From neural coding to decision-making

Kevin Berlemont

PSL UNIVERSITÉ PARIS

THÈSE DE DOCTORAT

DE L'UNIVERSITÉ PSL

Préparée à l'Ecole Normale Supérieure

Du codage neuronal à la prise de décision

From neural coding to decision-making

Soutenue par

Kevin BERLEMONT

Le 29 Juin 2020

École doctorale n°564

Physique en Île-de-France

Spécialité

Physique

Composition du jury :

Zhaoping LI Sensory and Sensorimotor Systems, Max-Planck-Institut	*Présidente du jury*
Bruno CESSAC INRIA	*Rapporteur*
Peter LATHAM Gatsby Computational Neuroscience Unit - UCL London	*Rapporteur*
Wulfram GERSTNER Laboratory of Computational Neuroscience, EPFL	*Examinateur*
Claire SERGENT Université Paris-Descartes	*Examinatrice*
Gianluigi MONGILLO Institut de la Vision - CNRS	*Invité*
Jean-Pierre NADAL CNRS - EHESS	*Directeur de thèse*

1794 ENS | PSL

Remerciements

En tout premier lieu je tiens à remercier mon directeur de thèse, Jean-Pierre Nadal, pour son accompagnement durant cette thèse multi-disciplinaire. Tout d'abord, je tiens à le remercier pour ses nombreux conseils d'orientations durant mes études. Merci également de toujours avoir pris le temps de combler mes nombreuses lacunes en sciences cognitives et biophysiques, de m'avoir ouvert à de nombreux sujets autre que les neurosciences durant ce doctorat et d'avoir pris le temps de partager la rigueur scientifique nécessaire à un travail de recherche. Pour finir, et sûrement le plus important, je tiens à le remercier de m'avoir laissé une si grande liberté dans les thèmes que j'ai souhaité aborder et d'y avoir manifesté un grand intérêt durant ces trois années. Je ne peux qu'être reconnaissant pour cet encadrement qui fut tout ce que j'aurai pu espérer.

Je remercie vivement Bruno Cessac et Peter Latham pour leur relecture attentive du manuscrit en tant que rapporteurs. Je tiens également à remercier Wulfram Gerstner, Zhaoping Li et Claire Sergent d'avoir accepté de faire partie du jury de ma thèse.
Je remercie également Martine Ben Amar et Srdjan Ostojic pour avoir accepté de faire partie de mon comité de suivi de thèse et d'avoir suivi ma thèse durant ces trois années.
Je tiens à remercier Gianluigi Mongillo qui m'a accompagné durant toute cette période du doctorat. Continuer notre travail de recherche commencé pendant mon M2 m'a permis de faire une pause dans le sujet principal de la thèse lorsque la motivation venait à baisser. Merci pour les nombreuses discussions scientifiques et pour les encouragements durant toutes ces années.

J'adresse également mes remerciements à Jérôme Sackur et Jean-Rémy Martin avec qui j'ai pu travailler sur la partie expérimentale, bien qu'étant totalement novice en sciences cognitives expérimentales. Sans eux, les résultats théoriques n'auraient pas pu avoir la même portée. Un remerciement particulier à Jérôme qui m'a expliqué les nombreuses techniques statistiques utilisées pour l'analyse de données expérimentales.
Je tiens ici à citer Laurent Bonnasse-Gahot, Vincent Hakim et Xiao-Jing Wang pour les nombreuses discussions que nous avons eues sur mes recherches, et les conseils qu'ils m'ont apportés.

Je remercie les différents membres du service administratif du laboratoire pour leur présence durant ces années. En particulier, je tiens à remercier Annie pour avoir toujours été disponible et qui va maintenant pouvoir récupérer sa chaise de bureau. Un merci également à Frédéric, Nabil et Zahir pour leur aide sur les nombreux sujets informatiques.

A l'ENS je tiens à remercier les différentes personnes que j'ai pu croiser, dans un ordre non exhaustif : Marco, Victor, Lorenzo, Clément, Aldo, Simone, Elie, Fabio, François, Trang-Ahn, Ling, Martin, Arnaud, Hugo, et Francesca. Un merci particulier à Marco, Clément, Elie et Aldo pour leur aide dans l'écriture du manuscrit et la préparation de la soutenance.

Je tiens à remercier l'ensemble des personnes présentes (professeurs et étudiants) lors de l'école d'été Woodshole 2018. Cette école fut l'occasion de découvrir de nombreux domaines des neurosciences dont je n'avais pas connaissance, et de faire de nouvelles rencontres. Une mention particulière à Manuel et Ulises que j'ai pu connaître lors de cette école.

Sur un plan plus personnel, merci à Anne-Lise pour tout ce que nous partageons depuis déjà 13 ans.
A Alice, merci d'avoir partagé cette aventure avec moi.

J'adresse également mes remerciements aux membres de l'EKSTA (et associés) pour

les innombrables soirées qu'ils ont organisées: Alexandre, Auriane, Baptiste, Bertrand, Célestine, Charlie, Laetitia, Pierre et Rémy. Plus particulièrement, à Baptiste, je tiens à rappeler ces ~~vacances~~ stages que nous avons faits à Palma et en particulier La Cuadra del Maño.
De manière plus générale, un grand merci à tous les membres du Play[Mobile].

Au groupe du LOL : Damien, Florian, Hugo, Joyce, Lucas, Mariam, Nais, Philippe, Robin, Thomas, Yann, déjà 10 ans de chochoye et autres frivolités !

Un merci aux différentes personnes qui ont fait partie d'une (ou de plusieurs) de ces années : Camille, Karen, Loic, Guillaume, Florian, encore Camille, Jérémy, en espérant que les personnes oubliées me le pardonneront.

Je tiens également à remercier mes parents, mon frère et ma sœur pour le support qu'ils m'ont apporté durant toutes ces années. Un grand merci également à mes parents pour l'aide qu'ils m'ont apportée dans la préparation de la version *en ligne* de la soutenance. Pour finir, Florie, merci pour la relecture complète de ce manuscrit.

Introduction

Neuroscience and neuropsychology study decision-making using experimental and theoretical methods that are usually based on simple paradigms. Decision-making can consist in either a choice between many alternatives, a decoding process, or a choice to perform a motor action. Perceptual decision-making, the most experimentally studied (with humans and animals) type of decision-making, is the subject of my PhD. Perceptual decision-making studies the mechanisms through which one is able to process more or less ambiguous stimuli to reach a decision.

Last decades have shown a surge in neurophysiological study of decision-making, particularly in the behavior of monkeys in various decision tasks. These different experiments have shed some light into the general mechanism of decision making. The stimuli, coming from the sensory receptors, are coded in a specific neuronal area. Let us consider the task of associating a stimulus to a discrete choice, for example showing a picture to a subject and asking him to categorize it as either the picture of a cat or a dog. Experiments have shown that information collected from the stimulus is encoded in a neuronal layer.

Different models have been proposed to reproduce the decision-making process. They all have in common the fact that they try to account for the three key features of decision-making: accuracy, decision time, and confidence. During my PhD, I have studied a dynamical model of decision-making that is thought to model cortical activity. The manuscript is divided in three parts. In the first part I introduce a dynamical model of decision-making. The model is used to explain different effects that I have observed in a decision-making experiment I have performed during my PhD. The second part focus on a learning problem: how can a network learn to decode the information that has been encoded by the sensory receptors? Finally, the last part consists in my M2 project carried out under supervision of Gianluigi Mongillo and completed during my PhD. I explain how one can use statistical physics method to study neuronal networks.

- *Chapter 2:* I introduce the concept of decision-making and the neuronal underpinning. I describe different models that have been used to study decision-making.

- *Chapter 3:* I present the model that I have used during my PhD. I show that it can explain many effects that are observed in human decision-making.

- *Chapter 4:* I describe the experiment I have developped, in collaboration with Jean-Rémy Martin and Jérôme Sackur; The experiment was designed to study confidence in decision-making.

- *Chapter 5:* I show that an attractor neural network can be fitted to behavioral data, and that it reproduces most of the behavioral results of humans in a decision-making task.

- *Chapter 6:* I introduce the notion of categorical perception: which impact do categories have on decision-making?

- *Chapter 7:* I focus on the question of learning in a decision-making model. More specifically, I study the effect of categorical perception on the learning process in neuronal neworks.

- *Chapter 8:* I study a specific type of neuronal networks for which classical analytical methods, such as mean-field, cannot be used. I show that more advanced statistical physics techniques, such as the cavity method, can be used in order to analyze the network.

- *Chapter 9:* I conclude the manuscript with this chapter. I summarize everything that I have presented and I give some perspectives on what the next step could be for these different thematics.

Contents

1

Résumé substantiel en français

1.1. Introduction

1.1.1 La prise de décision en sciences cognitives et neurosciences.

Le mécanisme d'intégration temporelle constitue une idée centrale dans l'étude de la prise de décision. Cette idée d'accumulation temporelle d'évidence possède une longue histoire en psychologie cognitive. Typiquement, lorsqu'il s'agit de prendre une décision plus difficile, nous prenons plus de temps (Hick, 1952; Vickers, 1970). De plus, il y a un équilibre entre vitesse et précision; les performances sont meilleures lorsque les temps de décision sont plus longs (Wickelgren, 1977). Cependant, il n'est pas évident de déterminer l'échelle temporelle à laquelle ce mécanisme d'accumulation d'évidence a lieu.

