, le même test a été réalisé avec des animaux entraînés (Figure 3.7). Une fois encore, l'effet sur le second pic est visible et significatif (ANOVA, $p < 0,05$) et le rapport des effectifs est doublé car il passe de 0.875 (28/32) à 1.857 (39/21). Ceci confirme le fait que l'entraînement ne changerait pas l'action naturelle de la lumière polarisée sur le comportement des animaux dans l'arène. En revanche, la courbe gaussienne passe de 95% d'apparition à 70% entre le test en lumière normale et celui en lumière polarisée. Elle serait donc moins robuste sous lumière polarisée, bien que cela reste très représentatif des résultats obtenus.

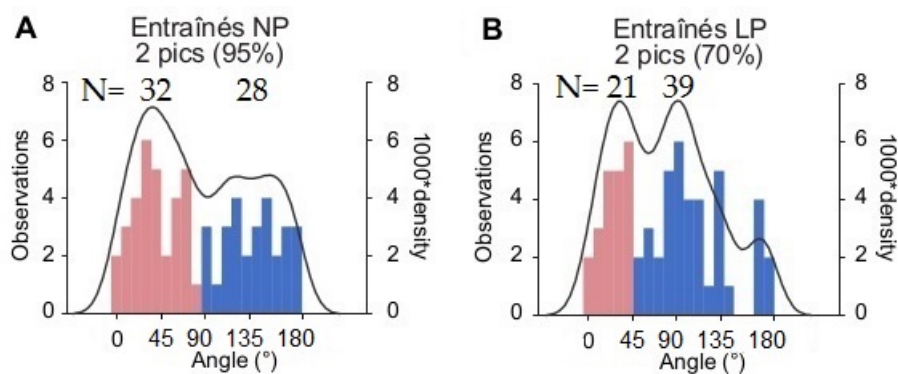


FIGURE 3.7 – Groupe d'animaux entraînés testés (A) sous lumière non polarisée (NP) et (B) sous lumière polarisée (LP). N représente l'effectif d'animaux pour une fourchette d'angles donnée (représentée par les différentes couleurs). Les pourcentages au dessus des figures donnent le nombre des cas pour lesquels la distribution à deux ou trois pics sort dans un set de 20 analyses en prenant en compte le critère d'information d'Akaike (rappel : la zone considérée comme « positive » après un entraînement se situe entre 68 et 112 °).

Nous avons ensuite étudié l'écart entre les deux pics pour chacun des graphiques (pour la Figure 3.6-A, les deux premiers pics à N=21 et N=9 ont été regroupés sous la forme d'une tendance centrale moyenne). L'écart entre les deux est d'environ 90° (plus ou moins 5°) excepté pour les animaux entraînés testés sous lumière polarisée. Pour ce dernier groupe, l'écart se resserre à 75°. L'entraînement ainsi que le test sous lumière polarisée aurait donc tendance à homogénéiser les résultats obtenus, en réduisant l'écart entre les deux ensembles de réponses. Les animaux entraînés montreraient donc un tropisme plus marqué pour la zone considérée comme positive dans l'arène (68 à 112 °) lorsque celle-ci est éclairée d'une lumière polarisée.

Selon ces résultats, l'entraînement semble ne pas avoir d'effet sur les distributions du nombre d'animaux mais il aurait en revanche un effet sur la moyenne des pics (tendance centrale) en rapprochant ces derniers de la zone de réussite (entre 68° et 112°). Ce phénomène se traduirait donc par une augmentation du nombre d'animaux aboutissant dans cette zone. Ces observations

montrent que les paramètres des pics obtenus, indépendamment du succès ou non de la tâche, pourraient être influencés par un ou plusieurs facteurs. Ainsi, la lumière polarisée présenterait un effet de renforcement sur le nombre d'animaux inclus dans ce cluster (présents dans le second pic de réponse).

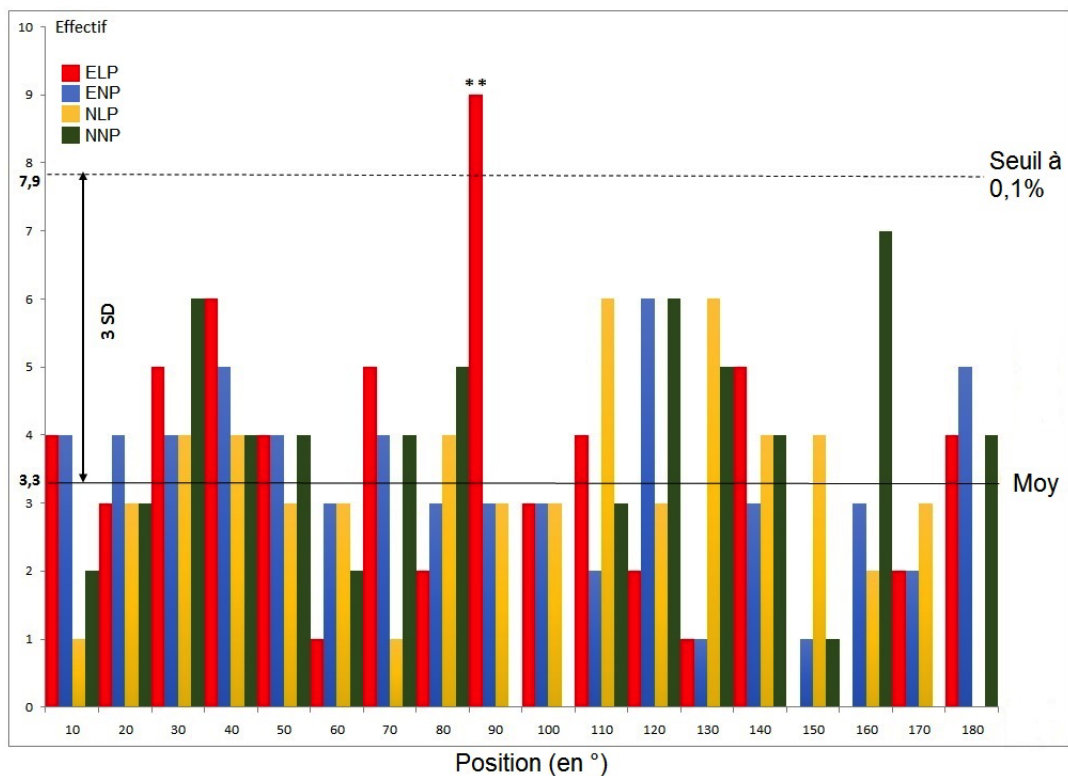


FIGURE 3.8 – Effectifs d’animaux par décile d’angles. ELP : Entraînés testés sous lumière polarisée ; ENP : Entraînés testés sous lumière Non polarisée ; NLP : Naïfs testés sous lumière polarisée ; NNP : Naïfs testés sous lumière Non polarisée ; Moy : moyenne ; SD : Ecart-type. Le seul pic ressortant et significatif est celui à 90° pour les animaux entraînés testés sous lumière polarisée.

Pour vérifier que l’entraînement améliorerait bien les performances des animaux dans la réussite de ce test, un histogramme de distribution par décile en fonction de la position dans l’arène (indiquée par un angle en degré) a été réalisé (Figure 3.8) et un seuil de significativité à $p < 0,001$ (moyenne + 3 SD) a été posé comme repère. Le seul pic ressortant de cette analyse est celui qui est observé à 90° pour les animaux entraînés et testés sous lumière polarisée. Ces animaux s’orienteraient donc préférentiellement vers la zone dans laquelle l’abri est censé se trouver, et ce uniquement sous lumière polarisée, ce qui démontrerait ainsi un entraînement effectif et valide ce conditionnement.

Enfin, le temps d’exécution de la tâche comportementale a été évalué en comparant les temps de réponse entre les quatre différentes conditions (Figure 3.9). En raison de la faible taille des échantillons, aucune relation statistique n’a pu être établie entre les animaux entraînés et les animaux naïfs selon les conditions de lumière polarisée ou non. En revanche un test non-paramétrique de Wilcoxon montre une différence significative lorsque l’on regroupe les données obtenues uniquement en fonction de l’entraînement ou non des animaux et il semble donc bien que les animaux entraînés mettent globalement moins de temps pour atteindre les bords de l’arène.

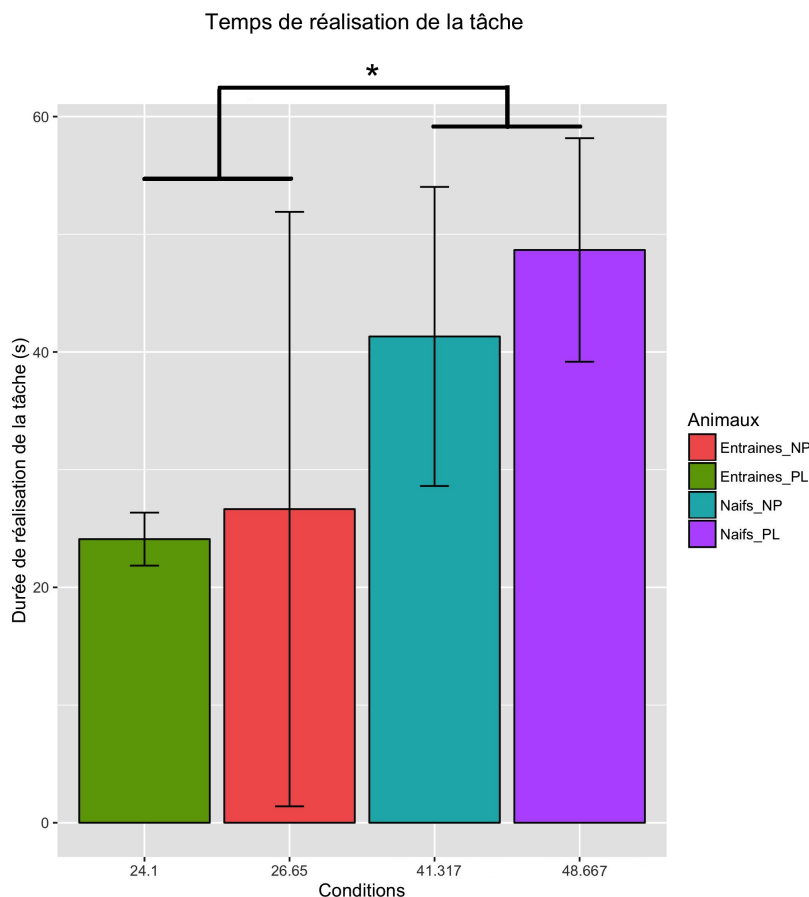


FIGURE 3.9 – Temps d’exécution de la tâche comportementale en fonction des conditions. Le temps (en ordonnées) est représenté en secondes. NP : Lumière non polarisée ; LP : Lumière polarisée. Le test de Wilcoxon est appliqué en comparant les données regroupées sur les animaux entraînés avec celles regroupées sur les animaux naïfs ($N = 6$, $p < 0.05$).

3.4.2 Lésions dopaminergiques

3.4.2.1 Analyses vidéos

En observant les pleurodèles après les avoir lésés, nous avons pu remarquer certains changements dans leur comportement. En effet, aux premiers jours post-lésionnels, ils paraissaient hypokinétiques, avec une marche et une nage plus saccadées, indiquant un comportement cohérent avec l’équivalent urodèle d’un tableau de type parkinsonien [Parish et al., 2007].

Nous avons effectué plusieurs mesures post-lésionnelles pour quantifier les effets des lésions dopaminergiques sur les comportements moteur et cognitif. Nous avons donc analysé le nombre moyen de pas, de cycles de nage ainsi que la durée moyenne de ces derniers au cours de la récupération des animaux (voir la section 3.3.2.3 pour plus de détails sur les moyens d’obtention de ces résultats préliminaires).

La Figure 3.10 représente l’évolution du nombre moyen de pas consécutifs (au moins trois d’affilés) au cours des jours suivant la lésion. Pour l’animal 1 (en rose sur la Figure), nous pouvons observer une diminution du nombre de pas à J+1 post-lésionnel, ce qui est en accord avec les enregistrements vidéo. Pour ce même animal, le nombre de moyen de pas effectués consécutivement va progressivement augmenter au fil du temps, pour revenir à la normale au

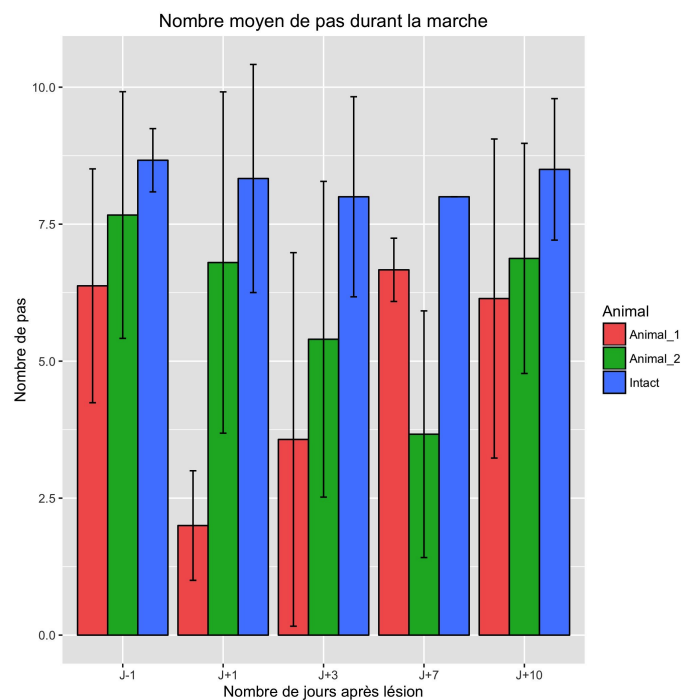


FIGURE 3.10 – Évolution du nombre de pas moyen après lésion dopaminergique. *Aucun test statistique ne permet de valider les différences observables sur la figure. Les tailles d'échantillon varient entre 3 et 8.*

bout d'une semaine. Pour l'animal 2 en revanche, les effets de la lésions sont beaucoup moins visibles et cohérents. En effet, ce paramètre montre une évolution qu'il est difficile d'expliquer et de relier à la lésion. Il semble que la lésion n'ait pas d'effet (entre J-1 et J+1) mais qu'en revanche il manifeste une diminution graduelle de son activité avant une reprise à J+10.

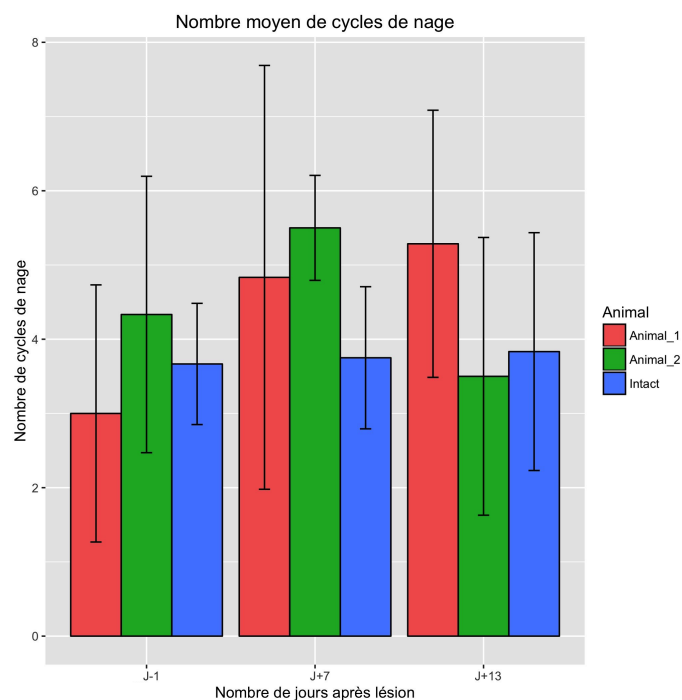


FIGURE 3.11 – Évolution du nombre de cycles de nage moyen après lésion dopaminergique.

La Figure 3.11 représente l'évolution du nombre moyen de cycles de nage consécutifs (au moins trois d'affilés également) après la lésion. Nous rappelons qu'il n'a pas été possible de réaliser ce test de nage avant J+7 post-lésionnel en raison de la cicatrice laissée par la chirurgie. Même s'ils ne sont pas significatifs, ces résultats et les analyses vidéos tendent à montrer une tendance à l'augmentation du nombre de cycle de nage après la lésion.

Nous nous sommes dans second temps intéressés à la mesure, non plus du nombre d'occurrences des pas ou des cycles de nage mais à leur durée pour tenter de détecter un effet plus net.

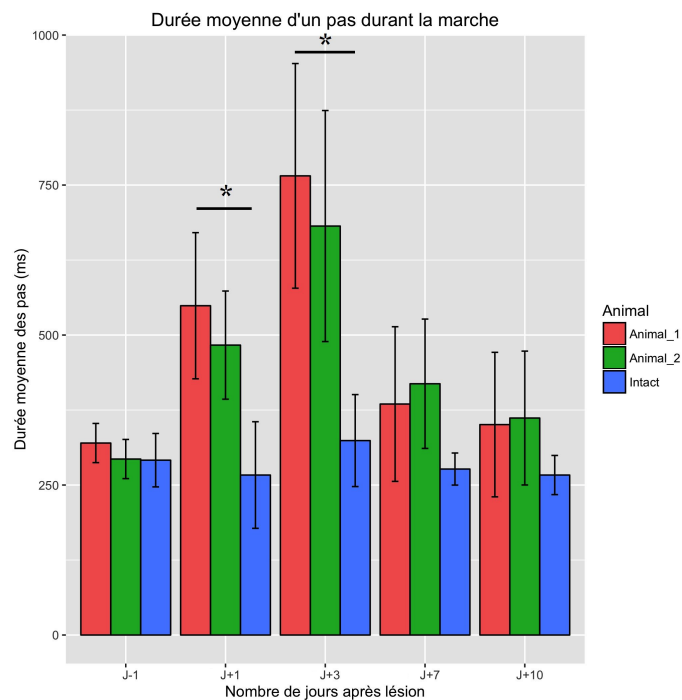


FIGURE 3.12 – Évolution de la durée moyenne des pas après lésion dopaminergique. Le test de Wilcoxon montre une différence significative $p < 0.05$ à J+1 et à J+3 entre les animaux 1 et 2 et l'animal intact (les tailles totales des échantillons sont respectivement $N_T = 76$ et $N_T = 93$ pour ces analyses).

La Figure 3.12 représente l'évolution de la durée moyenne des pas (dans le cas où l'animal fait au moins trois pas d'affilés). Pour ce paramètre, nous observons un allongement significatif de la durée des pas après lésion vs. contrôle à J+1 et à J+3. De plus, à J+3 après la lésion, les deux animaux ont presque doublé la durée moyenne de leurs pas. C'est donc à ce moment que ce paramètre semble le plus marqué et qu'il est possible d'observer les effets les plus visibles des lésions dopaminergiques. À J+7 après la lésion, nous observons une stabilisation de la durée moyenne des pas, qui conserve cependant une tendance à être plus élevée que chez l'animal intact.