Une des tâches perceptuelles de prise de décision la plus étudiée consiste en la *random dot motion task* (RDM). Les participants, le plus souvent des humains ou des singes, regardent un écran sur lequel certains points se déplacent aléatoirement et d'autres de manière cohérente dans une des deux directions possibles. Les participants doivent indiquer la direction nette du mouvement des points en effectuant une saccade visuelle dans cette direction. Lorsque la fraction de points se déplaçant de manière cohérente est réduite, la décision devient plus difficile. La difficulté de cette tâche peut être décrite par une quantité c appelée niveau de cohérence. Cette variable appartient à l'intervalle $[0,1]$ et correspond à une représentation à une dimension de la difficulté de la tâche. Par exemple, dans le cas de la tâche RDM, le niveau de cohérence c correspond directement à la fraction des points se déplaçant de manière cohérente. Dans cette tâche, lorsque la difficulté varie, les temps de réactions peuvent aller de 300 ms jusqu'à la seconde (Britten et al., 1992, 1993; Roitman and Shadlen, 2002). La tâche *two-alternative-forced choices* (2AFC) a également été étudiée avec des rongeurs, dans le cas d'une discrimination olfactive (Lak et al., 2014). Dans ces tâches, les rats doivent décider de la composante majoritaire dans un mélange binaire d'odeurs. Ces décisions sont rapides, de l'ordre de 300 ms. Les performances varient entre 50% et 100% de bonnes réponses suivant la difficulté de la tâche. Cependant, la différence en temps de réaction entre les stimuli les plus simples et les plus difficiles est plus faible pour les rats que pour les humains, de l'ordre de 35 ms pour Uchida and Mainen (2003) et 80 ms pour Abraham et al. (2004). Ces différents résultats témoignent de la diversité de l'intégration temporelle à travers les tâches et espèces.

Des études électrophysiologiques sur des singes ont eu lieu dans le but de comprendre les mécanismes neuronaux sous-jacents à la prise de décision perceptuelle (Roitman and Shadlen, 2002; Gold and Shadlen, 2007). Les recherches ont été concentrées sur les neurones corticaux qui encodent les signaux sensoriels dans différentes tâches de prise de décision, comme la tâche RDM ou des tâches sensorimotrices (Figure 1.1).

Les neurones sensoriels de l'aire visuelle MT encodent la direction du mouvement du

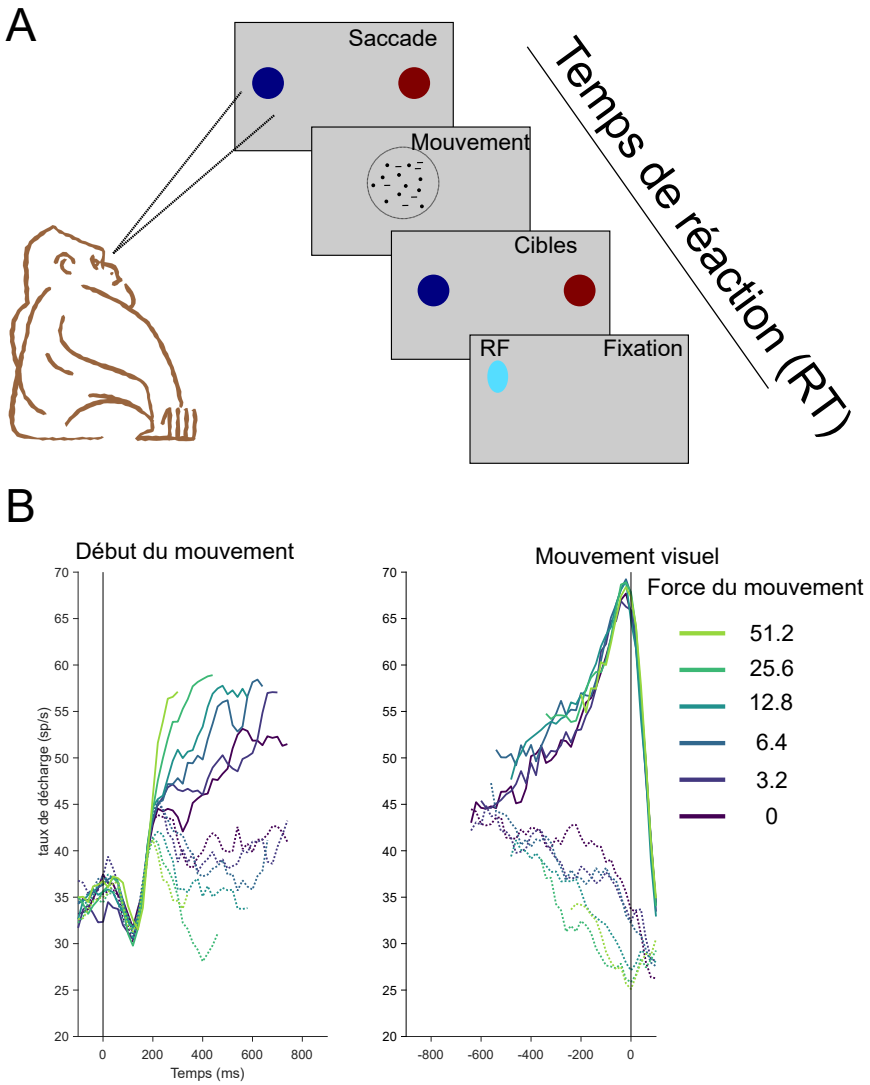


Figure 1.1: **Tâche RDM dans le cas où le singe choisit quand il répond.** (A) Le singe observe un ensemble de points qui se déplacent sur un écran et doit décider du mouvement global des points. La décision est indiquée par un mouvement saccadique visuel vers l'une des deux cibles périphériques. Le champ en bleu ciel correspond au champ réceptif d'un des neurones LIP enregistrés. (B) Réponse moyenne des neurones du LIP durant la décision. Les données viennent de Roitman and Shadlen (2002) et sont disponibles publiquement. Le taux de décharge moyen des 57 neurones du LIP est représenté pour 6 difficultés de décisions. Les taux de décharge sont regroupés par difficulté du mouvement des points et par direction du choix (les courbes en pointillés correspondent au choix de la direction opposée au champ réceptif du neurone). Le panneau de gauche représente l'activité moyenne durant la formation de la décision et le panneau de droite l'activité centrée sur le mouvement visuel.

stimulus (Newsome et al., 1989; Britten et al., 1992, 1993), mais la prise de décision n'a pas lieu dans cette aire neuronale. Shadlen and Newsome (1996) ont montré que l'activité des neurones dans le cortex latéral intrapariétal (LIP) est corrélée au choix perceptuel du singe. De plus, dans une version de la tâche où les temps de réactions sont contrôlés par les singes, un bon nombre de résultats suggère que les neurones du LIP régulent les

décisions entre saccades visuelles rivales (Roitman and Shadlen, 2002; Huk and Shadlen, 2005; Gold and Shadlen, 2007; Huk and Meister, 2012). L'activité moyenne des neurones du LIP sélectifs pour la cible visuelle de la saccade augmente du début de la présentation du stimulus jusqu'à la décision (Figure 1.1). Ce taux de croissance dépend de la qualité de l'information sensorielle; les stimuli plus forts sont associés à des pentes plus élevées d'intégration de l'information. Finalement, la décision est prise lorsque le taux d'activité des neurones du LIP (correspondant à ce choix) atteint un seuil qui est indépendant de la qualité du signal et du temps de réponse. De plus, les microstimulations des neurones des zones MT et LIP ont un effet sur les performances et les temps de réaction. Ceci concorde avec l'idée que les neurones du LIP intègrent l'information sensorielle (Ditterich et al., 2003; Hanks et al., 2006). Ces différentes études appuient l'idée que les neurones du LIP agissent comme des intégrateurs neuronaux. Malgré la variabilité stochastique de ces neurones entre les essais, ces neurones montrent en moyenne une croissance de leur activité durant la prise de décision.

Nouvelles approches pour l'étude de la prise de décision perceptuelle

Les études présentées montrent que le LIP et plus généralement le cortex préfrontal font partie d'un circuit qui est impliqué dans l'élaboration de tâches simples de prise de décision. Cependant, il y a toujours débat pour savoir où a lieu le processus d'accumulation. Les neurones du LIP montrent différents motifs de réponse aux entrées sensorielles; ces motifs peuvent être modulés par les variables sensorielles et motrices (Bennur and Gold, 2011; Park et al., 2014). Il est donc possible que les réponses neuronales du LIP ne ressemblent à un processus d'accumulation d'évidence que lorsqu'elles sont moyennées ensemble. Récemment, Katz et al. (2016) ont montré que des inactivations pharmacologiques unilatérales du LIP possèdent des effets négligeables sur le comportement. Cependant, l'inactivation de la zone du LIP affecte fortement l'analyse sensorielle durant la prise de décision, plus que les aspects moteurs (Zhou and Freedman, 2019).

Pour répondre à ces différentes questions, les chercheurs ont étudié la prise de décision chez les rongeurs. En effet, les rongeurs exhibent différents éléments de la prise de décision perceptuelle telle que l'accumulation d'évidence (Brunton et al., 2013). Il est possible de considérer de nombreuses modalités sensorielles comme des stimuli visuels (Brunton et al., 2013), tactiles (Guo et al., 2014), olfactifs (Uchida et al., 2006), en utilisant des instruments à la pointe de la technologie (Guo et al., 2014; Znamenskiy and Zador, 2013). Ces nouvelles méthodes ont permis l'étude de plusieurs aires neuronales en simultané lors de la prise de décision et de caractériser de quelle manière les cellules réagissent aux entrées sensorielles et les transforment en variable de décisions (Brody and Hanks, 2016). Les expériences futures vont permettre de préciser de manière plus spécifique les interactions entre les différentes régions du cerveau (Siegel et al., 2015).

1.1.2 Modèles de prise de décisions

De nombreux modèles ont été proposés pour expliquer la prise de décision chez les humains et les animaux. Il y a deux catégories de modèles: les modèles dynamiques et les modèles non-dynamiques. Je me concentre ici sur les modèles dynamiques car ils sont plus adéquats pour modéliser la dynamique neuronale lors de la prise de décision. Les modèles non-dynamiques seront abordés dans le chapitre 5.