Enfin, la Figure 3.13 représente l'évolution de la durée moyenne des cycles de nage (encore une fois dans le cas d'au moins trois cycles consécutifs). Ce paramètre est influencé par la lésion dopaminergique comme le montre la différence observée entre l'animal intact et l'animal 1 à J+7.

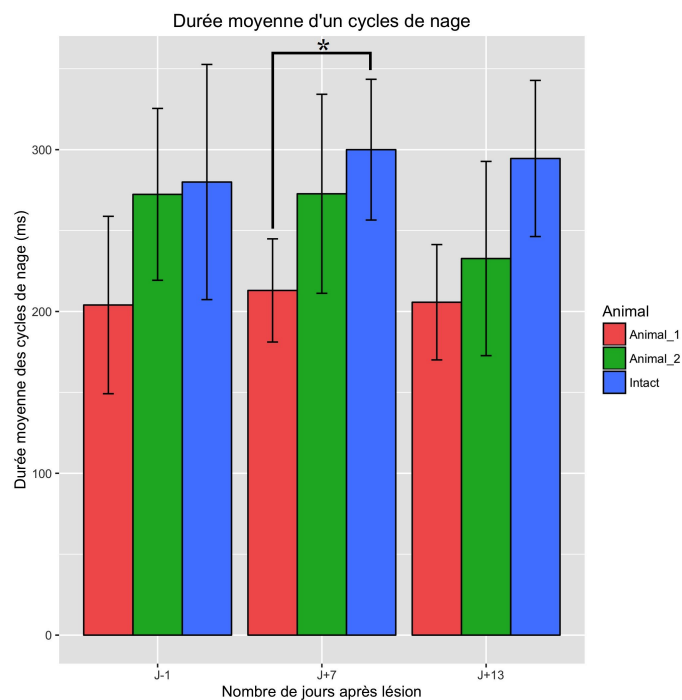


FIGURE 3.13 – Évolution de la durée moyenne des cycles de nage après lésion dopaminergique. Le test de Wilcoxon montre une différence significative à J+7 ($p < 0.05$) dans la durée des cycles de nage (la taille totale de l'échantillon est $N_T = 63$).

Tous ces résultats regroupés montrent que, dans le cas d'une lésion au niveau des neurones dopaminergiques diencéphaliques/mésencéphaliques, les pleurodèles ont tendance à montrer un tableau de type parkinsonien ce qui est avéré ici par leurs pas qui seraient moins nombreux et plus lents qu'en condition normale. Pour ce qui est de la nage, ils nageraient de façon plus saccadée en réalisant plus de cycles de nage, plus rapidement. N'ayant pas pu poursuivre plus avant ces investigations, nous n'avons malheureusement pas pu obtenir de résultats statistiquement plus solides et pour la plupart de nos métriques seules des tendances sont visibles.

3.4.2.2 Observations au microscope à fluorescence

Afin de déterminer la localisation, la nature et l'amplitude des lésions effectuées, nous avons analysé nos marquages au microscope à fluorescence. La Figure 3.14 montre un exemple de marquage DAPI et TH (destinés respectivement à révéler les noyaux et les cellules dopaminergiques). Ces clichés ont préalablement été comparés avec des schémas de neuroanatomie chez le pleurodèle [Marín et al., 1997a, Marín et al., 1997b] afin de déterminer si nous observions bien la bonne zone. La partie gauche a été réalisée avec un animal sain non lésé, et la partie droite avec un animal lésé (un mois post-lésion) afin de voir si la récupération motrice observée coïncidait avec la récupération neuronale.

Sur cette figure en bleu nous pouvons observer les noyaux des cellules du diencéphale et du mésencéphale des pleurodèles, alors qu'en rouge nous distinguons uniquement les neurones dopaminergiques. Sur ces images, nous pouvons également observer les fibres neuronales pour les deux conditions. Nous n'avons pas relevé ici de différence visible entre le nombre de neurones dopaminergiques avant et après lésion.

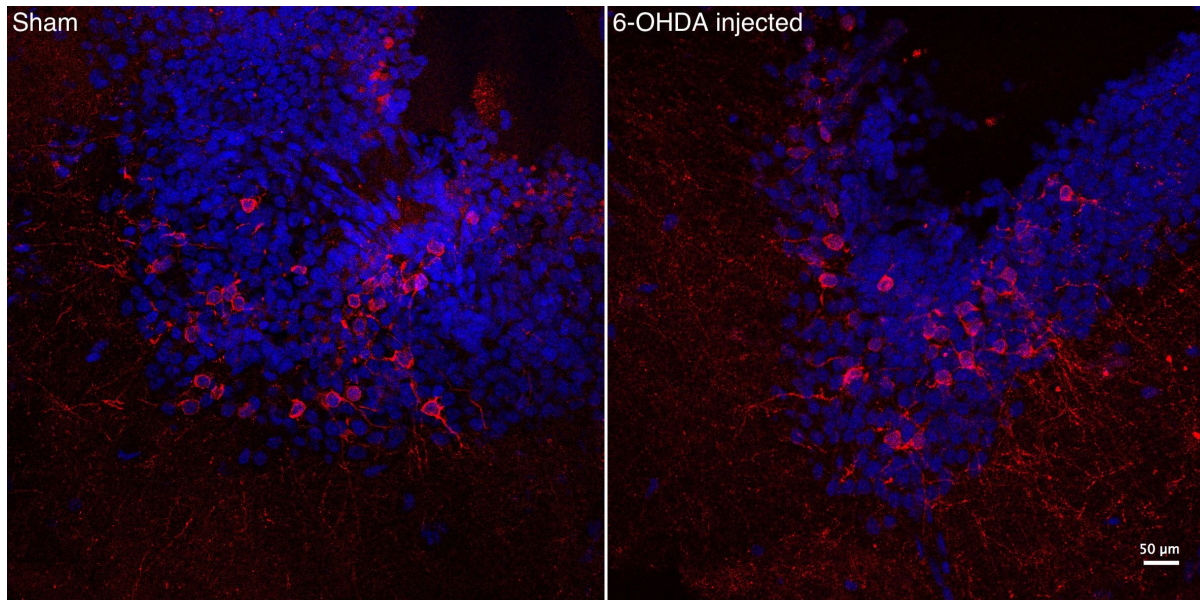


FIGURE 3.14 – *Cliché de microscopie à fluorescence montrant les marquage DAPI/TH des neurones du mésencéphale/diencéphale des pleurodèles en condition contrôle et un mois après une lésion. Le cliché de gauche est issu d'un animal témoin (« Sham ») alors que celui de droite a été réalisé à partir d'un animal lésé sacrifié un mois après la lésion. Les cellules révélées par le DAPI apparaissent en bleu et celles révélées par la TH en rouge.*

3.5 Discussion

3.5.1 Synthèse des travaux

Le travail d'analyse que nous avons réalisé à partir des résultats initiaux de Manon BONNET-SAVE montre un conditionnement effectif mais partiellement reproductible chez l'urodèle, en mettant à profit sa sensibilité naturelle à la lumière polarisée pour s'orienter. Ce modèle de conditionnement opérant constitue une première étape dans la mise en place d'un modèle comportemental de la prise de décision chez cet animal.

Pour compléter ce travail, nous avons effectué des lésions au niveau des neurones dopaminergiques diencéphaliques/mésencéphaliques afin de tester leur effet sur les BG et de rajouter la possibilité d'étudier en plus du modèle physiologique comportemental, un modèle pathologique de type parkinsonien. Le protocole mis en place ici a du être adapté du fait de l'originalité de ce modèle animal, mais les résultats préliminaires obtenus sont encourageants. Certains points restent cependant à améliorer à commencer par la taille des échantillons d'animaux utilisés.

3.5.2 Points sensibles des méthodes d'analyse

3.5.2.1 Effet du champ magnétique terrestre

Lors des expérimentations, un effet du champ magnétique terrestre sur l'orientation des urodèles a été observé, mais nous ne sommes pas parvenu à le caractériser précisément. En effet, nous ignorons encore le degré d'implication de ce vecteur sur le sensorium des urodèles et donc ses effets sur les résultats obtenus. Une étude complémentaire sur l'effet de ce champ magnétique serait un apport déterminant pour étayer nos résultats. Pour cela, un électro-aimant

d'une puissance convenable devrait permettre de contrôler le champ afin de tester l'effet de sa direction sur l'orientation des animaux, de la même manière que pour la lumière polarisée. Cet aimant pourrait influencer la perception des animaux pour l'axe Nord/Sud et donc nous permettre de voir à quel point les animaux l'utilisent dans leur orientation.

3.5.2.2 Apathie d'après lésion

Après lésion, l'apathie observée chez les urodèles peut biaiser nos interprétations lors des tests comportementaux. Ce phénomène peut être dû à plusieurs facteurs, comme par exemple l'atteinte de la mémoire acquise de la mémoire après apprentissage ou bien encore une baisse de la composante « motivationnelle ». Certains animaux ayant dû être stimulés physiquement pour avancer les premiers jours après lésion (contact répété au bout de l'appendice caudal), une perturbation dans le choix de la direction à prendre pourrait également être induite indépendamment d'une atteinte du système moteur. Pour vérifier cela, il aurait fallu réaliser le même genre de craniectomie mais sans l'injection de 6-OHDA afin de pouvoir différencier les effets de la lésion dopaminergique de ceux de la chirurgie.

3.5.2.3 Analyses vidéos

D'après les résultats préliminaires obtenus, l'analyse de la durée moyenne d'un pas serait la meilleure métrique pour quantifier l'effet des lésions dopaminergiques. Pour une meilleure reproductibilité et robustesse des résultats, il faudrait réaliser le même genre d'analyses sur un groupe d'animaux plus important, à différents stades avant et après lésion. Ceci permettrait un meilleur suivi de l'évolution des effets moteurs et de la récupération des animaux. Des enregistrements plus réguliers post-opératoires (comme par exemple tous les deux ou trois jours) rendraient également mieux compte de cette évolution.

3.5.2.4 Immunofluorescence

Sur les images issues de microscopie à fluorescence, il n'y a pas de différence observable entre les coupes de cerveaux d'un animal intact comparé à un animal lésé un mois après la lésion. Cela pourrait s'expliquer par le fait que, l'immunofluorescence étant réalisée tardivement, un mécanisme de régénération neuronale a pu se mettre en place entre temps chez nos animaux. Cette théorie irait dans le sens de la très grande faculté de récupération motrice observée au cours du temps et analysée dans la partie précédente. Cependant, il aurait été utile de réaliser le même genre de marquage aux mêmes jours que les enregistrements vidéo afin de pouvoir confirmer les effets de la lésion directement au niveau neuronal et ainsi de pouvoir les comparer.

3.5.3 Perspectives

Les résultats présentés ici indiquent que le manque de dopamine au niveau des BG des pleurodèles lésés aurait tendance à provoquer des troubles moteurs parkinsoniens typiques. Ces résultats préliminaires sont encourageants et montreraient que les lésions des neurones dopaminergiques que nous avons effectué sur ces animaux fonctionnent. Il faudrait donc simplement augmenter le nombre d'animaux étudiés. En effet, le fait d'effectuer ce type de lésions sur

plus d'animaux permettrait non seulement d'augmenter la significativité de nos résultats, mais également de réaliser des marquages d'immunofluorescence aux mêmes stades que les analyses vidéos afin de les corrélérer directement avec les effets observés au niveau neuronal.

Une autre application intéressante de ce travail lésionnel serait de soumettre les animaux lésés (naïfs et entraînés) au test comportemental en arène afin de mieux appréhender l'implication des BG dans le processus d'apprentissage et de prise de décision. De plus, il serait également utile pour les protocoles lésionnels à venir d'avoir un atlas de neuroanatomie pour les pleurodèles comme c'est le cas pour les rongeurs et primates classiquement utilisés en laboratoire. Un tel ouvrage demeure inexistant à ce jour. Une reconstitution numérique de l'intégralité de coupes obtenues à partir d'animaux sains permettrait d'avoir un exemple de l'organisation générale de leur CNS. Cependant, il s'agirait là d'un travail long et fastidieux car ces animaux ne sont pas « standardisés » et leurs dimensions ne sont pas ou peu conservées entre eux.

Enfin, nous avons envisagé initialement d'appliquer les neurosciences computationnelles et notamment la modélisation à l'étude de ce modèle animal. En effet, les pleurodèles possédant des structures analogues aux BG mais avec une complexité moindre, il serait théoriquement possible de réaliser un modèle informatique des BG pour cet animal. Il serait même envisageable d'inclure d'autres régions du CNS pouvant jouer un rôle dans les processus de prise de décision afin d'avoir une vision plus globale de ce mécanisme. Nous avons montré qu'il était possible dans une certaine mesure d'entraîner ces animaux, et ce beaucoup plus rapidement que des primates (la simplification de la tâche comportementale jouant également un rôle dans la durée de l'apprentissage). Ce point permettrait d'éprouver les capacités du modèle computationnel plus rapidement et de le rendre plus robuste, les prédictions *in silico* du modèle pouvant être testés et validés (ou non) assez rapidement *in vivo*.

Discussion générale

« The most exciting phrase to hear in science, the one that heralds new discoveries, is not 'Eureka!' but 'Hum ... That's funny...' . »

Isaac Asimov

Sommaire

4.1	Résumé des travaux	113
4.1.1	Apports de l'approche de modélisation	113
4.1.2	Apports de l'approche expérimentale	116
4.2	Perspectives et compléments d'étude	117
4.2.1	Rôles de l'inhibition latérale au niveau thalamique et cortical	117
4.2.2	Ajout de structures associées aux BG	117
4.2.3	Voie indirecte dans le modèle computationnel	118
4.2.4	Influence du champ magnétique terrestre	118
4.2.5	Amélioration du suivi de la récupération des animaux	118
4.2.6	Apathie des animaux après chirurgie	119
4.2.7	Implémentation d'un modèle computationnel des BG des pleurodèles	119

Ce chapitre propose une synthèse portant sur les résultats présentés auparavant, en soulignant leurs apports respectifs et leurs limites ainsi que les améliorations et perspectives envisageables.

4.1 Résumé des travaux

Ce travail de thèse s'est inscrit dans la continuité des travaux réalisés antérieurement par les membres de l'équipe d'accueil du laboratoire et de leurs collaborateurs [Guthrie et al., 2013, Garenne et al., 2011, Leblois et al., 2006a, Pasquereau et al., 2007]. Son objectif principal était le développement d'un modèle computationnel des BG décrit à l'échelle du neurone et de la synapse afin de mieux comprendre les mécanismes de la sélection de l'action et de la prise de décision dans cette structure cérébrale. En parallèle de cette approche de modélisation, une étude purement expérimentale a été mise en place dans laquelle l'objectif, complémentaire de l'approche théorique, consistait à développer un modèle physiopathologique et comportemental permettant l'étude des mécanismes de prise de décision à l'aide d'un modèle animal novateur et présentant potentiellement des avantages, notamment au regard de l'utilisation des primates non humains.

4.1.1 Apports de l'approche de modélisation

Le modèle computationnel des BG développé dans le cadre de cette thèse est fonctionnel et capable de fournir des résultats cohérents et exploitables, que nous allons résumer dans cette partie. Pour rappel, le modèle des BG a été soumis à une tâche de prise de décision en condition d'incertitude. Lors de cette tâche, deux cibles différentes (choisies au hasard sans remise parmi les quatre possibles, chacune avec sa propre probabilité de récompense) étaient affichées à deux positions différentes (encore une fois choisies au hasard parmi les quatre possibles). Le réseau devait apprendre à choisir quelle cible avait la meilleure probabilité de récompense afin de maximiser globalement son taux de récompense. Ce modèle a ensuite été soumis à une variante de cette tâche où cette fois-ci il n'y avait que deux cibles possibles.

4.1.1.1 Implémentation du réseau

Le modèle computationnel développé dans cette thèse est composé de cinq structures cérébrales - à savoir le cortex, les BG (striatum, GPi, STN) et le thalamus - reliées par des boucles cortico-corticales de rétroaction positive et négative - dans ce cas respectivement les boucles directe et hyperdirecte.

Pour répondre à des contraintes anatomo-fonctionnelles et aussi pour prendre en compte la réalisation de la tâche, chaque structure cérébrale à l'intérieur du réseau a été divisée en plusieurs sous-régions. En nous appuyant sur la préservation de la somatotopie au sein des différentes structures des boucles des BG et de leurs projections thalamiques et corticales, et en reprenant une partie des postulats de modèles précédents, nous avons décrit chacune de ces sous-régions autour d'une composante cognitive pour l'encodage de la reconnaissance de la forme de la cible présentée, et d'une composante motrice pour l'encodage de la direction des cibles. Ces régions

ont également été mises à profit pour interpréter les choix et décisions du modèle. Nous avons également inclus les composantes associatives du cortex et du striatum qui d'après les données anatomo-fonctionnelles disponibles constituent des zones d'interaction des parties motrices et cognitives. Enfin nous avons inclus les propriétés d'inhibition latérales ayant été mises en évidence dans le striatum. Nous avons ensuite construit ce réseau à l'aide de populations de neurones de type LIF et de synapses à saut de potentiel individualisés et simulés en parallèle dans un programme que nous avons développé en C++. Ils nous a alors fallu ajuster les paramètres du modèle de façon à ce qu'il se comporte qualitativement et quantitativement de manière cohérente avec son homologue biologique.

4.1.1.2 Bruit et activité basale dans le réseau

En conditions physiologiques, chaque neurone d'un réseau biologique est soumis à un bombardement synaptique et à des effets thermiques. Nous avons donc appliqué un bruit blanc gaussien sur chaque neurone afin de mimer cet effet à l'intérieur du réseau. Après avoir testé plusieurs niveaux de bruit sur les capacités de prise de décision du réseau, nous l'avons fixé à 50 pA, valeur pour laquelle nous obtenions les meilleurs résultats.

De la même façon et comme les structures cérébrales que nous avons modélisées sont spontanément actives physiologiquement, nous avons ajouté un courant constant dans le réseau afin de reproduire cette activité basale. Après avoir effectué plusieurs mesures nous avons pu déterminer une valeur optimale pour ce courant, et nous l'avons fixé à 100 pA. L'activité des différentes populations cellulaires s'est ainsi trouvée cohérente avec les taux de décharge neuronale enregistrés *in vivo*.