Drift diffusion model

Des tests mathématiques abstraits ont été développés afin de choisir entre deux hypothèses probabilistes, comme le test de probabilité séquentielle (Wald and Wolfowitz, 1948). Ce test est optimal dans le sens où il permet d'obtenir un taux d'erreur voulu avec un temps moyen

de décision minimal. Selon ce test, les décisions sont instanciées lorsque la cumulative d'une variable aléatoire d'évidence atteint un seuil spécifique (Ratcliff, 1978; Bogacz et al., 2006; Gold and Shadlen, 2007). Le modèle le plus utilisé parmi les modèles dits de course est le drift-diffusion model (DDM). Ce modèle consiste en un intégrateur unique qui accumule la différence entre les évidences pour les deux alternatives. Le choix est fait lorsque le niveau d'activité de l'intégrateur dépasse un certain seuil, positif ou négatif suivant l'alternative (Figure 1.2).

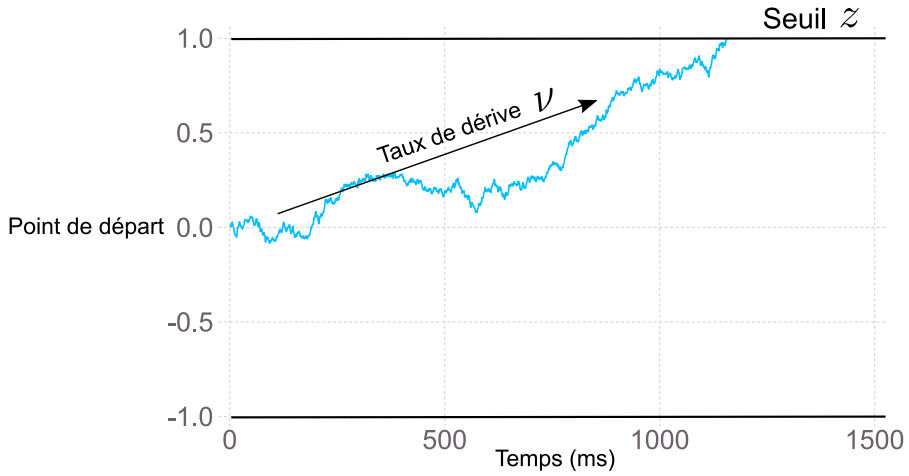


Figure 1.2: **Drift diffusion model.** Exemple de la dynamique du DDM. Les deux lignes noires représentent le seuil z et $-z$ correspondant aux deux alternatives. La course se termine lorsque la dynamique atteint l'une des deux bornes. Dans cet exemple, la décision prise correspond à un essai correct car le taux de dérive ν a été choisi positivement.

Le succès des DDMs réside dans le fait qu'ils permettent de décomposer un choix observé en un processus cognitif. Comme il s'agit d'un modèle dynamique, il prend en compte performance et temps de réponses et permet donc d'étudier le phénomène de balance entre vitesse et précision. Les DDMs intègrent l'information durant un temps plus court si l'évidence envers l'alternative gagnante est forte par rapport à celle de l'alternative perdante. Respectivement, il faudra plus de temps pour prendre la décision si la différence entre les évidences est faible. Le modèle est basé sur les paramètres suivants. Le point de départ a de l'accumulation d'évidence représente un potentiel biais en faveur d'une des deux alternatives. Cette accumulation est effectuée avec une certaine dérive ν et dépend de la qualité de l'information présente dans le stimulus. Typiquement, dans l'expérience RDM, ce paramètre va varier suivant l'intensité du mouvement des points. La frontière représente le niveau de prudence : plus la frontière est haute moins le système est sensible au bruit et moins le modèle effectuera le mauvais choix. Finalement, le temps de non-décision qui représente un délai additif constitue le dernier paramètre, par exemple dû au système moteur, chez les participants (Luce et al., 1986).

Les DDMs rendent compte des résultats comportementaux pour de nombreux paradigmes de prise de décision (Ratcliff and Rouder, 2000; Ratcliff et al., 2003; Ratcliff and Smith, 2004). Il est important de noter que les DDMs ne reproduisent pas uniquement les taux d'erreurs et temps de réponse (RT) mais également la forme de la distribution des temps de réponse (Ratcliff, 1978). Cependant,il est nécessaire d'ajouter de la variabilité dans les paramètres entre les différents essais afin de modéliser correctement la distribution des RTs dans le cas des essais qui ont conduit à une erreur. Dans le cas contraire, la linéarité du

modèle donne des distributions strictement identiques pour les essais corrects et incorrects; ce qui ne correspond à ce qui est observé expérimentalement (Ratcliff and Tuerlinckx, 2002).

Réseau cortical récurrent

Deux critiques majeures peuvent être faites à propos des DDMs lorsque le but est de modéliser l'activité neuronale durant la prise de décision. En premier lieu, l'activité neuronale est non-linéaire mais le DDM est strictement linéaire. Deuxièment, les DDMs n'ont pas de fondations biophysiques et ne permettent pas d'expliquer de quelle manière ce mécanisme d'intégration est effectué dans le cerveau. Différents modèles ont été proposés pour expliquer les processus corticaux lors de la prise de décision: Shadlen and Newsome (2001); Usher and McClelland (2001); Wang (2002). J'exposerai plus particulièrement l'approche de Wang (2002) car les deux autres sont plus proches d'une extension d'un DDM.

Réseau attracteur

Les neurones dans le LIP et le cortex préfrontal exhibent une activité qui est spécifique à l'orientation (Gnadt and Andersen, 1988; Funahashi et al., 1989). Ceci suggère qu'un mécanisme commun pourrait relier mémoire de travail et prise de décision. Un mécanisme qui génère une activité persistante similaire à celle de la mémoire de travail est celui d'une excitation récurrente forte dans circuit cortical local. Ce mécanisme donne lieu à des états attracteurs spécifiques du stimulus (Freeman, 1995; Goldman-Rakic, 1995; Brunel and Wang, 2001). Wang (2002) a construit un modèle inspiré de la biophysique pour étudier la tâche de discrimination RDM. Dans ce modèle, l'intégration est effectuée à l'aide d'un mélange entre une excitation feedback (canaux N-methyl-D-aspartate (NDMA) avec une constante de temps relativement longue) et un mécanisme d'inhibition. Ce modèle a ensuite été réduit en une version à champ moyen qui est beaucoup plus rapide à simuler numériquement (Wong and Wang, 2006). Le modèle est constitué de deux populations neuronales dont une partie des neurones sont sélectifs pour un des deux choix, C_1 et C_2 (Figure 1.3). Les deux populations neuronales sont en compétition par l'intermédiaire d'interneurones inhibiteurs. Les deux populations sélectives reçoivent des entrées sensorielles conflictuelles, avec l'intensité du mouvement caractérisée par le niveau de cohérence c.

Figure 1.3 représente un exemple de la dynamique de ce modèle. Au démarrage du stimulus, les taux de décharges des deux populations sont proches et croissent jusqu'à ce qu'ils divergent l'un de l'autre. Cette divergence est due à une dynamique de type *winner-take-all* qui a lieu dans ce réseau par l'intermédiaire de l'excitation récurrente et l'inhibition. Le choix est basé sur la population qui gagne cette compétition. Dans la version de la tâche avec les temps de réaction, ceci est indiqué par le fait qu'une des deux populations dépasse un certain seuil (Figure 1.3.B).

Modélisation des données comportementales

Les réseaux attracteurs neuronaux reproduisent un bon nombre de résultats comportementaux. La figure 1.4 montre la variation de la performance et des temps de réponses en fonction de l'intensité du mouvement. Plus le stimulus est fort, plus la décision est rapide et correcte dans le modèle. Ceci est dû au fait qu'un stimulus fort va entraîner une croissance plus rapide de l'activité. Comme mentionné précédemment, pour obtenir les distributions des RTs pour les essais corrects et incorrects avec les DDMs, il faut introduire une variabilité entre chaque essai. La différence des distributions entre les deux types d'essais est un élément clef du réseau d'attracteurs car elle est naturellement présente dans

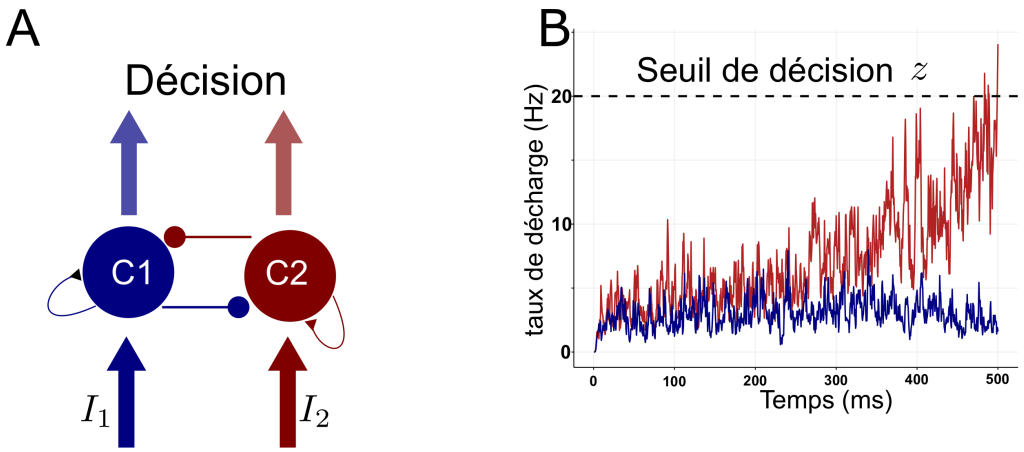


Figure 1.3: **Réseau neuronal attracteur de Wong and Wang (2006).** (A) Version schématique du circuit local de prise de décision. Deux populations neuronales (C_1 et C_2) sont en compétition par l'intemédiaire d'une inhibition latérale et sont sujettes à une excitation récurrente. Ce modèle correspond à une version champ moyen de Wang (2002). (B) Dynamique du réseau durant le processus de prise de décision. Chaque population neuronale montre une croissance de l'activité jusqu'à ce que l'une des deux atteigne le seuil. Dans ce cas, la population gagnante est C_2 et correspond au choix fait par le système.

ces modèles-ci.