4.1.1.3 Apprentissage dans le réseau

À l'état naif (avant apprentissage) notre réseau est capable de choisir spontanément une cible et donc d'effectuer la sélection d'une action. Ne connaissant pas les taux de récompenses des cibles à l'avance et n'ayant aucun pré-câblage pour aucune des cibles, le réseau va explorer l'espace des solutions et réussir la tâche environ une fois sur deux. Des récompenses accompagnent avec une certaine probabilité les actions sélectionnés par le réseau. Grâce à ces récompenses, un processus d'apprentissage par renforcement se met progressivement en place au cours des essais. Ce processus est en effet matérialisé par la modification des poids des synapses cortico-striatales cognitives sous l'influence de la libération de DA phasique accompagnant les récompenses. Ces libérations de DA vont moduler au niveau de ces synapses les mécanismes de LTP et de LTD. Ce protocole d'apprentissage par renforcement permet à notre réseau de sélectionner préférentiellement la cible optimale en suivant au cours des essais successifs, une cinétique similaire à celle qui est observée chez l'animal.

4.1.1.4 Effet de l'apprentissage sur le temps de réaction du modèle

Lors de l'analyse de nos résultats, nous avons mesuré les délais d'initiation des mouvements, c'est-à-dire les délais séparant la présentation des cibles à t_0 de la sélection de l'action issue de la composante « motrice » (sélection finale de la direction associée à la forme de la cible

choisie). En pratique, nous avons donc mesuré le délai entre t_0 et le moment de la rupture de la symétrie (seuil des 40 Hz atteint) observée au niveau des sous-populations corticales motrices et qui montre qu'une l'action est effectivement sélectionnée et engagée. Ces analyses ont montré que ce temps de réaction est diminué avec l'apprentissage.

4.1.1.5 Rôle de l'inhibition striatale latérale

Nous avons également étudié le rôle de l'inhibition striatale latérale dans les capacités d'apprentissage du modèle. Nous avons observé que cette inhibition jouait un rôle important dans le contrôle du gain au sein du réseau et dans les capacités globales d'apprentissage du modèle. En effet, nos résultats ont montré un meilleur taux de réussite, au niveau cognitif et moteur, lorsqu'elle était présente dans le réseau. Ces résultats théoriques, en accord avec d'autres données expérimentales et de modélisation semble bien confirmer que cet élément de connectivité doit être pris en considération pour étudier les relations structure-fonction au sein des réseaux des BG.

4.1.1.6 Protocole d'inversion de l'apprentissage

Lors d'un tel protocole, les conditions de délivrance de la récompense ont été changées sans indication complémentaire afin d'étudier les capacités d'adaptation du modèle à un changement de l'environnement. Nous avons divisé ce protocole en trois étapes : lors de l'étape 1 un entraînement classique est entrepris afin d'apprendre les valeurs respectives des cibles au modèle ; lors de l'étape 2 ces probabilités étaient permutées et lors de l'étape 3 elles étaient remises dans les conditions initiales de l'étape 1. Un délai plus ou moins long était nécessaire au modèle pour une optimisation de son taux de réussite à chaque nouvelle étape. Il était notamment plus important lors de l'étape 2. Ce phénomène s'expliquait facilement dans la mesure où le choix préférentiel du modèle continuait de se porter vers des cibles dont la probabilité de récompense était désormais beaucoup plus faible. Néanmoins il a montré qu'après un temps raisonnable d'exploration du nouvel espace des solutions il était capable de se comporter de nouveau de façon optimale, ce qui est cohérent ici encore avec les données expérimentales disponibles.

4.1.1.7 Protocole d'extinction

Pour ce protocole, nous avons stoppé complètement la délivrance des récompense après l'apprentissage par renforcement et observé le comportement du réseau pour déterminer la façon dont notre modèle réagissait. Après une session d'entraînement similaire aux précédentes nous avons continué les enregistrement des actions et des activités neuronales en l'absence totale de récompense. Les performances de notre modèle en terme de GC ont été significativement différentes avant et après extinction montrant que le réseau retournait à une phase purement exploratoire en l'absence de corrélation entre la forme d'une cible et la récompense associée. Après un temps plus long, le taux de rupture de symétrie (SB) lui même a tendance à baisser ce qui se traduit dans la composante motrice par une baisse générale de mouvement. Ce comportement qui pourrait être perçu chez un animal comme une baisse de la composante motivationnelle s'explique dans notre modèle de façon assez mécanique par une diminution des connexions

cortico-striatales cognitives associées aux cibles. Une seconde série d'essais a porté sur un rétablissement des récompenses après une phase d'extinction sur les mêmes bases qu'avant cette phase. Dans ce cas nous observons un retour de l'activité du modèle en terme de sélection des actions et aussi de choix optimaux (GC) et de décisions optimales (GD).

4.1.1.8 Tests lésionnels réalisés *in silico*

Différents types de lésions (striato-pallidales, cortico-thalamiques et pallido-thalamiques) ont été appliquées spécifiquement au sein du réseau afin d'avoir une vue globale de sa dynamique interne et de sa cohérence. Nos résultats ont montré que dans le cas de lésions striato-pallidales, le taux moyen de décharge neuronal est diminué dans l'ensemble du réseau. Ce phénomène était même amplifié dans le cas de lésions cortico-thalamiques alors que nous avons observé l'effet inverse (augmentation des taux de décharge) lors de l'application de lésions pallido-thalamiques. Ces résultats montrent que la dynamique interne de notre modèle des BG est cohérente. De plus ils nous permettent également de tester les degrés d'implication relative de chaque structure du modèle dans les mécanismes d'apprentissage et de prise de décision. Nous avons également pu montrer, en accord avec des résultats expérimentaux observés au laboratoire d'accueil chez le primate, qu'une lésion sur le GPi empêchait l'acquisition d'un apprentissage par renforcement par rapport à une condition contrôle.

4.1.2 Apports de l'approche expérimentale

Cette approche s'inscrit dans la continuité de travaux sur le conditionnement de pleurodèles en utilisant la sensibilité naturelle de ces animaux à la lumière polarisé pour leur orientation. Si nous avons tenté d'utiliser ce stimulus c'est que ces animaux sont par ailleurs très peu accessibles aux tests comportementaux réalisables par des vertébrés supérieurs (rats, souris, primates ...). Nous avons en parallèle commencé la mise en place d'un protocole lésionnel spécifique permettant l'étude d'un modèle parkinsonien du pleurodèle afin d'étudier l'implication des BG lors de la prise de décision en condition normale et pathologique.

4.1.2.1 Conditionnement fonctionnel et reproductible

Les animaux utilisés dans cette étude sont des urodèles (*Pleurodeles waltlii*) juvéniles en raison de leurs grandes capacités de régénération. Ils ont été entraînés à suivre le vecteur de polarisation d'une lumière artificielle afin de trouver leur récompense (un endroit où s'abriter). Même si quelques paramètres restent à améliorer pour que ce protocole puisse être publié et employé en routine, les résultats de l'entraînement ont été très prometteurs et montrent qu'ils peuvent être conditionnés pour des études comportementales impliquant des mécanismes de prise de décision.

4.1.2.2 Protocole lésionnel fonctionnel

Faisant suite au protocole de conditionnement des pleurodèles, nous avons mis en place un protocole lésionnel afin de produire des animaux parkinsoniens. A cette fin, il nous a fallu léser sélectivement les neurones dopaminergiques du mésencéphale/diencéphale de ces animaux.

Les lésions ont été effectuées au niveau du troisième ventricule en raison de l'homologie fonctionnelle entre cette population neuronale et celles de l'aire tegmentale ventrale et de la substance noire chez les mammifères. Dans le but de vérifier l'efficacité de ces lésions, nous avons suivi l'évolution du comportement des pleurodèles avant et plusieurs jours après l'opération. Nos résultats, bien que préliminaires, ont montré un allongement significatif de la durée moyenne des pas lors de la marche après lésion. Les animaux adoptaient également les caractéristiques comportementales d'un syndrome parkinsonien après la lésion avec un appauvrissement de l'initiation des mouvements, mais affichaient par ailleurs une tendance à récupérer très rapidement leurs facultés motrices (au bout d'une semaine environ). Nous avons également effectué une analyse par marquages immunohistochimiques afin d'observer l'effet des lésions dopaminergiques *in situ* au niveau neuronal.

4.1.2.3 Mise en place de méthodes d'analyse

Ce type d'étude sur ce modèle animal étant peu courante, nous avons dû mettre en place et adapter les protocoles expérimentaux ainsi que les méthodes d'analyses des résultats. Pour le test comportemental, nous avons analysé la direction prise par les pleurodèles lorsqu'ils étaient placés au centre d'une arène circulaire en présence ou non de lumière polarisée. Nous avons ensuite relevé les angles associés à ces positions et analysé le nombre d'animaux présents dans la zone de réussite du test afin de vérifier de l'efficacité de l'apprentissage. Pour l'étude de l'effet des lésions dopaminergiques, nous avons effectué des enregistrements vidéos des pleurodèles en train de marcher et de nager les jours suivant leur opération. Nous avons quantifié le nombre moyen de pas et de cycles de nage réalisés, ainsi que leurs durées moyennes.

4.2 Perspectives et compléments d'étude

4.2.1 Rôles de l'inhibition latérale au niveau thalamique et cortical

La prise en compte des effets d'inhibition latérale au niveau striatal nous a permis d'aller plus loin dans la compréhension des capacités fonctionnelles du réseau des BG. Ce type d'inhibition, bien que diffus en apparence, semble y jouer un rôle important dans le traitement de l'information dans le réseau, et notamment dans le contrôle du gain. Parmi les pistes à envisager pour faire évoluer notre modèle, nous avons envisagé de prendre en compte ce type d'inhibition latérale au niveau thalamique également. Il serait possible alors de tester ces effets sur l'encodage temporel des signaux dans le réseau et par conséquent d'améliorer l'information transmise au cortex [Wang et al., 2011]. Dans un deuxième temps il serait intéressant de tester les effets de l'inhibition latérale qui est également présente au niveau cortical [Piron et al., 2016] avec pour objectif ultime de présenter un modèle le plus générique et cohérent possible en ce qui concerne les contributions des différentes régions interagissant avec les BG.

4.2.2 Ajout de structures associées aux BG

Notre modèle comporte actuellement cinq structures cérébrales. Des régions supplémentaires pourraient y être ajoutées afin d'améliorer ses performances, notamment l'amygdale ou l'hippo-

campe. L'hippocampe pourrait permettre de perfectionner le processus d'apprentissage du fait de son rôle dans l'apprentissage procédural et inconscient [Atallah et al., 2004]. Par ailleurs, dans les mécanismes liés au protocole d'extinction, il existe des relations étroites entre le striatum dorso-latéral et le noyau central de l'amygdale qui pourraient jouer un rôle dans la diminution du temps de rappel après un tel protocole. L'ajout de ces régions actuellement non prises en compte dans notre modèle pourraient donc le rendre plus performant et nous permettrait d'aller plus loin dans la compréhension des mécanismes qui nous intéressent. L'approche intégrative du fonctionnement du CNS est du reste le projet principal et la raison d'être de l'équipe d'accueil de mon travail de thèse.

4.2.3 Voie indirecte dans le modèle computationnel

La voie indirecte, autre voie majeure des BG avec les voies directe et hyperdirecte, n'a pas été incluse dans notre modèle. Nous avons en effet construit une partie de notre étude sur l'utilisation du minimum de circuiterie possible afin de montrer dans un réseau connexionniste les mécanismes de sélection de l'action et d'étendre cette propriété à d'autres protocoles et à d'autres situations. Notre modèle comprend une boucle de sélection motrice et une boucle de sélection cognitive, qui interagissent au niveau striatal pour permettre un échange des informations et ainsi le processus de sélection de l'action. Le fait de compléter notre modèle par cette voie indirecte impliquerait également l'ajout du GPe et constituerait donc une étape prioritaire à considérer. Elle permettrait sans doute un meilleur contrôle de la durée du processus de sélection de l'action et des propriétés du réseau en général.

4.2.4 Influence du champ magnétique terrestre

Dans l'étude comportementale menée sur les pleurodèles, nous avons observé un effet du champ magnétique terrestre sur leur orientation (cf. section 3.5.2.1). Cependant, nous ignorons encore le degré d'implication de ce vecteur dans le comportement de nos animaux et donc dans les résultats que nous avons obtenus. Pour corriger ce phénomène ou au moins le contrôler, il faudrait réaliser une étude complémentaire où un électro-aimant serait utilisé pour perturber la perception de l'axe Nord-Sud par les animaux lors de leur orientation. Les pleurodèles seraient entraînés et testés dans les mêmes conditions que pour la lumière polarisée et ces résultats permettraient de discriminer les effets du champ magnétique terrestre de ceux de la lumière polarisée utilisée expérimentalement.

4.2.5 Amélioration du suivi de la récupération des animaux

Il serait intéressant dans l'étude de la prise de décision d'effectuer les tests en arène sur des animaux lésés afin d'observer les effets de la lésion sur les capacités d'apprentissage des animaux comme cela a été déjà suggéré dans la section 3.5.3. Cela nécessiterait un nombre d'animaux plus important du fait des nombreuses conditions à tester. Il faudrait en effet entraîner un premier groupe, le tester en arène, effectuer la lésion sur certains d'entre eux et ensuite les tester à nouveau dans l'arène régulièrement pendant leur récupération pour les comparer. Il serait également intéressant d'effectuer l'entraînement sur des animaux naïfs après lésion pour

voir s'ils seraient capables d'apprendre la tâche et ainsi avoir une idée plus précise du degré d'implication des BG dans l'apprentissage.

4.2.6 Apathie des animaux après chirurgie

Après les avoir opérés, nous avons observé une apathie générale chez tous nos animaux les premiers jours post-lésionnels (cf. section 3.5.2.2). Cela induit un biais dans leur analyse comportementale et donc dans nos résultats. Ce phénomène peut avoir plusieurs origines, comme une baisse de la motivation par exemple, ou encore une difficulté dans la prise de décision de la direction à suivre. Durant ces mêmes jours, certains pleurodèles nécessitaient une stimulation mécanique afin d'initier la marche alors qu'ils ne semblaient pas avoir de séquelles purement motrices. Pour discriminer les effets de la lésion de ceux de la chirurgie à proprement parler, il faudrait réaliser le même type d'opération sur des animaux sains, mais sans l'injection de 6-OHDA et observer leur récupération de la même façon que les animaux lésés.

4.2.7 Implémentation d'un modèle computationnel des BG des pleurodèles

Enfin, l'un des aspects les plus intéressants à développer serait de mêler l'approche de modélisation et l'approche expérimentale en implémentant le réseau des BG des pleurodèles sur la base de celui développé ici. Du fait de la simplification des structures, de la tâche comportementale et aussi de la taille moindre du réseau chez un tel animal (comparé aux rongeurs et aux primates), ce nouveau modèle pourrait être validé dans de meilleurs délais à la fois sur le plan expérimental et théorique car s'il existe une analogie structurelle intéressante et un invariant dans leur organisation, les tailles des populations neuronales sont nettement réduites entre les primates et les urodèles. De plus, les résultats obtenus *in silico* pourraient être plus facilement contrôlés *in vivo* et cela permettrait de tester des hypothèses sur les mécanismes de prise de décision avec plus d'efficacité et de plausibilité.

Conclusion

CE travail de thèse aura duré trois ans, au sein de l'équipe Mnémosyne de l'IMN et de l'INRIA de Bordeaux. Il a porté sur l'étude des mécanismes de prise de décision et de sélection de l'action dans la structure cérébrale des ganglions de la base. Pour développer notre modèle, nous sommes appuyés en partie sur des travaux antérieurs réalisés au sein du laboratoire et par des collaborateurs. Il nous a en effet paru logique d'essayer d'apporter des éléments de réflexion nouveaux et d'étendre les résultats précédents plutôt que de tenter de partir sur un concept radicalement différent. Il y a plusieurs raisons à cela. Outre le gain de temps réalisé (la mise en place d'un modèle computationnel est un processus très chronophage et toujours incertain) il y a un bénéfice scientifique réel dans le fait de valider puis d'étendre les résultats acquis auparavant. Il y a enfin la satisfaction ressentie lorsque l'on participe à la progression d'une idée qui nous semble pertinente.

Le modèle computationnel que nous avons développé s'appuie donc sur des tâches comportementales menées chez le primate, montrant l'implication des BG dans des protocoles de prise de décision en condition d'incertitude [Pasquereau et al., 2007, Piron et al., 2016]. Lors de ces études, des données expérimentales sous la forme d'enregistrements électrophysiologiques et de mesures comportementales ont été effectués. En parallèle plusieurs modèles computationnels ont été développés suite à cette étude parmi lesquels les travaux de [Leblois et al., 2006a] et de [Guthrie et al., 2013]. Nous nous sommes inspirés des résultats expérimentaux et de la filiation des modèles théoriques précédents pour guider notre réflexion et ajouter des éléments et un formalisme nouveau sur les modèles des BG. Nous avons développé une approche connexionniste décrivant les structures étudiées à l'échelle neuronale et synaptique (microscopique) en lieu et place des modèles de populations (mésoscopiques) employées précédemment. L'implémentation et la mise au point d'un tel modèle pose des problèmes supplémentaires d'ordre méthodologique mais à terme, il permet de mieux étudier l'impact des mécanismes neurobiologiques considérés à l'échelle cellulaire ainsi que les théories reliant les modalités de plasticité synaptique avec l'apprentissage. Une fois le modèle développé, nous l'avons validé en le soumettant à de nouveaux protocoles expérimentaux tels que différents types de tâches de prise de décision mais également en testant l'effet de lésions spécifiques locales en plusieurs points du réseau et enfin l'effet d'un protocole d'extinction sur ses capacités de sélection de l'action. Ce modèle étendu contribue selon nous à combler le fossé existant entre le traitement de l'information telle qu'elle est réalisée au niveau neuronal et les propriétés fonctionnelles des réseaux de grande taille.

Nous avons également commencé à travailler sur le développement d'un protocole novateur de prise de décision sur le pleurodèle de Waltham afin de mieux comprendre le rôle respectifs des constituants des BG. Ce type d'animal étant peu étudié, nous avons donc mis en place des moyens de conditionnement pour l'apprentissage et l'expérimentation, mais également pour l'analyse des résultats comportementaux. La plupart de ces résultats ne sont que des résultats préliminaires pour le moment, mais ils s'annoncent prometteurs pour des études ultérieures sur la prise de décision et l'apprentissage chez le pleurodèle. Enfin nous avons expérimenté et commencé à mettre en place au sein du laboratoire d'accueil un modèle parkinsonien [Parish et al., 2007] chez le pleurodèle qui là encore a donné des résultats préliminaires encourageants et qui constitueront, en complément de la tâche de prise de décision, un modèle complet d'étude physiopathologique unifié chez une nouvelle espèce animale.