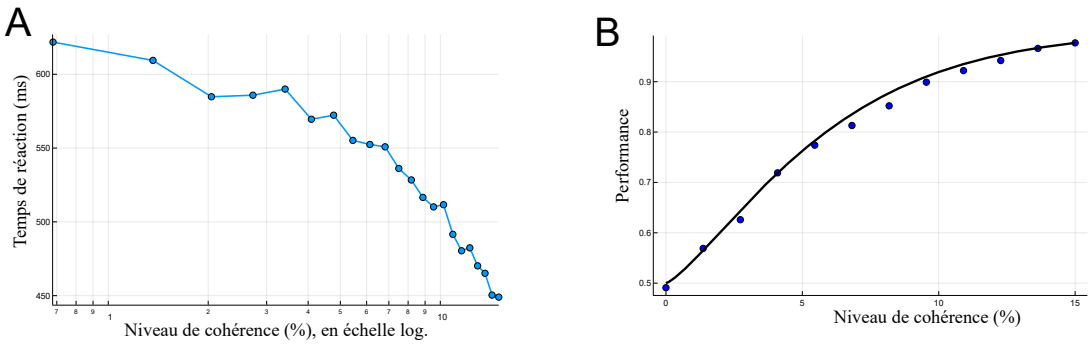


Figure 1.4: **Performances comportementales du réseau attracteur neuronal.** (A) Temps de réaction du circuit suivant le niveau de cohérence. Le niveau de cohérence est noté c et représente la différence de courant arrivant à chaque unité ($I_i \propto (1 \pm c)$). (B) Performances du système selon le niveau de cohérence.

La figure 1.5 représente le processus de décision dans le plan de phase (activité synaptique de la population 1 contre celle de la population 2). Lorsque la décision est en faveur de 1, le bassin d'attraction correspondant à cet attracteur est plus grand (représenté par la variété instable). Le système commence naturellement dans le bassin de 1 dans ce cas. Pour prendre la mauvaise décision, il faudrait que la dynamique traverse la frontière entre les deux bassins d'attraction. Traverser une frontière dans un système dynamique est un processus lent; ceci explique la raison pour laquelle les temps de réactions sont plus longs dans le cas des erreurs (Wong and Wang, 2006). Au contraire, une implémentation neuronale du DDM entraînerait un effet opposé (Mazurek et al., 2003b). Une autre différence majeure entre le DDM et le modèle à attracteur porte sur le comportement du système lorsque le stimulus est présenté durant une longue durée. Kiani et al. (2008) ont effectué une tâche RDM avec

des singes mais en faisant varier la durée des stimuli. Ils ont observé que les performances atteignent un plateau lorsque la durée du stimulus augmente. Ce phénomène n'est pas en accord avec les modèles de DDM, car dans ces modèles, les performances augmentent indéfiniment avec la durée du stimulus. Au contraire, les performances sont bornées pour les stimuli longs dans le modèle de réseau attracteur (Wang, 2002) car l'activité finit par atteindre un état attracteur.

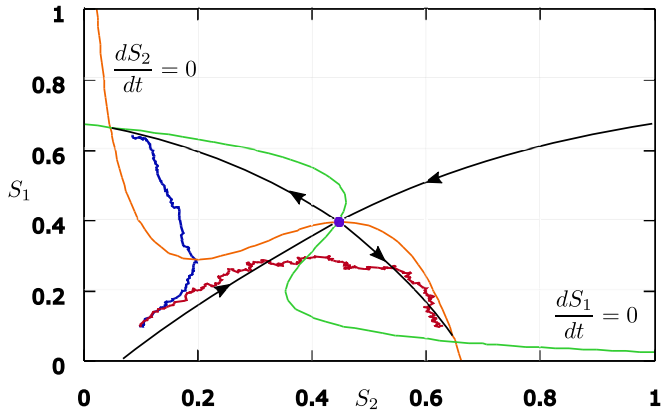


Figure 1.5: **Analyse dans le plan de phase du réseau attracteur pour un niveau de cohérence de** 10**% en faveur de la catégorie 1.** Les axes représentent les activités synaptiques dans le modèle. Les lignes verte et orange représentent les nullclines. Leur intersection au milieu correspond au point fixe instable. Les lignes noires représentent les variétés stables et instables. En bleu et rouge j'ai représenté les activités de deux essais différents.

Succès des modèles de circuits corticaux

Un avantage à utiliser les modèles de réseaux attracteurs pour modéliser les prises de décision est qu'il est possible d'étudier plus en détail les processus neuronaux de la prise de décision. Premièrement, on peut observer une chute de l'activité des neurones du LIP avant la croissance de l'activité. Le même comportement est observé dans les réseaux attracteurs (Wong et al., 2007). Dans cette étude, les auteurs étudient l'influence d'un afflux d'évidence sur le processus de décision. Ils ont montré que cette influence diminue avec le temps. Il s'agit d'un effet qui a été observé dans des expériences avec des singes (Huk and Shadlen, 2005; Wong et al., 2007).

Dans les modèles à attracteur, l'excitation récurrente doit être compensée par de l'inhibition. Ceci est effectué par l'inhibition synaptique latérale entre les populations neuronales. Hanks et al. (2006) ont effectué des microstimulations, chez les singes, des neurones du LIP sélectifs aux directions. Ils ont observé que, non seulement la décision vers la direction préférée des neurones était plus rapide mais la décision dans la direction opposée est plus lente également. Cette observation est congruente avec un modèle de réseau neuronal récurrent avec inhibition. Les modèles à attracteurs n'ont pas seulement été développés pour la tâche RDM mais également pour des tâches de discrimination somatosensorielle (Machens et al., 2005). De tels modèles peuvent reproduire l'activité persistante observée dans les neurones préfrontaux (Romo et al., 2004). De plus, ce type de circuit peut effectuer le calcul de discrimination, comme $f_1 > f_2$, pendant la période de comparaison (Machens et al., 2005).

Une limite aux modèles de diffusion et au modèle attracteur que j'ai présentés relève de la base biologique du seuil de la décision. Avant qu'une saccade visuelle ne soit effectuée

(dans le cas d'une tâche oculomotrice), les neurones du champ visuel frontal (FEF) et du superior colliculus (SC) émettent un burst de spikes (Hanes and Schall, 1996; Munoz et al., 2002). Ces neurones sont sélectifs de l'amplitude et de la direction de la saccade. En utilisant le cadre initial des attracteurs neuronaux, Lo and Wang (2006) ont étudié une version étendue de ce circuit qui utilise un système cortico-basal ganglia pour terminer le processus de décision. L'idée provient du fait que le passage du seuil par l'activité neuronale peut être détecté par des neurones dans le centre moteur. Dans le cas des mouvements saccadiques visuels, il s'agit principalement du superior colliculus et de la basal ganglia (Munoz and Schall, 2003). Dans ce modèle, le mécanisme de seuil est obtenu par l'intermédiaire d'une inhibition vers les aires corticales. Ce modèle suggère que la régulation du mécanisme de seuil constitue un mécanisme computationnel spécifique qui peut être obtenu dans un réseau qui connecte le cortex, la basal ganglia et le supérior colliculus. De plus, ce seuil peut être modulé par de la plasticité synaptique cortico-striatal. Ainsi, il peut être optimisé pour minimiser le temps de décision et le taux d'erreur.

1.2. Effets séquentiels et réseaux attracteurs

Les expériences typiques pour l'étude de la prise de décision consistent en une série d'essais successifs qui sont séparés par un court intervalle de temps. De nombreuses études ont mis en évidence la dépendance séquentielle dans les décisions perceptuelles. Dans cette section, je m'intéresserai à ces différents effets en proposant une modélisation à l'aide d'attracteurs neuronaux.

Dynamique de relaxation

Je considère un réseau neuronal récurrent qui simule un processus de décision grâce à de l'excitation récurrente et de l'inhibition. Le modèle est constitué de deux populations excitatrices qui sont sélectives pour une des deux catégories (correspondant au mouvement des points du RDM). Les deux populations s'inhibent l'une l'autre et sont soumises à une excitation récurrente. La dynamique consiste en une dynamique non-linéaire entre deux attracteurs (Wong and Wang, 2006; Berlemont and Nadal, 2019). Lors des tâches de prise de décision les neurones montrent une décroissance rapide de leur activité après que les sujets aient répondu(Roitman and Shadlen, 2002; Ganguli et al., 2008). Pour obtenir ce comportement dans le réseau à attracteur, il est nécessaire d'introduire une courte période de relaxation entre chaque essai. Je propose de modéliser cette période à l'aide d'un bref courant inhibiteur qui permet la décroissance de l'activité.

La figure 1.6 décrit le comportement du modèle attracteur lorsqu'il est soumis à un courant inhibiteur. Dans le cas d'un courant constant dans le temps (Figure 1.6 A-B-C-D), si le courant est trop faible, le système est inchangé. Si le courant est suffisamment fort, une bifurcation a lieu et seul un point fixe existe à présent. Pour un tel courant inhibiteur, le système est donc réinitialisé dans un état neutre en ayant perdu la mémoire de son état précédent. Dans le cas d'un courant dépendant du temps (Figure 1.6 E-F-G-H), le réseau voit sa dynamique de relaxation ralentie au cours du temps. La première partie de relaxation, rapide, permet de quitter l'état attracteur puis la seconde partie permet de garder une sorte de mémoire de l'état précédent.