Bibliographie

- [Abraham et al., 2014] Abraham, A. D., Neve, K. A., , and Lattal, K. M. (2014). Dopamine and extinction : A convergence of theory with fear and reward circuitry. *Neurobiology of learning and memory*, 108 :65–77. 30
- [Agatonovic-Kustrin and Beresford, 2000] Agatonovic-Kustrin, S. and Beresford, R. (2000). Basic concepts of artificial neural network (ann) modeling and its application in pharmaceutical research. *Journal of Pharmaceutical and Biomedical Analysis*, 22 :717–727. 34
- [Albin et al., 1989] Albin, R. L., Penney, J. B., and Young, A. B. (1989). The functional anatomy of basal ganglia disorders. *Trends in Neuroscience*, 12(10) :366–375. 5, 8, 12, 13, 15, 16, 43, 55, 86
- [Albus, 1971] Albus, J. S. (1971). A theory of cerebellar function. *Mathematical Biosciences*, 10 :25–61. 27
- [Albus, 1975] Albus, J. S. (1975). *A new approach to manipulator control : The cerebellar model articulation controller (CMAC)*. Trans. ASME. 28
- [Alegre et al., 2013] Alegre, M., Lopez-Azcarate, J., Obeso, I., Wilkinson, L., Rodriguez-Oroz, M. C., Valencia, M., Garcia-Garcia, D., Guridi, J., Artieda, J., Jahanshahi, M., and Obeso, J. A. (2013). The subthalamic nucleus is involved in successful inhibition in the stop-signal task : A local field potential study in parkinson’s disease. *Experimental Neurology*, 239 :1–12. 5
- [Alexander and Crutcher., 1990] Alexander, G. E. and Crutcher., M. D. (1990). Functional architecture of basal ganglia circuits : neural substrates of parallel processing. *Trends in Neurosciences*, 13(7) :266–271. 87
- [Alexander et al., 1986] Alexander, G. E., Long, M. R. D., and Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience*, 9 :357–381. 3, 9, 10, 55, 87
- [Alexander and Wickens, 1993] Alexander, M. E. and Wickens, J. R. (1993). Analysis of striatal dynamics : The existence of two modes of behaviour. *Journal of the Theoretical Biology*, 163 :413–438. 86
- [Alkemade et al., 2015] Alkemade, A., Schnitzler, A., and Forstmann, B. U. (2015). Topographic organization of the human and non-human primate subthalamic nucleus. *Brain Structure and Function*, 220(6) :3075–3086. 5
- [Armstrong et al., 2007] Armstrong, D., Halliday, W., Hawkins, C., and Takashima, S. (2007). *Pediatric Neuropathology : A Text–Atlas*. Springer Science. 15

- [Atallah et al., 2004] Atallah, H. E., Frank, M. J., and O'Reilly, R. C. (2004). Hippocampus, cortex, and basal ganglia : Insights from computational models of complementary learning systems. *Neurobiology of Learning and Memory*, 82 :253–267. 118
- [Bar-Gad and Bergman, 2001] Bar-Gad, I. and Bergman, H. (2001). Stepping out of the box : information processing in the neural networks of the basal ganglia. *Current Opinion in Neurobiology*, 11(6) :689–695. 55, 86
- [Bar-Gad et al., 2003] Bar-Gad, I., Morris, G., and Bergman, H. (2003). Information processing, dimensionality reduction and reinforcement learning in the basal ganglia. *Progress in Neurobiology*, 71 (6) :439–73. 2, 3, 4, 5, 6, 7, 10, 12, 13, 27, 29, 55
- [Barker et al., 2014] Barker, J. M., Taylor, J. R., and Chandler, L. J. (2014). A unifying model of the role of the infralimbic cortex in extinction and habits. *Learning & Memory*, 21 :441–448. 80
- [Benabid, 2003] Benabid, A. L. (2003). Deep brain stimulation for parkinson's disease. *Current Opinion in Neurobiology*, 13 :696–706. 13
- [Berthet et al., 2012] Berthet, P., Hellgren-Kotaleski, J., and Lansner, A. (2012). Action selection performance of a reconfigurable basal ganglia inspired model with hebbian–bayesian go-nogo connectivity. *Frontiers in Behavioral Neuroscience*, 6(65) :10–3389. 55, 77
- [Berthet et al., 2016] Berthet, P., Lindahl, M., Tully, P. J., Hellgren-Kotaleski, J., and Lansner, A. (2016). Functional relevance of different basal ganglia pathways investigated in a spiking model with reward dependent plasticity. *Frontiers in Neural Circuits*, 10(53) :1–21. 50
- [Bevan et al., 1997] Bevan, M. D., Eaton, S. A., and Bolam, J. P. (1997). Synaptic targets of physiologically, neurochemically and morphologically characterized neurons of the rat globus pallidus. In *Society of Neuroscience Abstract*, volume 23, page 196. 8
- [Bloch et al., 2016] Bloch, A., Tamir, D., Vakil, E., and Zeilig, G. (2016). Specific deficit in implicit motor sequence learning following spinal cord injury. *PLoS ONE*, 11(6) :e0158396. 82
- [Boraud, 2015a] Boraud, T. (2015a). *Matière à décision*, chapter 7, pages 97–107. CNRS Éditions. 10
- [Boraud, 2015b] Boraud, T. (2015b). *Matière à décision*, chapter 8, pages 109–117. CNRS Éditions. 11
- [Bouton, 2004] Bouton, M. E. (2004). Context and behavioral processes in extinction. *Learning & Memory*, 11(5) :485–494. 81
- [Breese et al., 2005] Breese, G. R., Knapp, D. J., Criswell, H. E., Moy, S. S., Papadeas, S. T., and Blake, B. L. (2005). The neonate-6-hydroxydopamine-lesioned rat : a model for clinical neuroscience and neurobiological principles. *Brain Research Reviews*, 48(1) :57–73. 23
- [Breton et al., 2000] Breton, L., Zucker, J.-D., and Clément, E. (2000). A multi-agent based simulation of sand piles in a static equilibrium. In *Multi-agent-based simulation*, pages 108–118. Springer. 31
- [Brette, 2015] Brette, R. (2015). Philosophy of the spike : Rate-based vs. spike-based theories of the brain. *Frontiers in Systems Neuroscience*, 9(151) :1–14. 38, 85

- [Brette and Gerstner, 2005] Brette, R. and Gerstner, W. (2005). Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. *Journal of Neurophysiology*, 94 :3637–3642. 38
- [Brette et al., 2007] Brette, R., Rudolph, M., Carnevale, T., Hines, M., Beeman, D., Bower, J. M., Diesmann, M., Morrison, A., Goodman, P. H., Jr., F. C. H., Zirpe, M., Natschläger, T., Pecevski, D., and Mikael Djurfeldt, B. E., Lansner, A., Rochel, O., Vieville, T., Muller, E., Davison, A. P., Boustani, S. E., and Destexhe, A. (2007). Simulation of networks of spiking neurons : A review of tools and strategies. *Journal of Computational Neuroscience*, 23(3) :349–398. 41
- [Brown et al., 1998] Brown, L. L., Smith, D. M., and Goldbloom, L. M. (1998). Organizing principles of cortical integration in the rat neostriatum : corticostriate map of the body surface is an ordered lattice of curved laminae and radial points. *Journal of Computational Neurology*, 392(4) :468–488. 8, 88
- [Buelow et al., 2015] Buelow, M. T., Amick, M. M., Queller, S., Stout, J. C., Friedman, J. H., and Grace, J. (2015). Feasibility of use of probabilistic reversal learning and serial reaction time tasks in clinical trials of parkinson’s disease. *Parkinsonism and Related Disorders*, 21 :894–898. 77
- [Burguière et al., 2013] Burguière, E., Monteiro, P., Feng, G., and Graybiel, A. M. (2013). Optogenetic stimulation of lateral orbitofronto-striatal pathway suppresses compulsive behaviors. *Science*, 370(6137) :1243–1246. 16
- [Burlingame et al., 1973] Burlingame, A. L., Mcpherron, R. V., and Wilson, D. M. (1973). Nonheuristic computer determination of molecular structure based upon carbon-13 nuclear magnetic resonance data. branched alkanes. *Proceedings of the National Academy of Sciences*, 70(12) :3419–3422. 33
- [Calabresi et al., 2014] Calabresi, P., Picconi, B., Tozzi, A., Ghiglieri, V., and Filippo, M. D. (2014). Direct and indirect pathways of basal ganglia : a critical reappraisal. *Nature neuroscience*, 17(8) :1022–1030. 13
- [Carnevale and Hines, 2009] Carnevale, N. T. and Hines, M. L. (2009). *The NEURON Book*. Cambridge University Press. 42
- [Chang et al., 1981] Chang, H. T., Wilson, C. J., and Kitai, S. T. (1981). Single neostriatal efferent axons in the globus pallidus : a light and electron microscopic study. *Science*, 313 :915–918. 8
- [Chen et al., 2010] Chen, H., Zhuang, P., Miao, S.-H., Yuan, G., Zhang, Y.-Q., and Li, J.-Y. (2010). Neuronal firing in the ventrolateral thalamus of patients with parkinson’s disease differs from that with essential tremor. *Chinese Medical Journal*, 123 (6) :695–701. 66
- [Chersi et al., 2013] Chersi, F., Mirolli, M., Pezzulo, G., and Baldassarre, G. (2013). A spiking neuron model of the cortico-basal ganglia circuits for goal-directed and habitual action learning. *Neural Networks*, 41 :212–224. 50, 86
- [Chiu et al., 2016] Chiu, P.-Y., Tsai, C.-T., Chen, P.-K., Chen, W.-J., and Lai, T.-J. (2016). Neuropsychiatric symptoms in parkinson’s disease dementia are more similar to alzheimer’s disease than dementia with lewy bodies : A case-control study. *PLoS ONE*, 11(4) :1–10. 13

- [Chou et al., 2015] Chou, T.-S., Bucci, L. D., and Krichmar, J. L. (2015). Learning touch preferences with a tactile robot using dopamine modulated stdp in a model of insular cortex. *Frontiers in Neurorobotics*, 9(6) :1–18. 86
- [Coetzee et al., 2016] Coetzee, S. G., Pierce, S., Brundin, P., Brundin, L., Hazelett, D. J., and Coetzee, G. A. (2016). Enrichment of risk snps in regulatory regions implicate diverse tissues in parkinson’s disease etiology. *Scientific Reports*, 6 :30509. 13
- [Cohen et al., 1997] Cohen, D. J., Leckman, J. F., and Pauls, D. (1997). Neuropsychiatric disorders of childhood : Tourette’s syndrome as a model. *Acta Paediatrica*, 86(S422) :106–111. 16
- [Cohen and Frank, 2009] Cohen, M. X. and Frank, M. J. (2009). Neurocomputational models of basal ganglia function in learning, memory and choice. *Behavioral Brain Research*, 199 (1) :141–156. 4
- [Cools et al., 2002] Cools, R., Clark, L., Owen, A. M., and Robbins, T. W. (2002). Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *The Journal of Neuroscience*, 22(11) :4563–4567. 77
- [Cools et al., 2009] Cools, R., Frank, M. J., Gibbs, S. E., Miyakawa, A., Jagust, W., and D’Esposito, M. (2009). Striatal dopamine predicts outcome-specific reversal learning and its sensitivity to dopaminergic drug administration. *Journal of Neuroscience*, 29(5) :1538–1543. 63
- [Costa et al., 2015] Costa, V. D., Tran, V. L., Turchi, J., and Averbeck, B. B. (2015). Reversal learning and dopamine : A bayesian perspective. *The Journal of Neuroscience*, 35(6) :2407–2416. 77, 78
- [Cowan and Wilson, 1994] Cowan, R. H. and Wilson, C. J. (1994). Spontaneous firing patterns and axonal projections of single corticostriatal neurons in the rat medial agranular cortex. *Journal of Neurophysiology*, 71(1) :17–32. 88
- [de la Tourette, 1899] de la Tourette, G. (1899). La maladie des tics convulsifs. *La Semaine Médicale*, 19 :153–156. 15
- [DeLong, 1990] DeLong, M. R. (1990). Primate models of movement disorders of basal ganglia origin. *TINS*, 13(7) :281–285. 8, 12, 43
- [Delvolvé et al., 1997] Delvolvé, I., Bem, T., and Cabelguen, J.-M. (1997). Epaxial and limb muscle activity during swimming and terrestrial stepping in the adult newt, *Pleurodeles waltl*. *Journal of Neurophysiology*, 78(2) :638–650. 100
- [Dombrovski et al., 2015] Dombrovski, A. Y., Szanto, K., Clark, L., Aizenstein, H. J., Chase, H. W., 3rd, C. F. R., and Siegle, G. J. (2015). Corticostriatothalamic reward prediction error signals and executive control in late-life depression. *Psychological Medicine*, 45(7) :1413–1424. 77
- [Douglas and Adler, 1973] Douglas, H. and Adler, K. (1973). Spatial orientation by salamanders using plane-polarized light. *Science*, 181 :285–287. 92, 95
- [Doya, 1999] Doya, K. (1999). What are the computations of the cerebellum, the basal ganglia and the cerebral cortex ? *Neural Networks*, 12 :961–974. 26, 27, 28, 30, 31, 32

- [Doya, 2000] Doya, K. (2000). Complementary roles of basal ganglia and cerebellum in learning and motor control. *Current opinion in neurobiology*, 10(6) :732–739. 26, 28, 31, 32
- [Doya, 2008] Doya, K. (2008). Modulators of decision-making. *Nature Neuroscience*, 11(4) :410–416. 30, 31
- [Dunovan and Verstynen, 2016] Dunovan, K. and Verstynen, T. (2016). Believer-skeptic meets actor-critic : Rethinking the role of basal ganglia pathways during decision-making and reinforcement learning. *Frontiers in neuroscience*, 10(106) :1–15. 6, 7
- [Eppler et al., 2009] Eppler, J. M., Helias, M., Muller, E., Diesmann, M., and Gewaltig, M.-O. (2009). Pynest : A convenient interface to the nest simulator. *Frontiers in Neuroinformatics*, 2(12) :1–12. 43
- [Ermentrout, 2002] Ermentrout, B. (2002). *Simulating, analyzing, and animating dynamical systems : a guide to XPPAUT for researchers and students*. Siam. 42
- [Felleman and Essen, 1991] Felleman, D. J. and Essen, D. C. V. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, 1(1) :1–47. 9
- [Fitzhugh, 1955] Fitzhugh, R. (1955). Mathematical models of threshold phenomena in the nerve membrane. *The bulletin of mathematical biophysics*, 17(4) :257–278. 39
- [Flaherty and Graybiel, 1991] Flaherty, A. W. and Graybiel, A. M. (1991). Corticostriatal transformations in the primate somatosensory system. projections from physiologically mapped body-part representations. *Journal of Neurophysiology*, 66(4) :1249–1263. 88
- [Frank, 2005] Frank, M. J. (2005). Modulation in the basal ganglia : A neurocomputational account of cognitive deficits in medicated and nonmedicated parkinsonism. *Journal of Cognitive Neuroscience*, 17(1) :51–72. 86
- [Frank et al., 2001] Frank, M. J., Loughry, B., and O’Reilly, R. C. (2001). Interactions between frontal cortex and basal ganglia in working memory : A computational model. *Cognitive, Affective and Behavioral Neuroscience*, 1(2) :137–160. 55, 86
- [Freake and Phillips, 2005] Freake, M. J. and Phillips, J. B. (2005). Light-dependent shift in bullfrog tadpole magnetic compass orientation : Evidence for a common magnetoreception mechanism in anuran and urodele amphibians. *Ethology*, 111(3) :241–254. 93
- [Féger et al., 2009] Féger, J., Worbe, Y., Galineau, L., and Tremblay, L. (2009). Troubles psychiatriques et ganglions de la base : une validation expérimentale. *Annales Pharmaceutiques Françaises*, 67 :320–334. 16
- [Galvan et al., 2015] Galvan, A., Devergnas, A., and Wichmann, T. (2015). Alterations in neuronal activity in basal ganglia-thalamocortical circuits in the parkinsonian state. *Frontiers in Neuroanatomy*, 9(5) :1–21. 13
- [Garenne et al., 2011] Garenne, A., Pasquereau, B., Guthrie, M., Bioulac, B., and Boraud, T. (2011). Basal ganglia preferentially encode context dependent choice in a two-armed bandit task. *Frontiers in Systems Neuroscience*, 2 :23. 66, 113
- [Gerfen, 1985] Gerfen, C. R. (1985). Compartmental organization of projections from the striatum to the substantia nigra in the rat. *Journal of Computational Neurology*, 236(4) :454–476. 8

- [Giraudoux, 2016] Giraudoux, P. (2016). *pgirmess* : Data analysis in ecology. *R package version 1.6.4*. 57
- [Goodman and Brette, 2009] Goodman, D. F. M. and Brette, R. (2009). The brian simulator. *Frontiers in Neuroscience*, 3(2) :1–6. 42
- [Goodman and Packard, 2016] Goodman, J. and Packard, M. G. (2016). The memory system engaged during acquisition determines the effectiveness of different extinction protocols. *Frontiers in Behavioral Neuroscience*, 9(314) :1–13. 80
- [Grassia et al., 2011] Grassia, F., Buhry, L., Lévi, T., Tomas, J., Destexhe, A., and Saïghi, S. (2011). Tunable neuromimetic integrated system for emulating cortical neuron models. *Frontiers in Neuroscience*, 5 :134. 65
- [Graybiel et al., 1994] Graybiel, A. M., Aosaki, T., Flaherty, A. W., and Kimura, M. (1994). The basal ganglia and adaptive motor control. *Science*, 265(5180) :1826–1831. 87
- [Grillner et al., 2013] Grillner, S., Robertson, B., and Stephenson-Jones, M. (2013). Independent circuits in the basal ganglia for the evaluation and selection of actions. *Proc. Natl. Acad. Sci. USA*, 110(38) :E3687–9. 92
- [Groves, 1983] Groves, P. M. (1983). A theory of the functional organization of the neostriatum and the neostriatal control of voluntary movement. *Brain research reviews*, 286 :109–132. 86
- [Gurney et al., 2001a] Gurney, K., Prescott, T. J., and Redgrave, P. (2001a). A computational model of action selection in the basal ganglia. i. a new functional anatomy. *Biological Cybernetics*, 84(6) :401–410. 2, 45, 46, 55, 87
- [Gurney et al., 2001b] Gurney, K., Prescott, T. J., and Redgrave, P. (2001b). A computational model of action selection in the basal ganglia. ii. analysis and simulation behavior. *Biological Cybernetics*, 84(6) :411–423. 45, 47
- [Gusella et al., 1983] Gusella, J. F., Wexler, N. S., Conneally, P. M., Naylor, S. L., Anderson, M. A., Tanzi, R. E., Watkins, P. C., Ottina, K., Wallace, M. R., Sakaguchi, A. Y., Young, A. B., and Shoulson, I. (1983). A polymorphic dna marker genetically linked to huntington’s disease. *Nature*, 306 :234–238. 15
- [Guthrie et al., 2013] Guthrie, M., Leblois, A., Garenne, A., and Boraud, T. (2013). Interaction between cognitive and motor cortico-basal ganglia loops during decision making : A computational study. *Journal of Neurophysiology*, pages 1522–1598. xviii, 6, 8, 22, 48, 49, 57, 58, 59, 64, 65, 75, 82, 87, 89, 90, 113, 120
- [Haber and Calzavara, 2009] Haber, S. N. and Calzavara, R. (2009). The cortico-basal ganglia integrative network : The role of the thalamus. *Brain Research Bulletin*, 78 :69–74. 12
- [Hamker, 2013] Hamker, H. S. F. H. (2013). Computational models of basal-ganglia pathway functions : Focus on functional neuroanatomy. *Frontiers in Systems Neuroscience*, 7(122). 7, 8
- [Hazrati et al., 1995] Hazrati, L.-N., Pinault, D., and Parent, A. (1995). The thalamic reticular nucleus does not send commissural projection to the contralateral parafascicular nucleus in the rat. *Brain Research*, 679(1) :123–134. 2, 55

- [Herculano-Houzel, 2009] Herculano-Houzel, S. (2009). The human brain in numbers : a linearly scaled-up primate brain. *Frontiers in Human Neuroscience*, 3(31) :1–11. 8
- [Hill and Walsh, 2005] Hill, R. S. and Walsh, C. A. (2005). Molecular insights into human brain evolution. *Nature*, 437 :64–67. 9, 10
- [Hines et al., 2009] Hines, M. L., Davison, A. P., and Muller, E. (2009). Neuron and python. *Frontiers in Neuroinformatics*, 3(1) :1–12. 42
- [Hines et al., 2008] Hines, M. L., Markram, H., and Schurmann, F. (2008). Fully implicit parallel simulation of single neurons. *Journal of computational neuroscience*, 25 :439–448. 42, 56, 60
- [Hodgkin and Huxley, 1952] Hodgkin, A. L. and Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology*, 117(4) :500–544. 40
- [Houghton, 2014] Houghton, C. (2014). Supervised learning with complex spikes and spike-timing-dependent plasticity. *PloS one*, 9(6) :e99635. 28
- [Houk et al., 1994] Houk, J., D., J. D., and Beiser (1994). *Neostriatal Circuitry as a Scalar Memory : Modeling and Ensemble Neuron Recording*, pages 315–336. MIT Press. 86
- [Houk et al., 1995] Houk, J. C., Adams, J. L., and Barto, A. G. (1995). A model of how the basal ganglia generate and use neural signals that predict reinforcement. *Models of Information Processing in the Basal Ganglia*, pages 249–270. 30
- [Hu and Wang, 2016] Hu, S. and Wang, G. (2016). Mitochondrial dysfunction in parkinson’s disease. *Translational Neurodegeneration*, 5(1) :1–8. 13
- [Humphries et al., 2006] Humphries, M. D., Stewart, R. D., and Gurney, K. N. (2006). A physiologically plausible model of action selection and oscillatory activity in the basal ganglia. *The Journal of Neuroscience*, 26(50) :12921–12942. 41, 47, 48, 62
- [Hunter, 2007] Hunter, J. D. (2007). Matplotlib : A 2d graphics environment. *Computing in Science and Engineering*, 9 :90–95. 57
- [Izhikevich, 2000] Izhikevich, E. M. (2000). Neural excitability, spiking and bursting. *International Journal of Bifurcation and Chaos*, 10(6) :1171–1266. 89
- [Izhikevich, 2003] Izhikevich, E. M. (2003). Simple model of spiking neurons. *IEEE Transactions on Neural Networks*, 14(6) :1569–1572. 38, 39
- [Jacobs et al., 2009] Jacobs, A. L., Fridman, G., Douglas, R. M., Alam, N. M., Latham, P. E., Prusky, G. T., and Nirenberg, S. (2009). Ruling out and ruling in neural codes. *PNAS*, 106(14) :5936–5941. 85
- [Jaeger et al., 1994] Jaeger, D., Kita, H., and Wilson, C. J. (1994). Surround inhibition among projection neurons is weak or nonexistent in the rat neostriatum. *Journal of Neurophysiology*, 72(5) :2555–2558. 86
- [Jain et al., 1996] Jain, A. K., Mao, J., and Mohiuddin, K. M. (1996). Artificial neural networks : A tutorial. *IEEE computer*, 29(3) :31–44. 26, 27, 28, 29, 31, 33, 35

- [Jason T. Moyer and Wolf, 2014] Jason T. Moyer, Benjamin L. Halterman, L. H. F. and Wolf, J. A. (2014). Lateral and feedforward inhibition suppress asynchronous activity in a large, biophysically-detailed computational model of the striatal network. *Frontiers in Computational Neuroscience*, 8(152) :1–13. 58, 59, 76, 85, 86
- [Jones et al., 2001] Jones, E., Oliphant, T., Peterson, P., et al. (2001). Scipy : Open source scientific tools for python. URL <http://www.scipy.org>, 73 :86. 57
- [Kawato and Gomi, 1992] Kawato, M. and Gomi, H. (1992). A computational model of four regions of the cerebellum based on feedback-error learning. *Biological Cybernetics*, 68 :95–103. 27
- [Kerr and Wickens, 2001] Kerr, J. N. D. and Wickens, J. R. (2001). Dopamine d-1/d-5 receptor activation is required for long-term potentiation in the rat neostriatum in vitro. *Journal of Neurophysiology*, 85 :117–124. 63, 88
- [Kita and Kita, 1994] Kita, H. and Kita, S. T. (1994). The morphology of globus pallidus projection neurons in the rat : an intracellular staining study. *Brain Research*, 636(2) :308–319. 8
- [Klopf, 1988] Klopf, H., editor (1988). *Neural Models of Plasticity*. John H. Byrne et William O. Berry. 64
- [Koos et al., 2004] Koos, T., Tepper, J., and Wilson, C. J. (2004). Comparison of ipscs evoked by spiny and fast-spiking neurons in the neostriatum. *Journal of Neuroscience*, 24(36) :7916–7922. 86
- [Koós and Tepper, 1999] Koós, T. and Tepper, J. M. (1999). Inhibitory control of neostriatal projection neurons by gabaergic interneurons. *Nature neuroscience*, 2(5) :467–472. 4
- [Kotter and Wickens, 1998] Kotter, R. and Wickens, J. R. (1998). Striatal mechanisms in parkinson’s disease : new insights from computer modeling. *Artificial Intelligence in Medicine*, 13 :37–55. 86
- [Lambot et al., 2016] Lambot, L., Rodriguez, E. C., Houtteman, D., Li, Y., Schiffmann, S. N., Gall, D., and de Kerchove d’Exaerde, A. (2016). Striatopallidal neuron nmda receptors control synaptic connectivity, locomotor, and goal-directed behaviors. *The Journal of Neuroscience*, 36(18) :4976–4992. 6, 8, 30
- [Lapicque, 1905] Lapicque, M. (1905). *Recherches sur l’excitabilité électrique de différents muscles de vertébrés et d’invertébrés*. Laroche-Delattre. 37
- [Le et al., 2013] Le, Q. V., Ranzato, M., Monga, R., Devin, M., Chen, K., Corrado, G. S., Dean, J., and Ng, A. Y. (2013). Building high-level features using large scale unsupervised learning. In *2013 IEEE international conference on acoustics, speech and signal processing*, pages 8595–8598. IEEE. 31, 32
- [Leblois et al., 2006a] Leblois, A., Boraud, T., Meissner, W., Bergman, H., and Hansel, D. (2006a). Competition between feedback loops underlies normal and pathological dynamics in the basal ganglia. *Journal of Neuroscience*, 26(13) :3567–3583. xviii, 7, 45, 47, 48, 49, 55, 57, 58, 89, 113, 120

- [Leblois et al., 2006b] Leblois, A., Meissner, W., Bezard, E., Bioulac, B., Gross, C. E., and Boraud, T. (2006b). Temporal and spatial alterations in gpi neuronal encoding might contribute to slow down movement in parkinsonian monkeys. *European Journal of Neuroscience*, 24(4) :1460–9568. 86, 87
- [LeCun et al., 2015] LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, 521(7553) :436–444. 28
- [Lee et al., 2012] Lee, D., Seo, H., and Jung, M. W. (2012). Neural basis of reinforcement learning and decision making. *Annual Review of Neuroscience*, 35 :287–308. 2, 29, 55
- [Levy et al., 1997] Levy, R., Hazrati, L.-N., Herrero, M.-T., Vila, M., Hassani, O.-K., Mouroux, M., Ruberg, M., Asensi, H., Agid, Y., Féger, J., Obeso, J., Parent, A., and Hirsch, E. (1997). Re-evaluation of the functional anatomy of the basal ganglia in normal and parkinsonian states. *Neuroscience*, 76(2) :335–343. 8
- [Lindskog et al., 2006] Lindskog, M., Kim, M., Wikstrom, M. A., Blackwell, K. T., and Kotaleski, J. H. (2006). Transient calcium and dopamine increase pka activity and darpp-32 phosphorylation. *PLoS Comput. Biol.*, 2(9) :e119. 59
- [Lingawi and Balleine, 2012] Lingawi, N. W. and Balleine, B. W. (2012). Amygdala central nucleus interacts with dorsolateral striatum to regulate the acquisition of habits. *The Journal of Neuroscience*, 32(3) :1073–1081. 81
- [Liénard, 2013] Liénard, J. (2013). *Models of the Basal Ganglia : Study of the Functional Anatomy and Pathophysiology using Multiobjective Evolutionary Algorithms*. PhD thesis, Université Pierre et Marie Curie. 44
- [Mahdi et al., 2013] Mahdi, A., Sturdy, J., Ottesen, J. T., and Olufsen, M. S. (2013). Modeling the afferent dynamics of the baroreflex control system. *Plot Computational Biology*, 9(12) :e1003384. 38
- [Mandali et al., 2015] Mandali, A., Rengaswamy, M., Chakravarthy, S., and Moustafa, A. A. (2015). A spiking basal ganglia model of synchrony, exploration and decision making. *Frontiers in Neuroscience*, 9(191). 50, 86
- [Marr, 1969] Marr, D. (1969). A theory of cerebellar cortex. *J. Physiol.*, 202 :437–470. 27
- [Marín et al., 1997a] Marín, O., González, A., and Smeets, W. J. (1997a). Basal ganglia organization in amphibians : Afferent connections of the striatum and the nucleus accumbens. *The Journal of Comparative Neurology*, 378 :16–49. 11, 99, 108
- [Marín et al., 1997b] Marín, O., González, A., and Smeets, W. J. (1997b). Basal ganglia organization in amphibians : Efferent connections of the striatum and the nucleus accumbens. *The Journal of Comparative Neurology*, 380 :23–50. 11, 99, 108
- [Masson et al., 2014] Masson, G. L., Przedborski, S., , and Abbott, L. (2014). A computational model of motor neuron degeneration. *Neuron*, 83(4) :975–988. 40
- [McCulloch and Pitts, 1943] McCulloch, W. S. and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4) :115–133. 36, 37

- [Medina et al., 2001] Medina, J. F., Garcia, K. S., and Mauk, M. D. (2001). A mechanism for savings in the cerebellum. *The Journal of Neuroscience*, 21(11) :4081–4089. 86
- [Migliore et al., 2006] Migliore, M., Cannia, C., Lytton, W. W., Markram, H., and Hines, M. L. (2006). Parallel networks simulations with neuron. *Journal of Computational Neuroscience*, 21(2) :119–129. 42, 56
- [Mink, 1996] Mink, J. W. (1996). The basal ganglia : focused selection and inhibition of competing motor programs. *Progress in Neurobiology*, 50 :381–425. 7, 45, 55, 56, 87
- [Mink, 2004] Mink, J. W. (2004). Deep brain stimulation for treating tourette syndrome ? *TSA USA Newsletter*, 32 :5–7. 16
- [Montague et al., 1996] Montague, P. R., Dayan, P., , and Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive hebbian learning. *The Journal of Neuroscience*, 16(5) :1936–1947. 30
- [Morita et al., 2016] Morita, K., Jitsev, J., and Morrison, A. (2016). Corticostriatal circuit mechanisms of value-based action selection : Implementation of reinforcement learning algorithms and beyond. *Behavioral Brain Research*, 311 :110–121. 78
- [Morris and Lecar, 1981] Morris, C. and Lecar, H. (1981). Voltage oscillations in the barnacle giant muscle fiber. *Biophysical journal*, 35(1) :193. 39
- [Mutterer and Zinck, 2013] Mutterer, J. and Zinck, E. (2013). Quick-and-clean article figures with figurej. *Journal of Microscopy*, 252(1) :89–91. 101
- [Nakano et al., 2010] Nakano, T., Doi, T., Yoshimoto, J., and Doya, K. (2010). A kinetic model of dopamine- and calcium-dependent striatal synaptic plasticity. *PloS Computational Biology*, 6(2) :e1000670. 41
- [Nambu, 2011] Nambu, A. (2011). Somatotopic organization of the primate basal ganglia. *Frontiers in Neuroanatomy*, 5 :26. 3, 4, 5, 6, 12, 56
- [Namburi et al., 2015] Namburi, P., Beyeler, A., Yorozu, S., Calhoun, G. G., Halbert, S. A., Wichmann, R., Holden, S. S., Mertens, K. L., Anahtar, M., Felix-Ortiz, A. C., Wickersham, I. R., Gray, J. M., and Tye, K. M. (2015). A circuit mechanism for differentiating positive and negative associations. *Nature*, 520(7549) :675–678. 30
- [Neymotin et al., 2011] Neymotin, S. A., Lee, H., Park, E., Fenton, A. A., and Lytton, W. W. (2011). Emergence of physiological oscillation frequencies in a computer model of neocortex. *Frontiers in Systems Neuroscience*, 5 :19. 63
- [Nieuwenhuys et al., 1998] Nieuwenhuys, R., Nicholson, C., and Donkelaar, H. (1998). *The central nervous system of vertebrates.*, volume 2. Springer. 92
- [Obeso et al., 2008] Obeso, J. A., Rodriguez-Oroz, M. C., Benitez-Temino, B., Blesa, F. J., Guridi, J., Marin, C., , and Rodriguez, M. (2008). Functional organization of the basal ganglia : Therapeutic implications for parkinson’s disease. *Movement Disorders*, 23(Suppl. 3) :S548–559. 4
- [Oorschot, 1996] Oorschot, D. E. (1996). Total number of neurons in the neostriatal, pallidal, subthalamic, and substantia nigral nuclei of the rat basal ganglia : A stereological study using the cavalieri and optical disector methods. *Journal of Comparative Neurology*, 366 (4) :580–599. 5, 86

- [O'Reilly and Frank, 2006] O'Reilly, R. C. and Frank, M. (2006). Making working memory work : A computational model of learning in the prefrontal cortex and basal ganglia. *Neural Computation*, 18(2) :283–328. 86
- [Papachristou et al., 2013] Papachristou, H., Nederkoorn, C., Beunen, S., and Jansen, A. (2013). Dissection of appetitive conditioning. does impulsivity play a role ? *Appetite*, 69 :46–53. 80
- [Parent and Hazrati, 1995] Parent, A. and Hazrati, L.-N. (1995). Functional anatomy of the basal ganglia. ii. the place of subthalamic nucleus and external pallidum in basal ganglia circuitry. *Brain Research Reviews*, 20(1) :128–154. 3, 10, 12
- [Parent et al., 2000] Parent, A., Sato, F., Wu, Y., Gautier, J., Lévesque, M., and Parent, M. (2000). Organization of the basal ganglia : the importance of axonal collateralization. *Trends in Neurosciences*, 23 (Suppl. 10) :S20–S27. 87
- [Parent et al., 2001] Parent, M., Lévesque, M., and Parent, A. (2001). Two types of projection neurons in the internal pallidum of primates : Single- axon tracing and three-dimensional reconstruction. *The Journal of Comparative Neurology*, 439 :162–175. 5, 6
- [Parish et al., 2007] Parish, C. L., Beljajeva, A., Arenas, E., and Simon, A. (2007). Midbrain dopaminergic neurogenesis and behavioural recovery in a salamander lesion-induced regeneration model. *Development and Disease*, 134 :2881–2887. xix, 23, 24, 93, 98, 99, 100, 105, 121
- [Parthasarathy et al., 1992] Parthasarathy, H. B., Shall, J. D., and Graybiel, A. M. (1992). Distributed but convergent ordering of corticostriatal projections : analysis of the frontal eye field and the supplementary eye field in the macaque monkey. *Journal of Neuroscience*, 12(11) :4468–4488. 88
- [Pasquereau et al., 2007] Pasquereau, B., Nadjar, A., Arkadir, D., Bezdard, E., Goillandeau, M., Bioulac, B., Gross, C. E., and Boraud, T. (2007). Shaping of motor responses by incentive values through the basal ganglia. *Journal of Neuroscience*, 27(5) :1176–83. xviii, 18, 19, 20, 49, 56, 64, 66, 73, 87, 93, 113, 120
- [Pavlov, 1927] Pavlov, I. P. (1927). Conditioned reflexes : An investigation of the physiological activity of the cerebral cortex. *Annals of Neuroscience*, 17(3) :136–141. 79
- [Pawlak and Kerr, 2008] Pawlak, V. and Kerr, J. N. D. (2008). Dopamine receptor activation is required for corticostriatal spike-timing dependent plasticity. *Journal of Neuroscience*, 28(10) :2435–2446. 63, 88
- [Percheron et al., 1984] Percheron, G., Yelnik, J., and François, C. A. (1984). A golgi analysis of the primate globus pallidus. iii. spatial organization of the striato-pallidal complex. *Journal of Comparative Neurology*, 227(2) :214–227. 6, 57
- [Phillips, 1977] Phillips, J. B. (1977). Use of the earth's magnetic field by orienting cave salamanders (*Eurycea lucifuga*). *Journal of Comparative Physiology*, 121(2) :273–288. 93
- [Piochon et al., 2012] Piochon, C., Kruskal, P., MacLean, J., and Hansel, C. (2012). Non-hebbian spike-timing-dependent plasticity in cerebellar circuits. *Frontiers in Neural Circuits*, 6(124) :1–8. 28