Effets séquentiels

Avec cette relaxation, il est donc possible d'étudier les effets séquentiels observés lors de la prise de décision reproduits par les réseaux attracteurs. J'ai commencé par l'analyse des effets dits de biais séquentiels: les décisions sont biaisées vers la décision précédente lorsque les choix sont effectués en séquence. Lors de l'analyse de ce type d'effets dans le modèle, j'ai observé la présence de ces effets avec une variation selon le temps de relaxation

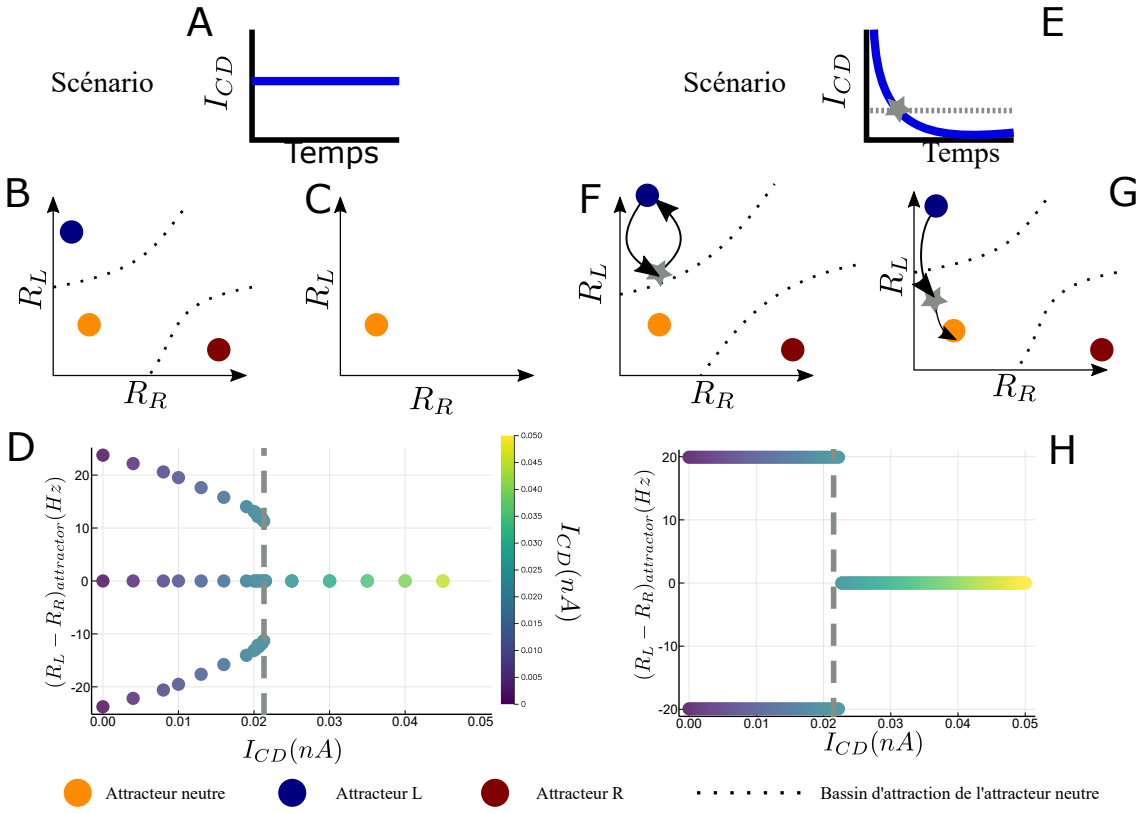


Figure 1.6: **Diagramme de bifurcation pour la prise de décision séquentielle pour deux types de courant inhibiteur.** (A) Scénario avec une valeur constante du courant inhibiteur (panneaux B, C et D). (B) Représentation plan de phase pour des faibles courants inhibiteurs. (C) Représentation plan de phase pour des forts courants inhibiteurs. (D) Etats attracteurs (différence d'activités, $R_L - R_R$) en fonction du courant inhibiteur. La ligne grise correspond à la ligne de bifurcation. (E) Scénario avec une valeur décroissante exponentiellement avec le temps du courant inhibiteur (panneaux F, G et H). (F) Représentation plan de phase pour des faibles courants inhibiteurs. (G) Représentation plan de phase pour des forts courants inhibiteurs. (H) Etats attracteurs du système à la fin de la relaxation. A gauche de la ligne grise, le système est bloqué dans sa décision initiale. A droite, le système est capable d'effectuer les décisions les unes à la suite des autres.

entre chaque essai et l'intensité du courant inhibiteur lors de la relaxation. Ceci m'a amené à étudier un autre type d'effet observé lors des expériences psychophysiques: les effets après-erreur (Danielmeier et al., 2011).

Les effets après-erreurs sont nombreux et consistent en des différences entre les temps de réponse et les performances des participants si leurs essais sont regroupés en deux groupes: les essais qui suivent des erreurs et les essais qui suivent des décisions correctes. L'effet le plus étudié expérimentalement consiste en le ralentissement après-erreur (PES). Les temps de réponse après une erreur sont plus longs que les temps de réponse après une décision correcte. De manière surprenante, cet effet est présent dans les tâches de décision où le participant n'est pas informé de l'exactitude de sa réponse.

Le modèle de réseau attracteur reproduit les effets après-erreur (Figure 1.7). Il met également en évidence une certaine région de paramètres où, au lieu d'un ralentissement après-erreur, le réseau montre une accélération après erreur. Il est important de noter que cet effet a également été observé chez les participants, qu'il s'agisse de singes ou d'humains.

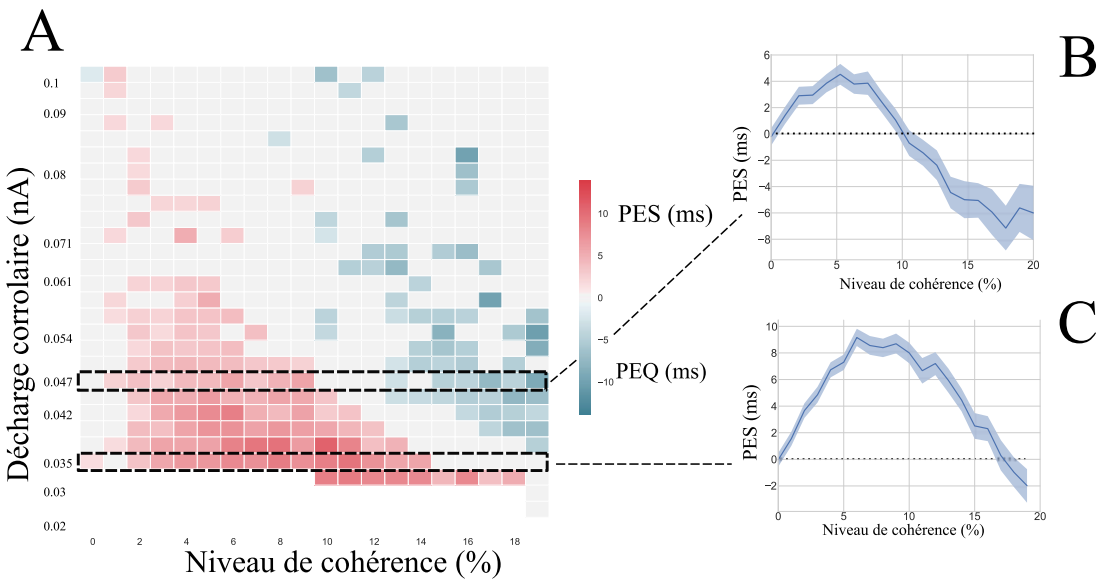


Figure 1.7: **Ralentissement après-erreur dans le modèle pour un intervalle entre stimuli de 500 ms.** (A) Diagramme de phase de l'effet PES. La zone inférieure blanche correspond aux paramètres pour lesquels la prise de décision en séquences n'est pas possible. Les carrés rouges correspondent à l'effet PES et en bleus à l'effet opposé. Les zones noires correspondent aux zones où les panneaux (B) et (C) zooment.

Jusqu'à présent il semblait contradictoire avec l'effet PES. Dans ce travail, j'ai montré que bien que ces deux effets soient opposés, ils peuvent être obtenus par le même système de prise de décision.

1.2.1 Dynamique non-linéaire

Pour reproduire les conditions expérimentales, le réseau neuronal ne reçoit aucune information sur sa réponse. Ainsi, l'effet après-erreur observé ne peut pas provenir d'interférences d'éléments externes. Il est surprenant qu'un tel système puisse exhiber, dans ces conditions, tous ces différents effets après-erreur. Cependant, ceci peut s'expliquer par la dynamique non-linéaire du modèle (Figure 1.8).

Si la décision est répétée (Figure 1.8.A et B), à la fin de la relaxation du système, les deux dynamiques après-erreur et après-correct se situent à l'intérieur du bon bassin d'attraction. Ainsi les taux d'erreurs de ces essais sont similaires. Cependant, les états atteints à la fin de la relaxation sont différents. Comparé à l'essai après-erreur, l'essai après-correct se situe plus proche de la frontière du nouvel attracteur associé à la décision suivante; ainsi cette nouvelle décision sera plus rapide. Dans le cas d'une alternance de réponse (Figure 1.8 C et D), les états atteints à la fin de la relaxation ne sont pas dans le bon bassin d'attraction. Pour l'essai après-correct, la dynamique est assez directe pour passer à travers la frontière; la décision est donc rapide. Pour le cas après-erreur, il s'agit du comportement opposé. Ainsi, c'est bien la dynamique non-linéaire du modèle qui explique les différents effets après-erreur qu'il reproduit.