- [Piron et al., 2016] Piron, C., Kase, D., Topalidou, M., Goillandeau, M., Orignac, H., N’Guyen, T.-H., Rougier, N., and Boraud, T. (2016). The globus pallidus pars interna in goal-oriented and routine behaviors : Resolving a long-standing paradox. *Movements Disorders*, pages 1–9. xix, 21, 22, 56, 73, 75, 81, 82, 93, 117, 120
- [Planert et al., 2013] Planert, H., Berger, T. K., and Silberberg, G. (2013). Membrane properties of striatal direct and indirect pathway neurons in mouse and rat slices and their modulation by dopamine. *PloS one*, 8(3) :e57054. 13, 30
- [Plantinga et al., 2016] Plantinga, B. R., Roebroek, A., Kemper, G., Uludag, K., Melse, M., Mai, J., Kuijf, M. L., Herrler, A., Jahanshahi, A., ter Haar, B. M., Romeny, and Temel, Y. (2016). Ultra-high field mri post mortem structural connectivity of the human subthalamic nucleus, substantia nigra, and globus pallidus. *Frontiers in Neuroanatomy*, 10(66) :1–10. 6
- [Pohlert, 2014] Pohlert, T. (2014). The pairwise multiple comparison of mean ranks package (pmcnr). *R package*. 57
- [Portelli et al., 2016] Portelli, G., Barrett, J. M., Hilgen, G., Masquelier, T., Maccione, A., Marco, S. D., Berdondini, L., Kornprobst, P., and Sernagor, E. (2016). Rank order coding : a retinal information decoding strategy revealed by large-scale multielectrode array retinal recordings. *eNeuro*, 3(3) :ENEURO–0134. 85
- [Postuma and Lang, 2003] Postuma, R. B. and Lang, A. E. (2003). Hemiballism : revisiting a classic disorder. *Lancet Neurol*, 2(11) :661–668. 15
- [R-Team, 2015] R-Team, editor (2015). *R : A Language and Environment for Statistical Computing v. 3.2.2*, R Foundation for Statistical Computing, Vienna, Austria. Vienna, Austria. 57
- [Rakic, 1988] Rakic, P. (1988). Specification of cerebral cortical areas. *Science*, 241(4862) :170–176. 9
- [Ramon y Cajal, 1894] Ramon y Cajal, S. (1894). The croonian lecture : La fine structure des centres nerveux. *Proceedings of the Royal Society of London*, 55 :444–468. 8
- [Raymond et al., 1996] Raymond, J. L., Lisberger, S. G., and Mauk, M. D. (1996). The cerebellum : A neuronal learning machine ? *Science*, 272(5265) :1126–1131. 27
- [Redgrave et al., 1999] Redgrave, P., Prescott, T. J., and Gurney, K. (1999). The basal ganglia : a vertebrate solution to the selection problem ? *Neuroscience*, 89(4) :1009–1023. 55
- [Reiner et al., 1998] Reiner, A., Medina, L., and Veenman, C. (1998). Structural and functional evolution of the basal ganglia in vertebrates. *Brain research reviews*, 28(3) :235–285. 92
- [Retailleau et al., 2013] Retailleau, A., Dejean, C., Fourneaux, B., Leinekugel, X., and Boraud, T. (2013). Why am i lost without dopamine ? effects of 6-ohda lesion on the encoding of reward and decision process in ca3. *Neurobiology of Disease*, 59 :151–164. 93
- [Reynolds et al., 2001] Reynolds, J., Hyland, B., and Wickens, J. R. (2001). A cellular mechanism of reward-related learning. *Nature*, 413(6851) :67–70. 63, 88
- [Rivest, 2009] Rivest, F. (2009). *Modèle informatique du coapprentissage des ganglions de la base et du cortex : L’apprentissage par renforcement et le développement de représentations*. PhD thesis, Université de Montréal. 31

- [Rodriguez-Oroz et al., 2009] Rodriguez-Oroz, M. C., Jahanshahi, M., Krack, P., Litvan, I., Macias, R., Bezard, E., and Obeso, J. A. (2009). Initial clinical manifestations of parkinson's disease : features and pathophysiological mechanisms. *The Lancet Neurology*, 8(12) :1128–1139. 4
- [Rosenblatt, 1958] Rosenblatt, F. (1958). The perceptron : a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6) :386–408. 33
- [Rougier and Fix, 2012] Rougier, N. P. and Fix, J. (2012). Dana : Distributed numerical and adaptive modelling framework. *Network : Computation in Neural Systems*, 23(4) :237–253. 23
- [Ryczko et al., 2015] Ryczko, D., Knüsel, J., Crespi, A., Lamarque, S., Mathou, A., Ijspeert, A. J., and Cabelguen, J. M. (2015). Flexibility of the axial central pattern generator network for locomotion in the salamander. *Journal of Neurophysiology*, 113 :1921–1940. 98
- [Saal et al., 2015] Saal, H. P., Wang, X., and Bensmaia, S. J. (2015). Importance of spike timing in touch : an analogy with hearing ? *Current Opinion in Neurobiology*, 40 :142–149. 85
- [Sandoz et al., 2016] Sandoz, J.-C., , and Pham-Delègue, M.-H. (2016). Spontaneous recovery after extinction of the conditioned proboscis extension response in the honeybee. *Learning & memory*, 11 :586–597. 81
- [SandraDiaz-Pier et al., 2016] SandraDiaz-Pier, Naveau, M., Butz-Ostendorf, M., and Morrison, A. (2016). Automatic generation of connectivity for large-scale neuronal network models through structural plasticity. *Frontiers in Neuroanatomy*, 10(57) :1–15. 41
- [Sanger, 1989] Sanger, T. D. (1989). Optimal unsupervised learning in a single-layer linear feedforward neural network. *Neural Networks*, 2 :459–473. 32
- [Sarvestani et al., 2013] Sarvestani, I. K., Kozlov, A., Harischandra, N., Grillner, S., and Örjan Ekeberg (2013). A computational model of visually guided locomotion in lamprey. *Biological Cybernetics*, 107(5) :497–512. 56, 92
- [Schlegel et al., 2006] Schlegel, P. A., Briegleb, W., Bulog, B., and Steinfartz, S. (2006). Revue et nouvelles données sur la sensibilité à la lumière et orientation non-visuelle chez proteus anguinus, calotriton asper et desmognathus ochrophaeus (amphibiens urodèles hypogés). *Bull. Soc. Herp. Fr*, 118 :1–31. 93
- [Schroll et al., 2015] Schroll, H., Beste, C., and Hamker, F. H. (2015). Combined lesions of direct and indirect basal ganglia pathways but not changes in dopamine levels explain learning deficits in patients with Huntington's disease. *Eur. J. Neurosci.*, 41(9) :1227–1244. 8, 59
- [Schroll et al., 2012] Schroll, H., Vitay, J., and Hamker, F. H. (2012). Working memory and response selection : A computational account of interactions among cortico-basalganglio-thalamic loops. *Neural Network*, 26 :59–74. 89
- [Schroll et al., 2014] Schroll, H., Vitay, J., and Hamker, F. H. (2014). Dysfunctional and compensatory synaptic plasticity in parkinson's disease. *Eur. J. Neurosci.*, 39(4) :688–702. 8
- [Schultz et al., 1997] Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275 :1593–1899. 29, 30, 31

- [Schutter, 2009] Schutter, E. D. (2009). *Computational modeling methods for neuroscientists*. The MIT Press. 41
- [Shepherd, 2011] Shepherd, G. M. (2011). The microcircuit concept applied to cortical evolution : from three-layer to six-layer cortex. *Frontiers in Systems Neuroanatomy*, 5(30) :1–15. 10
- [Shohamy et al., 2009] Shohamy, D., Myers, C. E., Hopkins, R. O., Sage, J., and Gluck, M. A. (2009). Distinct hippocampal and basal ganglia contributions to probabilistic learning and reversal. *The Journal of Cognitive Neuroscience*, 21(9) :1821–1833. 78
- [Shortliffe, 1976] Shortliffe, E. H. (1976). *Computer-based medical consultations : MYCIN*. Elsevier. 33
- [Skinner, 1979] Skinner, B., editor (1979). *The Shaping of a Behaviorist : Part Two of an Autobiography*. Alfred A. Knopf. 79
- [Smith et al., 1998] Smith, Y., Bevan, M. D., Shink, E., and Bolam, J. P. (1998). Microcircuitry of the direct and indirect pathways of the basal ganglia. *Neuroscience*, 86(2) :353–388. 7, 8
- [Smith and Bolam, 1989] Smith, Y. and Bolam, J. P. (1989). Neurons of the substantia nigra reticulata receive a dense gaba-containing input from the globus pallidus in the rat. *Brain Research*, 493 :160–167. 8
- [Stein, 2002] Stein, D. J. (2002). Obsessive-compulsive disorder. *The Lancet*, 360(9330) :397–405. 16
- [Stephenson-Jone et al., 2013] Stephenson-Jone, M., Kardamakis, A., Robertson, B., and Grillner, S. (2013). Independent circuits in the basal ganglia for the evaluation and selection of actions. *Proc Natl Acad Sci USA*, 110(2) :E3670–9. 92
- [Sukumar et al., 2012] Sukumar, D., Rengaswamy, M., and Chakravarthy, S. V. (2012). Modeling the contributions of basal ganglia and hippocampus to spatial navigation using reinforcement learning. *PLoS One*, 7(10) :e47467. 55
- [Suri and Schultz, 1999] Suri, R. E. and Schultz, W. (1999). A neural network model with dopamine-like reinforcement signal that learns a spatial delayed response task. *Neuroscience*, 91 (3,) :871–890. 55, 86
- [Sutton, 1984] Sutton, R. S. (1984). *Temporal credit assignment in reinforcement learning*. PhD thesis, Graduate School of the University of Massachusetts. 29
- [Sutton and Barto, 1998] Sutton, R. S. and Barto, A. G. (1998). *Reinforcement learning : An introduction*, volume 1. MIT Press Cambridge. 29, 30
- [Takada et al., 2001] Takada, M., Tokunu, H., Hamada, I., Inase, M., Ito, Y., Imanishi, M., Hasegawa, N., Akazawa, T., Hatanaka, N., and Nambu, A. (2001). Organization of inputs from cingulate motor areas to basal ganglia in macaque monkey. *European Journal of Neuroscience*, 14(10) :1633–1650. 88
- [Terman et al., 2002] Terman, D., Rubin, J. E., Yew, A. C., and Wilson, C. J. (2002). Activity patterns in a model for the subthalamopallidal network of the basal ganglia. *Journal of Neuroscience*, 22(7) :2963–2976. 60

- [Touzet, 1992] Touzet, C. (1992). *Les Réseaux de Neurones Artificiels : Introduction au Connexionnisme*, volume 1. EC2. 2
- [Tuckwell, 1989] Tuckwell, H. C., editor (1989). *CBMS-NSF Regional Conference Series in Applied Mathematics*, volume 56, Society for Industrial and Applied Mathematics, Philadelphia. SIAM. 70
- [Tunstall et al., 2002] Tunstall, M., Oorschot, D. E., Kean, A., and Wickens, J. R. (2002). Inhibitory interactions between spiny projection neurons in the rat striatum. *Journal of Neurophysiology*, 88(3) :1263–1269. 86
- [Umemura et al., 2016] Umemura, A., Oyama, G., Shimo, Y., Nakajima, M., Nakajima, A., Jo, T., Sekimoto, S., Ito, M., Mitsunashi, T., Hattori, N., and Arai, H. (2016). Current topics in deep brain stimulation for parkinson disease. *Neurologia medico-chirurgica*, 0 :1–13. 13
- [Utter and Basso, 2008] Utter, A. A. and Basso, M. A. (2008). The basal ganglia : An overview of circuits and function. *Neuroscience and Biobehavioral Reviews*, 32(3) :333 – 342. 5, 6, 13, 15, 16
- [Von der Malsburg, 1973] Von der Malsburg, C. (1973). Self-organization of orientation sensitive cells in the striate cortex. *Kybernetik*, 14(2) :85–100. 32
- [Wang et al., 2011] Wang, X., Vaingankar, V., Sanchez, C. S., Sommer, F. T., and Hirsch, J. A. (2011). Thalamic interneurons and relay cells use complementary synaptic mechanisms for visual processing. *Nature Neuroscience*, 14(2) :224–231. 117
- [Webster, 1961] Webster, K. E. (1961). Cortico-striate interrelations in the albino rat. *Journal of Anatomy*, 95(Pt 4) :532–545. 88
- [Wei et al., 2015] Wei, W., Rubin, J. E., and Wang, X. J. (2015). Role of the indirect pathway of the basal ganglia in perceptual decision making. *J. Neurosci.*, 35(9) :4052–4064. 89
- [Wendling et al., 2016] Wendling, F., Benquet, P., Bartolomei, F., and Jirsa, V. (2016). Computational models of epileptiform activity. *Journal of Neuroscience Methods*, 260 :233–251. xix
- [Wichmann and DeLong, 1996] Wichmann, T. and DeLong, M. R. (1996). Functional and pathophysiological models of the basal ganglia. *Current Opinion in Neurobiology*, 6 :751–758. 13, 15
- [Wickens, 1997] Wickens, J. (1997). Basal ganglia : structure and computations. *Network : Computation in Neural Systems*, 8(4) :R77–R109. 4
- [Wickham, 2009] Wickham, H. (2009). *ggplot2 : Elegant graphics for data analysis*. Springer Science and Business Media. 57
- [Wilke, 2016] Wilke, C. O. (2016). *cowplot : Streamlined plot theme and plot annotations for 'ggplot2'*. *R package version 0.6.2*. 57
- [Wilson, 1987] Wilson, C. J. (1987). Morphology and synaptic connections of crossed corticostriatal neurons in the rat. *Journal of Comparative Neurology*, 263(4) :567–580. 88
- [Wilson, 1995] Wilson, C. J. (1995). The contribution of cortical neurons to the firing pattern of striatal spiny neurons. In Houk, J. C., Davis, J. L., and Beiser, D. G., editors, *Models of*

information processing in the basal ganglia., pages 141–171. The MIT Press, Cambridge, MA. 4

[Xue et al., 2013] Xue, G., Xue, F., Drouman, V., Lu, Z.-L., Bechara, A., and Read, S. (2013). Common neural mechanisms underlying reversal learning by reward and punishment. *plus One*, 8(12) :e82169. 78

[Yin and Knowlton, 2006] Yin, H. H. and Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nature Reviews Neuroscience*, 7(6) :464–476. xix

Annexes

Sommaire

A.1	Glossaire	xxiii
A.2	Circuiterie détaillée du modèle	xxiv
A.3	Tables supplémentaires	xxv
A.4	Publications	xxviii
A.5	Posters	xxviii

A.1 Glossaire

Action Mouvement ou absence de mouvement, tout ou partie d'un processus tendant à un but.

Apprentissage Initiation par l'expérience à une activité. Ensemble des processus de mémorisation mis en œuvre par un sujet pour élaborer ou modifier les schèmes comportementaux spécifiques sous l'influence de son environnement et de son expérience ¹.

Apprentissage procédural Fait d'apprendre, stocker et utiliser des procédures perceptuelles, sensorimotrices et cognitives (comme conduire un véhicule, faire du vélo ou taper à l'ordinateur) ².

Sélection de l'action Tâche consistant à décider de l'action (ou de la séquence d'actions) à effectuer afin de réaliser un objectif, en se basant sur un répertoire d'actions disponibles, une certaine connaissance de son état interne ou encore sur des informations sensorielles concernant le contexte environnemental ³.

Somatotopie Représentation associée à une cartographie où une population de neurones dans une région spécifique du cerveau va être associée à une autre partie du système nerveux ou à une fonction ⁴.

Mot Groupe de caractère ou d'éléments binaires considérés comme une entité, susceptible d'une transition écrite individualisée et formant une unité de sens ⁵.

1. <http://www.larousse.fr/dictionnaires/francais/apprentissage/4748>

2. http://theses.univ-lyon2.fr/documents/getpart.php?id=lyon2.2010.bussy_g

3. http://www.scholarpedia.org/article/Action_selection

4. <http://medecine.univ-batna.dz/medecine/images/pdf/14-15/cours/Regul%20motricite%20par%20le%20cerveau%20PDF%20%2015.pdf>

5. <http://www.larousse.fr/dictionnaires/francais/mot/52767>

A.2 Circuiterie détaillée du modèle

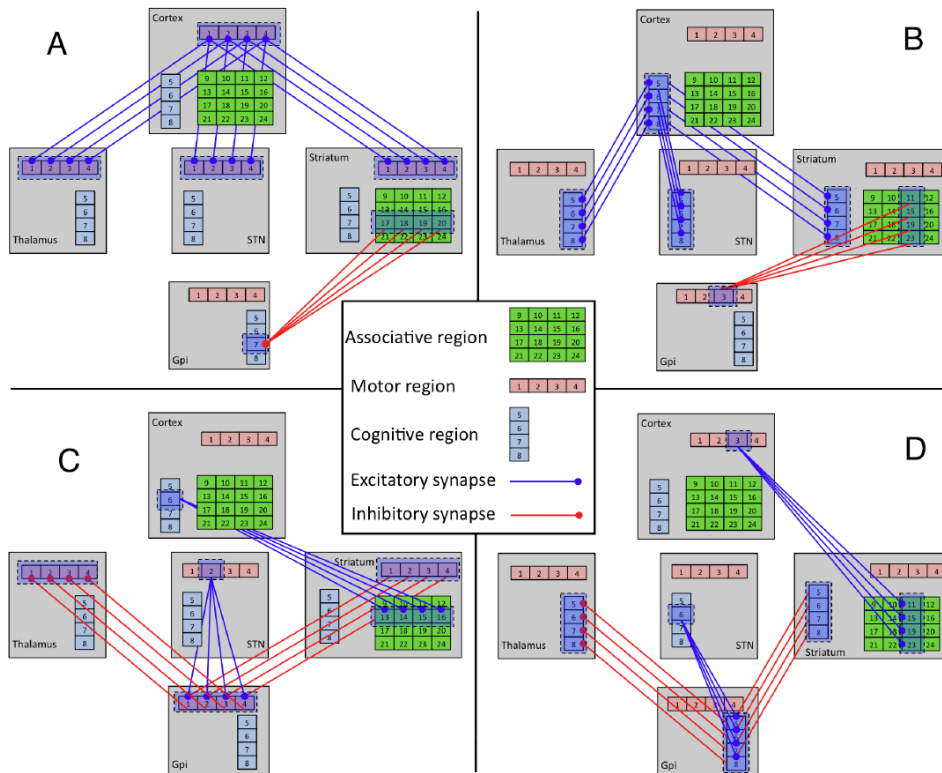


FIGURE A.1 – Circuiterie détaillée du modèle. Les boucles cortico-basales sont modélisées en prenant compte des deux niveaux de la prise de décision : un boucle cognitive (cellules en bleu) et une boucle motrice (cellules en rouges clair). Les régions associatives corticales et striatales sont représentées en vert. Une cellule représente une sous-population neuronale (par exemple 100 neurones pour une cellule corticale sur ce diagramme). **(A)** Projections des régions motrices corticales vers leurs analogues motrices au niveau du thalamus, STN et du striatum. Le striatum associatif va projeter vers le Gpi cognitif d'une manière convergente focalisée. **(B)** Projections des régions cognitives du cortex vers leurs analogues au niveau du striatum, thalamus et du STN. Le striatum associatif va projeter vers le Gpi moteur de manière convergente et focalisée. **(C)** Le cortex cognitif va projeter vers le striatum associatif de manière divergente et avec un motif de projection horizontal (« Horizontal Projection Pattern », HPP). Le STN moteur projette dans le Gpi moteur également de manière divergente, la spécificité d'un canal d'une cellule à une autre va être perdue. Le striatum moteur projette dans ses analogues pallidaux, qui à leur tour projettent dans leurs analogues thalamiques (avec cette fois-ci un maintien de la spécificité). **(D)** Le cortex moteur projette vers le striatum associatif d'une manière divergente et avec un motif de projection vertical (« Vertical Projection Pattern », VPP). Le STN cognitif projette vers le Gpi cognitif de manière divergente, où la spécificité d'une cellule à l'autre sera perdue. Le striatum cognitif projette dans ses analogues pallidaux, qui à leur tour projettent dans leurs analogues des structures thalamiques. Les biais entre les boucles motrice et cognitive du modèle sont pour une grande partie dus à l'effet croisé des HPP et des VPP. De plus (non montré sur la figure), le cortex et striatum associatif sont connectés avec un ratio de 1 : 1 et les effets des interneurons striataux inhibiteurs ont été pris en compte en fournissant une inhibition latérale des régions striatales. Chaque boucle consiste en une voie directe et focalisée de rétroaction positive (cortex - striatum - Gpi - thalamus - cortex) et en une voie hyperdirecte et divergente de rétroaction négative (cortex - STN - Gpi - thalamus - cortex). La sélection de l'action dans chaque boucle est une propriété émergente de l'interaction des voies directe et hyperdirecte.

A.3 Tables supplémentaires

TABLE A.1 – Détails de l’environnement de travail lors du développement du modèle et de l’analyse des résultats.

Environnement de travail		
SE	OSX 10.11 (El Capitan) et Ubuntu 14.04 LTS	
Machine	Macbook Air i7-8 Go et Asus G74S, i7-16 Go	
Langages	C++14 (clang-703.0.31 et GNU C++ 4.9.4)	(Développement du modèle)
	Python 2.7.11	(Analyse et visualisation des données)
	R 3.2.3	(Analyse et visualisation des données)
Librairies	Numpy 1.11.0	
	Matplotlib 1.5.1	
	Ggplot2 2.1.0	

TABLE A.2 – Détails des numéros des neurones et sous-populations neuronales du modèle (1/3).

Numéros des neurones	Sous-total	Sous-population
1 → 100	100	Cortex moteur 1
101 → 200	100	Cortex moteur 2
201 → 300	100	Cortex moteur 3
301 → 400	100	Cortex moteur 4
401 → 500	100	Cortex cognitif 1
501 → 600	100	Cortex cognitif 2
601 → 700	100	Cortex cognitif 3
701 → 800	100	Cortex cognitif 4
801 → 900	100	Cortex associatif 1
901 → 1000	100	Cortex associatif 2
1001 → 1100	100	Cortex associatif 3
1101 → 1200	100	Cortex associatif 4
1201 → 1300	100	Cortex associatif 5
1301 → 1400	100	Cortex associatif 6
1401 → 1500	100	Cortex associatif 7
1501 → 1600	100	Cortex associatif 8
1601 → 1700	100	Cortex associatif 9
1701 → 1800	100	Cortex associatif 10
1801 → 1900	100	Cortex associatif 11
1901 → 2000	100	Cortex associatif 12
2001 → 2100	100	Cortex associatif 13
2101 → 2200	100	Cortex associatif 14
2201 → 2300	100	Cortex associatif 15
2301 → 2400	100	Cortex associatif 16

TABLE A.3 – *Détails des numéros des neurones et sous-populations neuronales du modèle (2/3).*

Numéros des neurones	Sous-total	Sous-population
2401 → 2440	40	Striatum moteur 1
2441 → 2480	40	Striatum moteur 2
2481 → 2520	40	Striatum moteur 3
2521 → 2560	40	Striatum moteur 4
2561 → 2600	40	Striatum cognitif 1
2601 → 2640	40	Striatum cognitif 2
2641 → 2680	40	Striatum cognitif 3
2641 → 2720	40	Striatum cognitif 4
2721 → 2760	40	Striatum associatif 1
2761 → 2800	40	Striatum associatif 2
2801 → 2840	40	Striatum associatif 3
2841 → 2880	40	Striatum associatif 4
2881 → 2920	40	Striatum associatif 5
2921 → 2960	40	Striatum associatif 6
2961 → 3000	40	Striatum associatif 7
3001 → 3040	40	Striatum associatif 8
3041 → 3080	40	Striatum associatif 9
3081 → 3120	40	Striatum associatif 10
3121 → 3160	40	Striatum associatif 11
3161 → 3200	40	Striatum associatif 12
3201 → 3240	40	Striatum associatif 13
3241 → 3280	40	Striatum associatif 14
3281 → 3320	40	Striatum associatif 15
3321 → 3360	40	Striatum associatif 16
3361 → 3370	10	STN moteur 1
3371 → 3380	10	STN moteur 2
3381 → 3390	10	STN moteur 3
3391 → 3400	10	STN moteur 4
3401 → 3410	10	STN cognitif 1
3411 → 3420	10	STN cognitif 2
3421 → 3430	10	STN cognitif 3
3431 → 3440	10	STN cognitif 4
3441 → 3450	10	GPi moteur 1
3451 → 3460	10	GPi moteur 2
3461 → 3470	10	GPi moteur 3
3471 → 3480	10	GPi moteur 4
3481 → 3490	10	GPi cognitif 1
3491 → 3500	10	GPi cognitif 2
3501 → 3510	10	GPi cognitif 3
3511 → 3520	10	GPi cognitif 4

TABLE A.4 – Détails des numéros des neurones et sous-populations neuronales du modèle (3/3).

Numéros des neurones	Sous-total	Sous-population
3521 → 3540	20	Thalamus moteur 1
3541 → 3560	20	Thalamus moteur 2
3561 → 3580	20	Thalamus moteur 3
3581 → 3600	20	Thalamus moteur 4
3601 → 3620	20	Thalamus cognitif 1
3621 → 3640	20	Thalamus cognitif 2
3641 → 3660	20	Thalamus cognitif 3
3661 → 3680	20	Thalamus cognitif 4
Total	3680 neurones	

TABLE A.5 – Résumé du modèle.

Résumé du modèle	
Populations	Douze en tout : Cortex (moteur, cognitif et associatif), Striatum (moteur, cognitif et associatif), GPi (moteur et cognitif), STN (moteur et cognitif), Thalamus (moteur et cognitif)
Topologie	–
Connectivité	1 : 1, 1 : n (divergence), n : 1 (convergence)
Modèle neuronal	<i>Leaky Integrate-and-Fire</i> (LIF)
Modèle de canal	–
Modèle de synapse	Saut de potentiel
Plasticité	Règles d'apprentissage par renforcement
Entrée	Courant excitateur externe dans les zones corticales (motrice, cognitive et associative)
Mesures	spikes

A.4 Publications

Charlotte Héricé, Radwa Khalil, Marie Moftah, Thomas Boraud, Martin Guthrie and André Garenne. **Decision making under uncertainty in a spiking neural network model of the basal ganglia**. Accepté dans le *Journal of Integrative Neuroscience* le 15 septembre 2016, publication prévue dans le numéro de décembre 2016.

A.5 Posters

Dans cette partie sont présentés les posters exposés lors de conférences, congrès nationaux et internationaux :

- « *Fourth International Symposium on Biology of Decision Making* » - Paris, France - Mai 2014 (Figure A.2)
- « *Fifth International Symposium on Biology of Decision Making* » - Paris, France - Mai 2015 (Figure A.3)
- « *Multi-disciplinary Conference on Reinforcement Learning and Decision Making* » - Edmonton, Canada - Juin 2015 (Figure A.4)
- Rencontres INRIA-Industries (R2I) - Bordeaux, France - Octobre 2015 (Figure A.5)
- « *Sixth International Symposium on Biology of Decision Making* » - Paris, France - Mai 2016 (Figure A.6)

M. Guthrie^{1,2}, C. Héricé^{1,2}, R. Khalil², M. Moftah³, T. Boraud^{1,2}, A. Garenne^{1,2}
 1 Université Bordeaux-Segalen, UMR 5293, Institut des Maladies Neurodégénératives, Bordeaux, France.
 2 CNRS, UMR 5293, Institut des Maladies Neurodégénératives, Bordeaux, France
 3 Zoology Department, Faculty of Science, Alexandria University, 21151, Alexandria, Egypt

Introduction

Basal ganglia (BG) are known to hosts mechanisms of action selection and its adaptation to a changing environment. Their architecture consist of several parallel functional loops connecting back to distinct areas of cortex (motor, cognitive and limbic) and processing different modalities of decision making. The picture of parallel loops is complicated by convergence and divergence that implies that the various loops are interacting. A previous BG model was developed by Martin Guthrie. This model was built of interacting bloc-diagram based on rate-models and it was able to learn optimized action selection during a probabilistic reward task. The aim of the present work is to refine and extend these results to a cell-synapse level through a bottom-up approach. This should allow us to:

1. highlight structure-function relationship and circuitry emerging properties
2. Investigate cell-scale mechanisms impact on the whole model capacities (learning and decision-making)

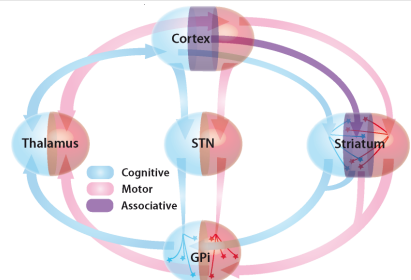


Figure 1. Overview of model architecture*.

Material and Method

Behavioral task

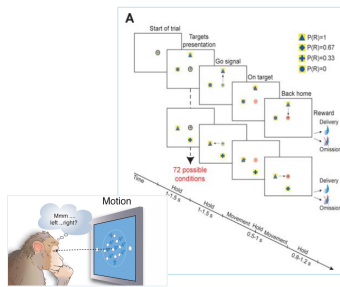


Figure 2. Task design**

4 target shapes are associated with different reward probabilities. At each trial, 2 targets, randomly chosen among the 4, are presented in random positions.

A choice is made by the model and the reward is given or not according to its own reward probability (probabilistic learning task).

The model aim is to learn to choose the optimum target on each trial, which is the target with the higher probability of reward.

Implementation

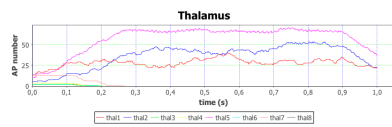
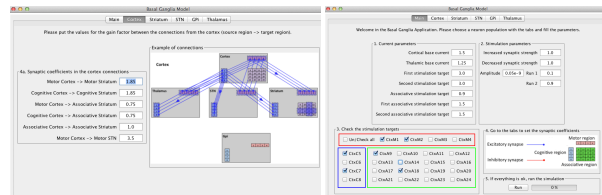


Figure 3. Modeling tool java implementation

Network architecture

The model is designed in a manner consistent with known anatomy and electrophysiology in three main areas.

Each neuron is implemented by a leaky integrate-and-fire model:

$$C_m \cdot \frac{dV_m(t)}{dt} + \frac{V_m(t)}{R_m} = I(t)$$

Each subpopulation is composed of a cell assembly which size is related to the respective cell ratios.

Learning is modeled as an increase in cortico-striatal cognitive projections (synaptic plasticity).

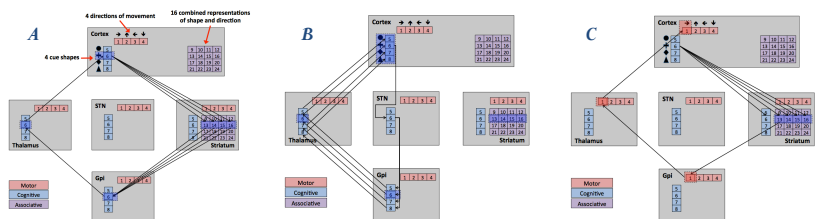


Figure 4. Model pathways. A) Direct pathway. B) Hyperdirect pathway. C) Pathway crosstalk

Results

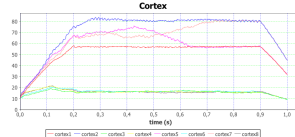
Global network properties

A manual tuning of the strength of the different projections has provided a model which is robust to noise and consistent regarding (i) the average and relative levels of activity of each cell assembly and (ii) the interactions between each subpopulation.

Feed-forward inhibition and action selection

Thanks to the architecture and the noise, the model spontaneously selects an action (defined here by the highest motor cortex activity) which breaks the symmetry. During this "exploration" phase, the naive network selects motor alternative with equal probabilities.

Figure 5. Cortical subpopulations recordings (Hz)



* Guthrie et al, 2013

Cognitive bias of motor selection

The symmetry breaking occurs with equal probabilities regarding the motor and cognitive activation in a naive model but with a superior probability for motor-cognitive coupled pairs.

Discussion

The neural network implementation of a BG stochastic model is developed, embedded in a sensory-motor closed-loop environment and applied to an action-selection task. This neuron level description model allows for a better understanding of the information flows dynamics in the BG and for an investigation of the physiopathological models including synapses, cells and neuromodulators properties. To achieve this study two steps have to be reached:

- role of the cognitive bias on action selection as a function of learnt reward values.
- implementation of a dopamine modulated STDP to reproduce a complete learning process with a bottom-up and model-free paradigm.

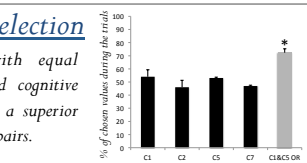


Figure 6. Average of chosen values

** Pasquereau et al, 2007

Setting A Behavioral Task For Assessing Decision Making In Urodele Amphibian

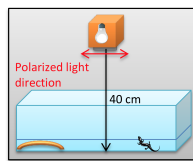
Charlotte H ERIC  1,3,4, Manon BONNET-SAVE 1,2,4, Andr  GARENNE 1,3,4, Jean-Marie CABELGUEN 2,4, Thomas BORAUD 1,2,4

INTRODUCTION

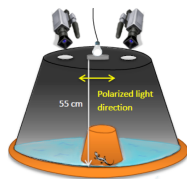
Decision making implies cortical-basal ganglia (BG) loops, but the respective roles of each structure are still debated. Urodeles, by their similitude with mammals in BG organisation, offer a much less level of complexity in understanding this process by the presence of fewer neurons. Therefore, we decided to set up a behavioral test of decision making in this animal. Moreover, their regenerative capacities make them a unique model to study post-lesional plasticity of this process networks. To our knowledge, urodele is the earliest group of limbed vertebrates in which decision-making has been successfully addressed yet.

BEHAVIORAL TEST

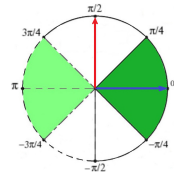
Animals have been trained in associating a polarized light direction to the position of an obscure shelter [1]. Then they were tested according to the feature of the light (non polarized light, NPL; polarized light, PL). Note that the direction of the polarization and the N-S direction of the magnetic field were perpendicularly oriented, both during training and test.



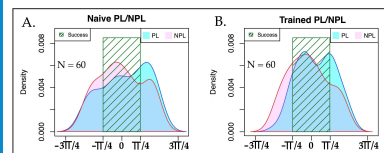
Training tank. Animals were housed under a 12h:12h polarized light:dark cycle and trained 4 times a day during 3 weeks.



Test arena. Testing sessions occurred once a day and by 3 tries for each tested condition.



Results field representation. Polarized light vector (blue arrow) and terrestrial magnetic orientation (red arrow).



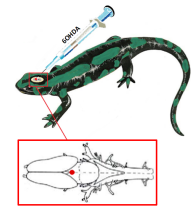
Angle occurrence probability. Following training, the number of animals in the success zone is increased, showing that training is effective.

ANIMALS

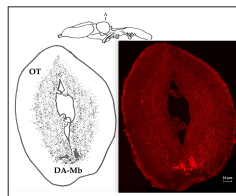
All the experiments were performed on fully metamorphosed amphibian urodeles (*Pleurodeles waltlii*) (13 months post-hatching). We used 20 animals for the behavioral test and 7 for the DA system lesion (preliminary results obtained for 2 out of 7 animals are reported).

DOPAMINERGIC SYSTEM LESION

The animals were deeply anaesthetised via immersion in a 0.1% aqueous solution of tricaine methanesulfonate (MS-222). The micro-pipette containing a 6-OHDA solution was inserted into the brain through a small hole surgically drilled at the junction of the parietal and frontal bones in the cranial midline.

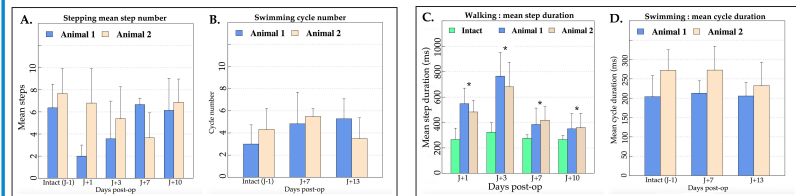


Injection site [3]. The red dot indicates the 6-OHDA injection site, directly into the third ventricle at the level of diencephalic/mesencephalic dopaminergic (DA) neurons.



Identification of DA neurons : Immunohistochemical identification. OT: Optic Tectum. DA-Mb: DA cells groups of the midbrain. Antibodies raised against tyrosine hydroxylase (TH). The study will be focused on the ventral diencephalic/mesencephalic TH+ cells, which are consistently affected by 6-OHDA administration in the third ventricle.

The number of successive steps during each walking episode (A) and the step duration (C) were measured for each animal before, 7 days and 10 or 13 days after the 6-OHDA injection. A similar quantification was performed for swimming cycles (B, D).



The effect of 6-OHDA. The 6-OHDA lesion induced clear short-term effects on the walking behavior. By contrast, the swimming behavior seemed unaffected.

FUTURE DIRECTION

This paradigm gives us access to a motor decision making mode. In a second step, we will perform selective electrical lesions of the pallidum (i.e. the cortex in lower vertebrates), the thalamus and the basal ganglia to determine the respective role of each structure. We will then study implication of dopamine in this behavioral task by 6-OHDA lesions.