Au-delà de reproduire les effets séquentiels lors de la prise de décision, le modèle effectue également certaines prédictions sur leurs conditions d'apparence : quand observe-t-on le ralentissement après-erreur ou bien l'accélération après-erreur ? Mais également certaines prédictions du point de vue neuronal, comme la corrélation entre la différence d'activité

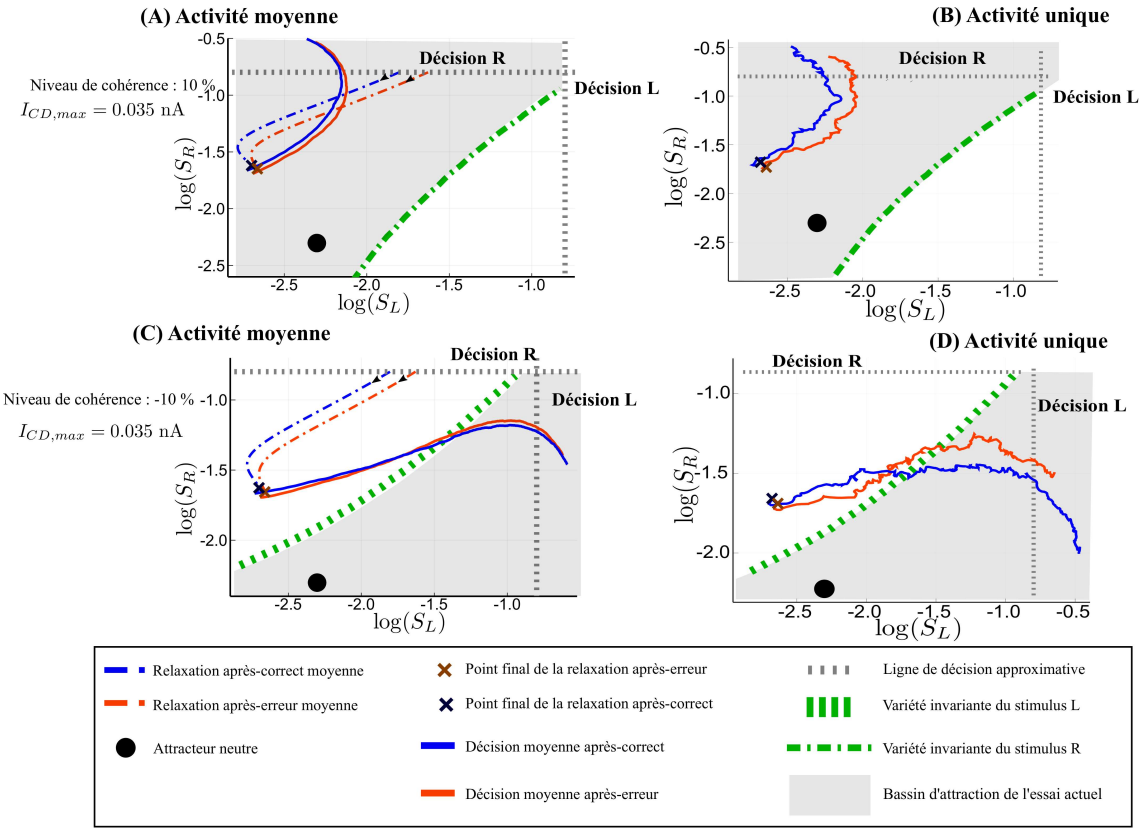


Figure 1.8: **Analyse des trajectoires après-erreur dans le cas PES.** Trajectoire dans le plan de phase des essais après-erreur (en rouge) et après-correct (en bleu).

entre les populations au moment de la décision et la force de l'effet observé lors de l'essai suivant. Ces différents résultats pourraient être testés expérimentalement dans le but de valider ou bien d'invalider le modèle.

1.3. Modélisation de la confiance

Dans une deuxième partie, je me suis intéressé à la notion de confiance des participants ont lorsqu'ils effectuent des décisions. Pour cela, j'ai mis en place, en collaboration avec Jean-Rémy Martin et Jérôme Sackur, une expérience psychophysique. Les participants sont face à un écran sur lequel s'affiche un cercle strié de nuances de gris (Figure 1.9). Les participants doivent décider si l'orientation des traits est dans le sens horaire ou anti-horaire. Dans certains des essais, il leur est également demandé d'indiquer leur degré de confiance en leur décision. La confiance reportée est corrélée avec les temps de réaction et les performances. Les décisions au niveau de confiance élevé sont en moyenne plus rapides.

Ces différents résultats ont été observés dans de nombreuses expériences de prise de décision. Cependant, ils n'ont jamais été étudiés dans le contexte des réseaux attracteurs. Dans ce travail, je propose une méthode pour calibrer les paramètres du modèle sur certaines données comportementales des participants: les temps de réponse moyens et les performances moyennes. Ceci permet, pour la première fois, de comparer quantitativement les résultats d'une expérience de prise de décision avec des humains et un modèle de réseaux attracteurs (Figure 1.10). La figure 1.10 montre que le modèle à attracteur reproduit quantitativement les différents comportement des participants, que ce soient leurs temps

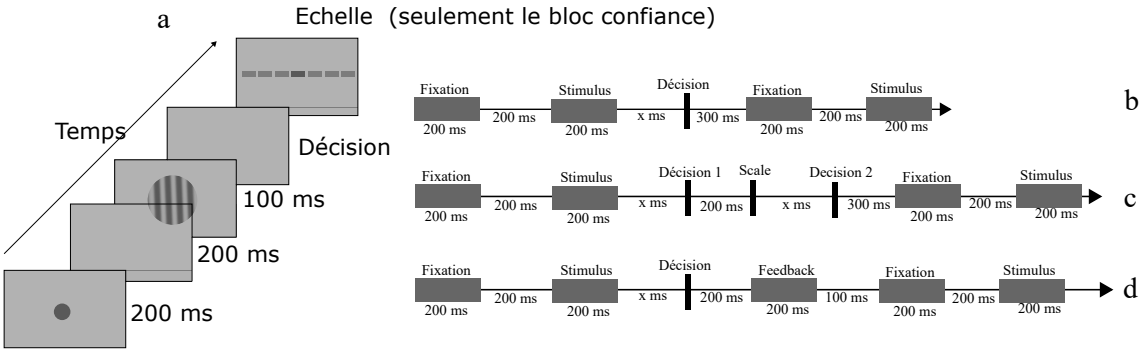


Figure 1.9: **Procédure de la tâche de catégorisation.** (A) Structure d'un essai: Après une période de fixation, le stimulus apparaît et les participants doivent effectuer une décision. Dans les essais avec confiance, après un délai, les participants indiquent leur confiance sur une échelle discrète à 10 niveaux. (B) Schéma temporel des blocs *purs*. (C) Schéma temporel des blocs *confiance*. (D) Schéma temporel des blocs *feedback*.

de réponse ou bien leurs performances. De plus, bien qu'assez proche les uns des autres, les paramètres du modèle pour chaque participant permettent de reproduire la diversité des comportements.

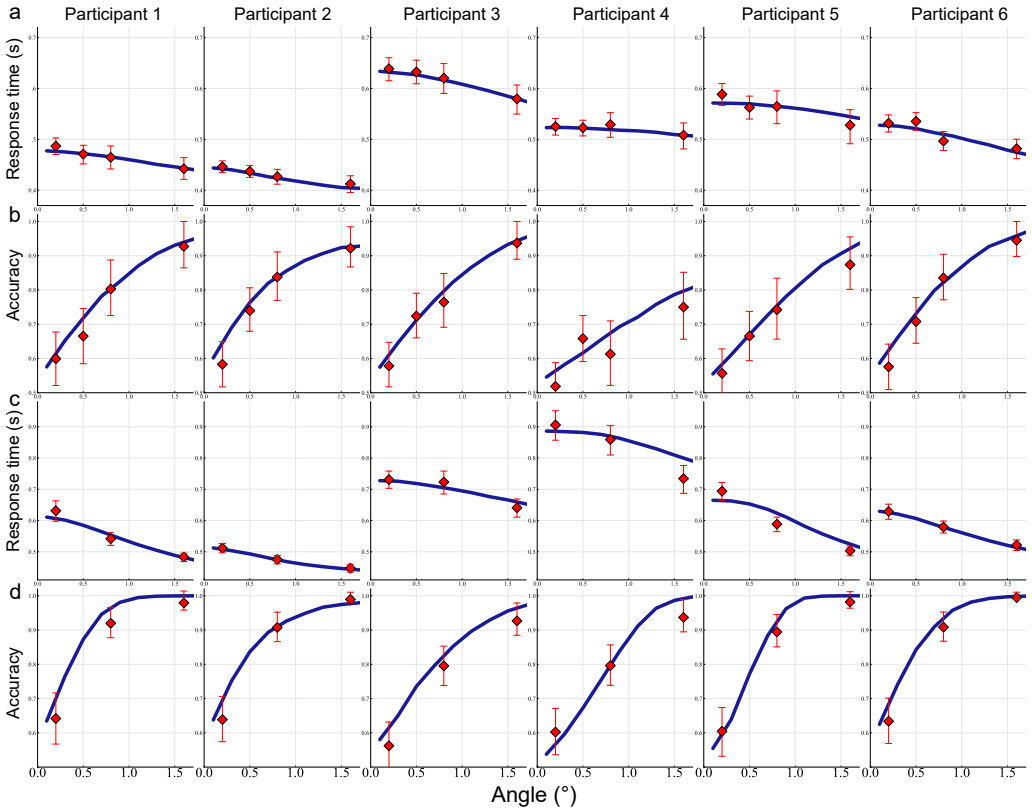


Figure 1.10: **Temps de réponse moyen (A,C) et performances (B,D) selon l'orientation absolue du stimulus**, dans le bloc *pure* (A and B) er *confiance* (C and D) blocks. Pour chaque participant, les données comportementales sont en rouge et le modèle calibré en bleu.

Deux résultats principaux ressortent de cette analyse. Premièrement, la confiance peut

être modélisée comme une fonction de la différence entre les activité neuronales. Je montre ceci en utilisant une méthode de détermination implicite de la relation entre confiance et activité neuronale, en me basant directement sur les données comportementales. Avec cette modélisation, le modèle reproduit les variations des temps de réaction et des performances selon le niveau de confiance des participants (Figure 1.11).