INSTITUTES

- 1 Institut des Maladies Neurod g n ratives, CNRS UMR 5293, Bordeaux, France
- 2 Institut Fran ois Magendie, INSERM U862, Bordeaux, France
- 3  quipe Mn mosyne INRIA, Talence, France
- 4 Universit  de Bordeaux, Bordeaux, France

SOURCES

Video recordings are available by flashing the QR-code. Mail contact: charlotte.herice@u-bordeaux.fr



REFERENCES

- [1] Taylor & al. Spatial Orientation by Salamanders Using Plane-Polarized Light, 2007 In *Science*, 181, 285-287
- [2] Retailleau & al. Why am I lost without dopamine? Effects of 6-OHDA lesion on the encoding of reward and decision process in CA3, 2013 In *Neurobiology of Disease*, 59, 151-164
- [3] Parish & al. Midbrain dopaminergic neurogenesis and behavioural recovery in a salamander lesion-induced regeneration model, 2007 In *Development and Disease*, 134, 2881-2887
- [4] Nieuwenhuys & al. The central nervous system of vertebrates, 1998 In *Edition Springer*, Volume 2, chapter 18

FIGURE A.3 – Poster pr sent  lors du « Fifth International Symposium on Biology of Decision Making » - Paris, France - Mai 2015

Decision making mechanisms in a connectionist model of the basal ganglia

Charlotte HÉRICE^{1,2,3}, Radwa KHALIL¹, Marie MOFTAH⁴,
Thomas BORAUD^{1,2,3}, Martin GUTHRIE^{1,3} and André GARENNE^{1,2,3}

INTRODUCTION

Basal ganglia (BG) are known to hosts mechanisms of action selection and its adaptation to a changing environment. Their architecture consist of several parallel functional loops connecting back to distinct areas of cortex (motor, cognitive and limbic) and processing different modalities of decision making. The picture of parallel loops is complicated by partial convergence and divergence connections that implies that the various loops are interacting.

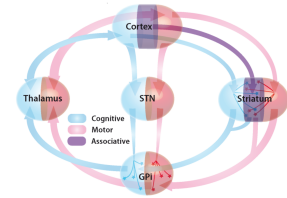


Figure 1. Overview of model architecture

A previous BG model (developed by M. Guthrie[1]) was built of interacting bloc-diagram based on rate-models. It was able to learn optimized action selection during a probabilistic reward task. The aim of the present work is to refine and extend these results to a cell-synapse level through a bottom-up approach:
→ Highlighting of the structure-function relationship and circuitry emerging properties
→ Investigation of cell-scale mechanisms impact on the whole model capacities (learning and decision-making)

NETWORK ARCHITECTURE

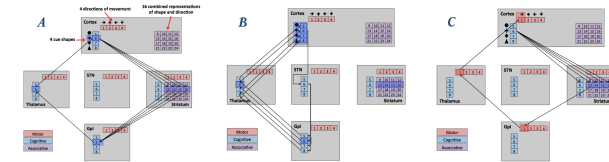


Figure 4. Model pathways. A) Direct pathway. B) Hyperdirect pathway. C) Pathway crosstalk

Leaky integrate-and-fire neuron model: $C_m \cdot \frac{dV_m(t)}{dt} + \frac{V_m(t)}{R_m} = I(t)$
Learning takes place when a reward is obtained. It is implemented as an increase in the synaptic weights corresponding to the projection from the cognitive cortex to the striatum.

RESULTS

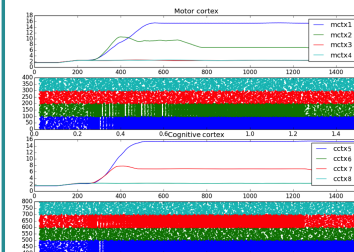


Figure 5. Example of the cortex activity in a naive network

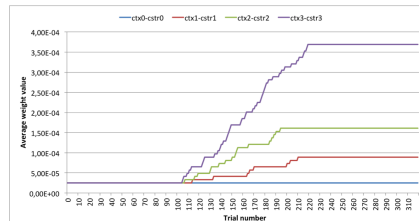


Figure 6. Example of the cortex activity after learning

Naive network

Ability of spontaneous action selection (defined here by a sufficient difference in the activity between two regions of the motor cortex).
→ Exploration phase ⇔ Target selection of the primate.

Trained network

Ability to preferentially select the optimum target (highest reward probability).
→ Exploitation phase ⇔ Successful learning from the primate.

Before/after learning

Better percentage of success (choosing the optimum target, p-value < 5%) after learning than before.

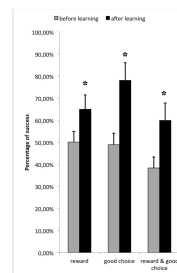
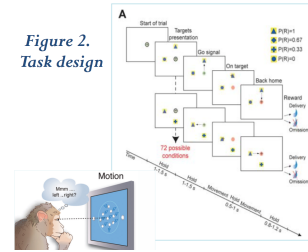


Figure 7. Successful trials before/after learning

BEHAVIORAL TASK

Figure 2. Task design



4 target shapes ⇔ 4 different reward probabilities. At each trial[2]:
- random presentation of 2 targets, at random positions
- choice made by the model
- reward given or not according to the reward probability (probabilistic learning task)
→ The model has to learn to choose the optimum target on each trial.

IMPLEMENTATION

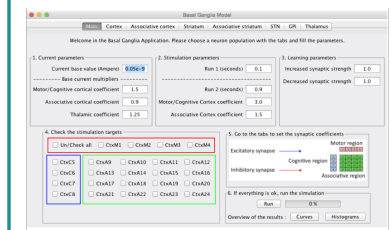
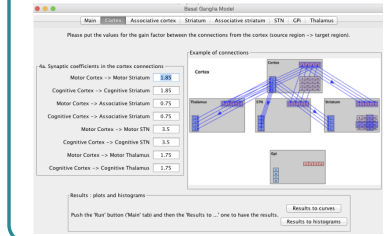


Figure 3. Modeling tool with Java implementation



DISCUSSION

The neural network implementation of a BG stochastic model is developed, embedded in a sensory-motor closed-loop environment and applied to an action-selection task. Once finalized, this neuron level description model allows for a better understanding of the information flows dynamics in the BG and for an investigation of the physiopathological models including synapses, cells and neuromodulators properties.

To achieve this study, an accurate implementation of a dopamine modulated LTP mechanism will be done to reproduce a complete learning. To validate the model, lesional tests and changes in the conditions of the task are planned.

REFERENCES

- [1] M. Guthrie et al. Interaction between cognitive and motor cortico-basal ganglia loops during decision making: a computational study, 2013. *J. Neurophysiology*
- [2] B. Pasquereau et al. Shaping of Motor Responses by Incentive Values through the Basal Ganglia, 2007. *J. Neuroscience*

INSTITUTES

- ¹ Institut des Maladies Neurodégénératives, CNRS UMR 5293, Bordeaux, France
- ² Equipe Mnémosyne INRIA, Talence, France
- ³ Université de Bordeaux, Bordeaux, France
- ⁴ Zoology Department, Faculty of Science, Alexandria University, Alexandria, Egypt

CONTACT

andre.garenne@u-bordeaux.fr
charlotte.herice@u-bordeaux.fr

FIGURE A.4 – Poster présentée lors de « Multi-disciplinary Conference on Reinforcement Learning and Decision Making » - Edmonton, Canada - Juin 2015

L'équipe **Mnemosyne** a pour but de **modéliser le cerveau** comme un système de **mémoires actives** en synergie et en interaction avec les mondes interne et externe et de le simuler en situation dans des conditions réalistes. Actuellement, nous nous concentrons sur certaines zones cérébrales : **amygdale, cortex, ganglions de la base et hippocampe**. Notre approche de **modélisation bio-inspirée** s'intéresse à des **phénomènes cognitifs** aussi bien **Globaux que Locaux** et vise des impacts dans les **domaines du Numérique et de la Santé**.

Le côté Global du Numérique

- **Apprentissages basés sur une récompense et algorithmes de prise de décision** : faire le parallèle entre ces éléments permet d'évaluer la performance des algorithmes d'apprentissage et de prise de décision existants et de les améliorer en s'inspirant de la biologie.



- **Robotique autonome** : l'étude de structures cérébrales précises permet d'obtenir des modèles robotiques (1) pour des tâches de cartographie (en lien avec l'hippocampe) ou de navigation vers un but (en lien avec le cortex frontal). De plus, cela apporte un point de vue différent sur les maladies neurodégénératives telles que la maladie d'Alzheimer.

Le côté Global de la Santé

- **Psychopathologie** (comprend l'étude de la psychologie et de la psychiatrie) et **synergie des structures cérébrales** : l'ensemble de ces structures collaborent ensemble et il n'est pas toujours évident de les observer. Le but principal est d'avoir un modèle global lié à notre approche systémique pour comprendre le fonctionnement de ces structures (3).

- Exemple de la **maladie d'Alzheimer** : maladie neurodégénérative globale, induisant des problèmes de mémoire (hippocampe) provoqués par des problèmes de connexion et de communication entre les différentes mémoires. C'est donc l'ensemble de l'architecture mnésique qui est touché.

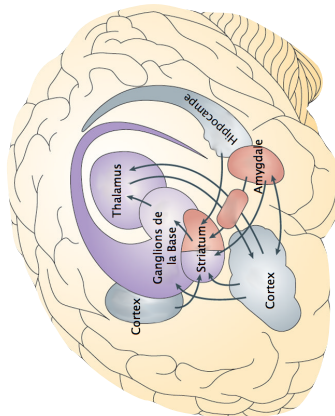


- **Additions** : l'étude de l'ensemble des structures vise à expliquer comment, même si « ça ne vaut pas le coup », l'action est quand même réalisée. En termes neurologiques, la valeur de l'action (au niveau des ganglions de la base) reste élevée même si celle du stimulus (au niveau de l'amygdale) est basse.

- **Le but de l'équipe Mnemosyne** est, à ce niveau, la réalisation d'un modèle cognitif général, offrant la possibilité d'étudier artificiellement ce modèle et ses variantes (tests léionnels par exemple).

Références :

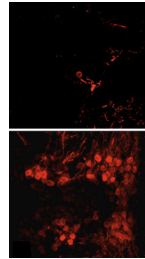
- (1) Nicolas Denoyelle, Florian Pouget, Thierry Viéville, Frédéric Alexandre. VirtualEnaction: A Platform for Systemic Neuroscience Simulation.. International Congress on Neurotechnology, Electronics and Informatics, Oct 2014, Rome, Italy, 2014
- (2) Charlotte Hélicé, Radwa Khalil, Maria Moffah, Thomas Boraud, Martin Guthrie, André Garenne. Decision making mechanisms in a connectionist model of the basal ganglia. The Multi-disciplinary Conference on Reinforcement Learning and Decision Making (RLDM 2015), Jun 2015, Edmonton, Canada.
- (3) Meroopi Topalidou, Camille Piron, Daisuke Kaise, Thomas Boraud, Nicolas P. Rougier. The formation of habits : The implicit supervision of the basal ganglia. Fifth International Symposium on Biology of Decision Making (SBDM 2015), May 2015, Paris, France.
- (4) Parish et al. Midbrain dopaminergic neurogenesis and behavioural recovery in a salamander lesion-induced regeneration model. *Development*, 2007
- (5) Maxime Carrère, Frédéric Alexandre. A pavlovian model of the amygdala and its influence within the medial temporal lobe. *Frontiers in Systems Neuroscience*, Frontiers, 2015, pp.14.



Charlotte Hélicé
13 Octobre 2015

Le côté Local de la Santé

- L'étude des **ganglions de la base** améliore la compréhension de la **maladie de Parkinson** : maladie neurodégénérative provoquée par la mort des neurones dopaminergiques, situés dans les ganglions de la base.



Exemple d'écriture d'un patient parkinsonien

Exemple de neurones dopaminergiques chez un sujet sain (à gauche) et un sujet parkinsonien (à droite) (4)

- **Stimulation Cérébrale Profonde** : méthode utilisée pour traiter localement les dégénérescences neuronales liées à la maladie de Parkinson.

- **Aspects pharmacologiques et neuromodulation** : la simulation des effets d'une molécule (5) permet de prédire les effets d'un médicament sur le cerveau, et a posteriori sur le comportement (maladies neurologiques et psychiatriques).



Charlotte HÉRICÉ et Ikram CHRAÏBI KAADOU

FIGURE A.5 – Poster présenté lors des Rencontres INRIA-Industries (R2I) - Bordeaux, France - Octobre 2015

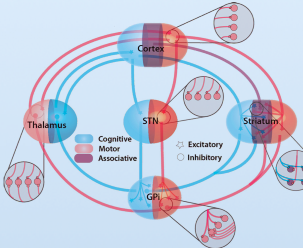
Decision-making in a neural network model of the basal ganglia

Charlotte HÉRICÉ^{1,2,3}, Radwa KHALIL¹, Marie MOFTAH⁴,
Thomas BORAUD^{1,2}, Martin GUTHRIE^{1,2} and André GARENNE^{1,2,3}

Contact: charlotte.herice@u-bordeaux.fr
andre.garanne@u-bordeaux.fr

INTRODUCTION

Basal ganglia (BG) are known to host mechanisms of **action selection** and its adaptation to a changing environment. Their architecture consist of several parallel functional **loops** connecting back to distinct areas of cortex (motor, cognitive and limbic) and processing different modalities of **decision making**. The picture of parallel loops is complicated by partial convergence and divergence connections that implies that the various loops are interacting.



A previous BG model¹ was built of interacting bloc-diagram based on **rate-models**. It was able to learn optimized action selection during a probabilistic reward task. The aim of the present work is to refine and extend these results to a cell-synapse level through a **bottom-up approach**.

- Highlighting of the **structure-function relationship** and circuitry emerging properties.
- Investigation of **cell-scale mechanisms** impact on the whole model capacities (learning and decision-making).

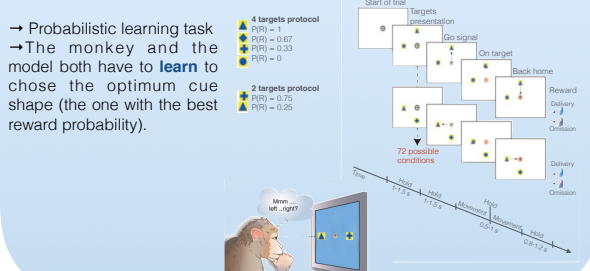
BEHAVIORAL TASK

We submit the model to a protocol^{2,3} for BG involvement in decision-making with monkeys in **conditions of uncertainty**. There are **4 different cue shapes**, each with its **own reward probability** and 4 possible positions.

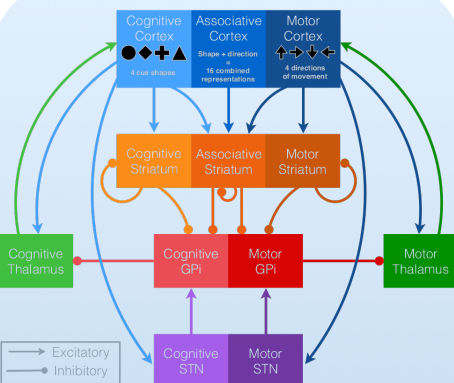
At each trial :

1. Random **presentation** of 2 cue shapes (at random positions)
2. **Choice** made by the monkey and the model
3. **Reward** given or not according to the reward probability of the shape

→ Probabilistic learning task
→ The monkey and the model both have to **learn** to chose the optimum cue shape (the one with the best reward probability).



NETWORK ARCHITECTURE

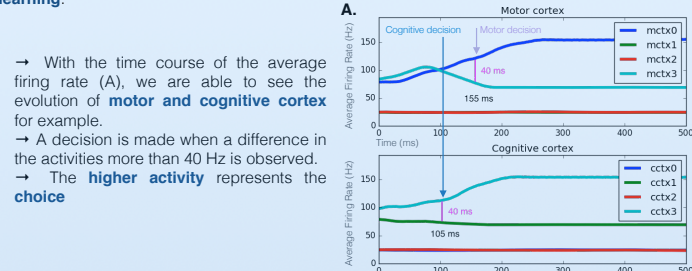


- **Spiking neurons**: Leaky Integrate-and-Fire (LIF) neurons and voltage jump synapses.
- **Learning**: adaptation of the cognitive corticostriatal projections strength modulated by a phasic dopamine release (≈ reward prediction error).

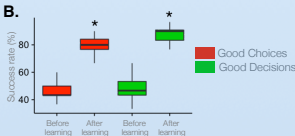
RESULTS

- Exploration:
An expected emergent property of the network is a **divergence** in the cortical activations of cognitive and motor sub-populations. In the absence of learning the network is still able to make a decision. This is equivalent to **decision-making** during the exploration phase of **reinforcement learning**.


→ With the time course of the average firing rate (A), we are able to see the evolution of **motor and cognitive cortex** for example.
→ A decision is made when a difference in the activities more than 40 Hz is observed.
→ The **higher activity** represents the **choice**



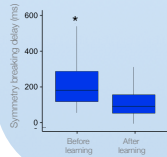
- Exploitation:
A Good Choice (GC) is made when the **optimal shape** is selected and a Good Decision (GD) when the **associated direction** is selected too. Both are improved during a standard learning session (B).



C. During training, the model learns to create a **dynamic link** between the cognitive and motor sensory component of a cue. This can be assessed by the learning curves profile of the model (C).



→ The **average reward** and **GC** rate gradually increase along the session (C).
→ The **optimum** cue shape direction is preferentially selected (B).
→ The movement onset delay is decreased by the **learning** (D).



CONCLUSION

We have presented here, for the first time, a biophysically based, **spiking neuron model of the BG** that is able to perform 2 levels action selection. This model is closely based on the known anatomy and physiology of the basal ganglia and demonstrates a reasonable mechanism of network level **action selection**. This cellular and synaptic level of description bridges the gap between top-down mesoscopic level of description¹ and a bottom-up approach relying on emerging properties of neuronal networks dynamics. Our model is also able to predicts some important behavioral characteristics like localized lesion consequences on **learning impairment** and intrinsic dynamics, **reversal learning** and **extinction protocol**.

REFERENCES

- ¹ M. Guthrie et al. Interaction between cognitive and motor cortico-basal ganglia loops during decision making: a computational study, 2013 *J. Neurophysiology*
- ² B. Pasquier et al. Shaping of Motor Responses by Incentive Values through the Basal Ganglia, 2007 *J. Neurosciences*
- ³ C. Piron et al. The GPI in goal-oriented and routine behaviors: resolving a long-standing paradox, 2015 *Movement Disorders*

INSTITUTES

- ¹ Institut des Maladies Neurodégénératives, CNRS UMR 5293, Bordeaux, France
- ² University of Bordeaux, Bordeaux, France
- ³ Team Mnemosyne INRIA Bordeaux Sud-Ouest, Bordeaux, France
- ⁴ Zoology Department, Faculty of Science of Alexandria University, Alexandria, Egypt



FIGURE A.6 – Poster présenté lors du « Sixth International Symposium on Biology of Decision Making » - Paris, France - Mai 2016