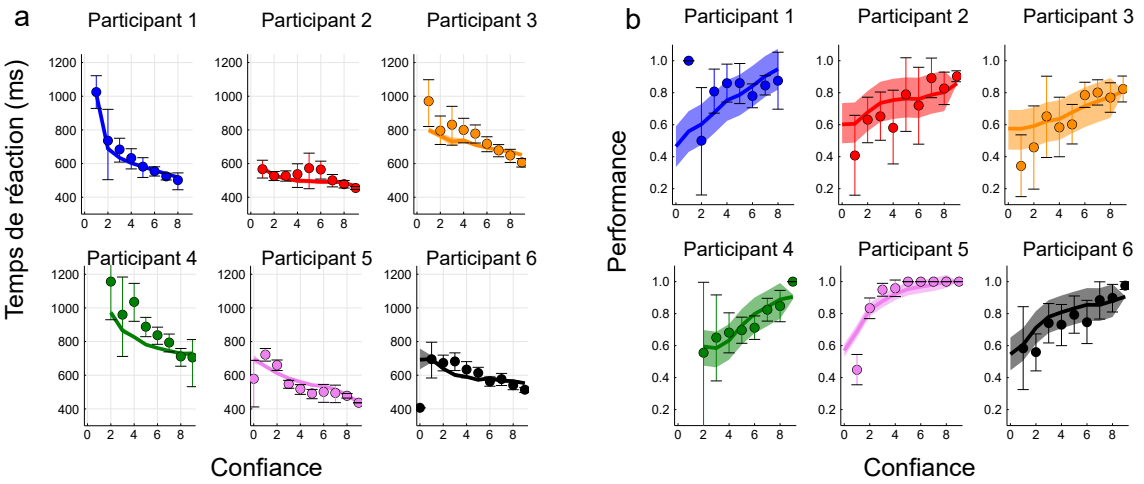


Figure 1.11: **Temps de réponse et performances en fonction de la confiance.** (A) Temps de réponses, (B) Performances. Chaque panneau représente un participant différent. Les lignes représentent les simulations numériques du modèle.

Le deuxième résultat concerne les effets séquentiels. En effet, la confiance à l'essai $n-1$ impacte l'essai n. Si l'essai correspond à un essai à forte confiance, alors l'essai suivant sera plus rapide (et vice-versa dans le cas d'un essai à faible confiance). De manière surprenante, le modèle reproduit également ces effets, alors qu'il n'a pas été calibré sur les séquences d'essais des participants mais sur les valeurs moyennes. Ceci met en évidence le fait que la dynamique non-linéaire du système permet de modéliser non seulement la confiance des participants mais également divers effets séquentiels.

1.4. Apprentissage d'un tâche de catégorisation et confiance

Dans les deux sections précédentes, j'ai présenté des résultats qui considéraient un modèle à attracteurs qui recevait une entrée sensorielle déjà caractéristique de la catégorie d'appartenance (de part le niveau de cohérence). La suite de mes travaux a consisté à analyser un système où le codage de la catégorie est effectué par une couche de codage neuronale (Figure 1.12). Dans un tel système, deux questions peuvent se poser. Quelles doivent-être les caractéristiques de la couche de codage ? Comment le système peut-il apprendre la tâche de décision ?

Pour la couche de codage, je me suis intéressé à deux distributions des neurones différentes: une distribution uniforme ou bien une distribution optimisée dans le but d'effectuer une tâche de catégorisation. Pour effectuer l'apprentissage, je considère un apprentissage Hebbien modulé par la récompense. Cependant, ce type d'apprentissage ne permet pas au système composé d'une couche de codage optimisé d'apprendre la tâche correctement car il obtient des performances moins bonnes que dans le cas d'une couche de codage uniforme. Je me suis donc intéressé à l'impact de la confiance sur le processus d'apprentissage. Lorsque la confiance est utilisée comme modulation de l'apprentissage Hebbien les performances du réseau sont grandement améliorées. La confiance module l'apprentissage de la façon suivante: si la confiance à l'essai n est plus grande qu'une

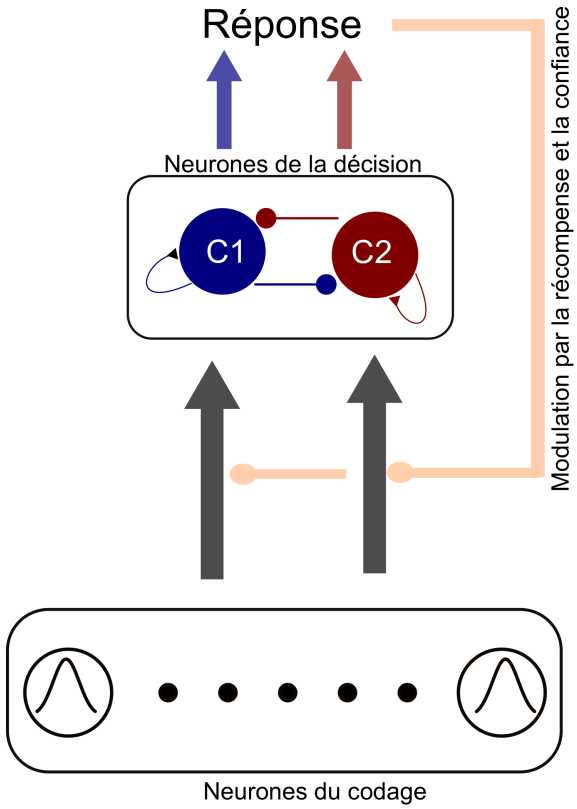


Figure 1.12: **Modèle du circuit neuronal.** (A) Schéma du modèle. Le réseau est composé de deux couches. Les neurones dans la couche de codage reçoivent directement l'entrée sensorielle. La couche de décision reçoit l'activité des neurones de codage grâce aux connections synaptiques. La catégorie de la décision est obtenue à l'aide d'une dynamique d'attracteurs en compétition. Les connections synaptiques subissent une plasticité synaptique Hebbienne modulée par un signal de récompense.

certaine valeur, alors le système n'effectue aucune mise à jour des poids synaptiques. Si l'essai correspond à une décision à faible confiance, la mise à jour des poids synaptiques a lieu.

Dans ces conditions, un réseau neuronal avec une couche de codage optimisée est plus performant qu'un réseau avec une couche de codage uniforme. Ainsi, la modulation par la confiance permet d'utiliser l'information qui a été optimisée lors du codage pour améliorer les performances. Il est également important de noter que ce type d'apprentissage ne nécessite pas de garder en mémoire les récompenses des essais précédents. En effet, la confiance se construit au fur et à mesure et cette quantité locale tient le rôle d'arrêt de l'apprentissage si les performances sont déjà maximales pour ce stimulus.

1.5. Méthode de la cavité pour les réseaux neuronaux

La dernière partie de mon manuscrit concerne un sujet différent de ce que j'ai présenté jusqu'ici. Il s'agit de l'étude de réseaux neuronaux à l'équilibre dynamique et partiellement connectés. Les réseaux neuronaux à l'équilibre dynamique ont fait l'objet de nombreuses études. Cependant dans la plupart des travaux, les connexions entre neurones sont aléatoires. Bien que ceci permette une simplification des calculs analytiques, ce n'est pas ce qui est observé expérimentalement. J'ai utilisé une méthode inspirée de la physique statistique, *la méthode de la cavité*, pour étudier de tels réseaux. Ceci m'a permis d'obtenir

le comportement de ces réseaux suivant les différents paramètres utilisés. J'ai mis en évidence deux zones du diagramme de bifurcation. Dans la première zone, il existe un point fixe stable vers lequel converge le système. Dans la seconde zone, une rupture d'ergodicité a lieu. Cette rupture existe à des niveaux de symétrie comparables avec ceux des régions corticales. De plus, dans cette zone, le système exhibe une dynamique de type verre en ayant, notamment, une très longue constante de temps d'oubli des conditions initiales.

Decision-making using attractor neural networks: focus on sequential effects

Part summary

At the beginning of the project, this part was only supposed to last for a few months. The goal was to study how an attractor network could perform sequences of decisions. In the end, this project lasted one year and half.

During a preliminary study of my PhD project I found some unexpected results. I had planned to study, during a few months, the relaxation dynamics of attractor neural networks in order to perform sequences of decisions. However, I found that attractor neural networks reproduce many non-trivial sequential effects that other models do not successfully account for. This led me to investigate these unexpected effects in more details.

During the first six months, I have studied the dynamical properties of an attractor network when it was subject to a relaxation between two decision trials. When humans perform decisions in sequences, many effects are observed, such as the fact that participants are slower after they have made an error. I have shown that an attractor neural network reproduces most of the observed sequential effects. In particular, for the first time, a model makes testable prediction on the different observation conditions of different post-error effects.

The next step of this work was to developp a decision-making paradigm in order to study confidence in decision-making. For the first time, I have shown that it is possible to fit an attractor neural network on human behavioral data in order to make quantitative comparisons. Attractor networks reproduce the behavioral measures of the participants: accuracies, response times and confidence. Moreover, I show that it explains the sequential effects due to confidence too: participants have a tendency to be faster after high-confidence trials.

The different results that I present in this part are published in:

- (Chapter 2) Berlemont, K., & Nadal, J. P. (2019). Perceptual decision-making: Biases in post-error reaction times explained by attractor network dynamics. Journal of Neuroscience, 39(5), 833-853.

- (Chapter 3 and 4) Kevin, B., Jean-Rémy, M., Jérôme, S., & Jean-Pierre, N. (2020). Nonlinear neural network dynamics accounts for human confidence in a sequence of perceptual decisions. Scientific Reports (Nature Publisher Group), 10(1).

2

Introduction to decision-making

2.1. Decision-making in cognitive science and neuroscience

A central idea to decision-making is time integration. The idea of temporal accumulation of evidence has a long history in cognitive psychology. Decision-makers typically take longer time to perform more difficult decisions (Hick, 1952; Vickers, 1970), and there is a tradeoff between speed and accuracy, i.e performances improve with slower response times (Wickelgren, 1977). However, it is not obvious to determine the timescale on which this accumulation of evidence occurs (Uchida et al., 2006).

One of the most-studied perceptual decision task is the random dot motion task (RDM). Participants (humans or monkeys) look at a screen where some dots are moving randomly and the others coherently into one of the two possible directions. The participants must report the net direction of motion by making a saccade in this direction. When the fraction of dots moving coherently is reduced, the decision becomes more difficult. The difficulty level of these tasks can be characterized by the coherence level c. This variable lies between 0 and 1 and corresponds to a 1D mapping of the task difficulty. For example, in the case of the RDM task, the coherence level would correspond directly to the fraction of dots moving coherently. In this task, when the difficulty is varied, reaction times can go from 300 ms to the order of the second (Britten et al., 1992, 1993; Roitman and Shadlen, 2002). The two-alternative-forced-choice (2AFC) task has been studied in rodents too, in the case of an olfactory discrimination (Lak et al., 2014). In such tasks, the rats needed to respond to the dominant component in binary odour mixtures. The first observation is that these decisions are fast, around 300 ms. The performances varied from chance level to almost 100% correct with the stimulus difficulty. However, the difference in reaction times between the easiest and the hardest stimuli is smaller for rats than for humans and monkeys, around 35 ms in Uchida and Mainen (2003) and 80 ms in Abraham et al. (2004). These different results highlight the diversity of time integration across tasks and species.

Electrophysiological studies in behaving monkeys have been performed in order to understand the neural mechanisms underlying perceptual decision-making (Roitman and Shadlen, 2002; Gold and Shadlen, 2007). Research has focused on cortical neurons that encode task-relevant sensory signals in different perceptual decision-making tasks, such as the random dot motion task or sensorimotor tasks (Figure 2.1.A and Figure 2.2.A).

Sensory neurons in the visual area MT encode the motion direction of the stimulus (Newsome et al., 1989; Britten et al., 1992, 1993), but the decision process does not occur in this area. Shadlen and Newsome (1996) found that the activity of neurons in the lateral intraparietal cortex (LIP) was correlated with the monkey's perceptual choice. Moreover, in reaction time version of the task (when response times are controlled by the monkeys), a number of findings suggest that LIP neurons mediate decisions between rival saccadic decisions (Roitman and Shadlen, 2002; Huk and Shadlen, 2005; Gold and Shadlen, 2007;

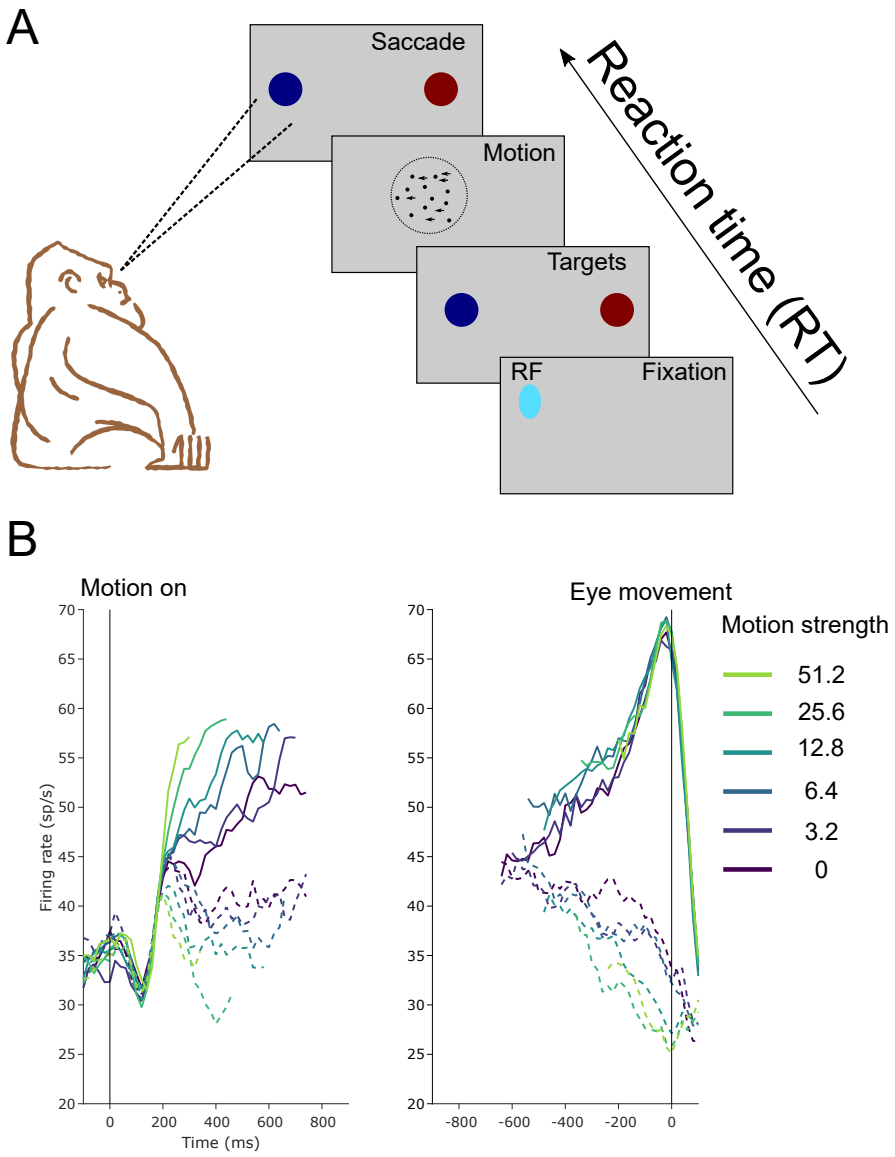


Figure 2.1: **Reaction time version of the random dot motion discrimination task.** (A) The monkey views a set of dots moving across the screen and decides the net direction of movement. The decision is indicated by a saccadic eye movement to one of the two peripheral target. The light blue field corresponds to the receptive field of one of the recorded LIP neurons. (B) Average response of LIP neurons during decision. The data are from Roitman and Shadlen (2002) and are publicly available. The average firing rate of 54 LIP neurons is shown for 6 degrees of difficulty. The firing rates are grouped by motion difficulty and direction of choice (the dashed lines corresponding to the choices out of the receptive field of the neuron). The left panel represents the average firing rate during decision formation starting from motion onset. The right panel shows the average firing rate centered at the time of the eye movement.

Huk and Meister, 2012). The average activity of the LIP neurons selective for the saccadic target increased from the stimulus onset until the saccadic eye movement (Figure 2.1.B). This buildup rate depends on the quality of the sensory information, with stronger evidence associated to steeper slopes of information integration. Finally, the decision choice is

made when the firing rate of the LIP neurons (corresponding to this choice) reaches a threshold that is independent of the signal quality and of the response time. In addition, microstimulations on MT and LIP neurons have an effect on accuracies and reaction times which is consistent with the idea that LIP neurons integrate sensory information (Ditterich et al., 2003; Hanks et al., 2006). These different studies support the notion that LIP neurons act as a neuronal integrator. Despite the stochastic variability of LIP neurons between trials, in average it displays a ramping of activity during decision-making.

Visual perception is not the only type of sensory modality that has been studied in decision-making tasks. Another task consists in a vibrotactile frequency (VTF) discrimination task (Hernández et al., 2002; Romo et al., 2004). In this task (Figure 2.2.A), the monkeys need to report a decision based on two different vibrations f_1 and f_2. The vibrations are received sequentially and the monkeys must indicate whether the frequency f_1 was greater than f_2 or not. In order to correctly perform such a task, the monkeys are required to hold in working memory the frequency of the first stimulus during the delay period. In these studies the researchears found that prefrontal cortex neurons showed persistent activity during the delay. A *plus cell* will have its firing rate monotically increasing with stimulus frequency during delay, and a *minus cell* will show a decrease in activity (Figure 2.2.B). This behavior changes during the decision as the neural activity becomes binary: some neurons will show high firing rates if $f_1 > f_2$ and low firing rates otherwise, and some neurons will show the opposite trend. It highlights the fact that the same circuit is involved in both working memory and decision-making (Lemus et al., 2007).

New approaches to study perceptual decision-making

The presented studies show that LIP, and more generaly prefrontal cortex form part of a circuit that is involved in implementating simple decision-making tasks. However there is still debate about where this accumulation process takes place. LIP neurons exhibit various response motives to sensory inputs, patterns that can be modulated by sensory and motor-related variables (Bennur and Gold, 2011; Park et al., 2014). It is thus possible that LIP neural responses only ressemble an evidence accumulation process when averaged together. Recently, Katz et al. (2016) have shown that unilateral pharmacological inactivation of the LIP has negligible effect on behavior. However, inactivation of the LIP area strongly impacted sensory processing during decision-making, more than motor aspects (Zhou and Freedman, 2019).

To adress these different questions, researchers have started to study perceptual decision-making in rodents. Indeed, rodents exhibit various elements of perceptual decision-making such as evidence accumulation (Brunton et al., 2013). It is possible to consider many different sensory modalities such as visual stimuli, auditory (Brunton et al., 2013), tactile (Guo et al., 2014) and olfactory (Uchida et al., 2006) using state of the art technologies (Guo et al., 2014; Znamenskiy and Zador, 2013). These new methods have allowed to study multi-area during decision-making and to characterize how single-cells react to sensory inputs and transform them in decision variables (Brody and Hanks, 2016). Future experiments will allow to characterize more specifically the interactions between multiple brain regions during decision-making (Siegel et al., 2015).

2.2. Models of decision-making

Many models have been proposed to explain decision-making in humans and animals. There are two categories of models: dynamical models and non-dynamical models. Here, I will focus on dynamical models as they are more adequate to model the neuronal dynamics during decision-making. Non-dynamical models will be discussed in Chapter 5.

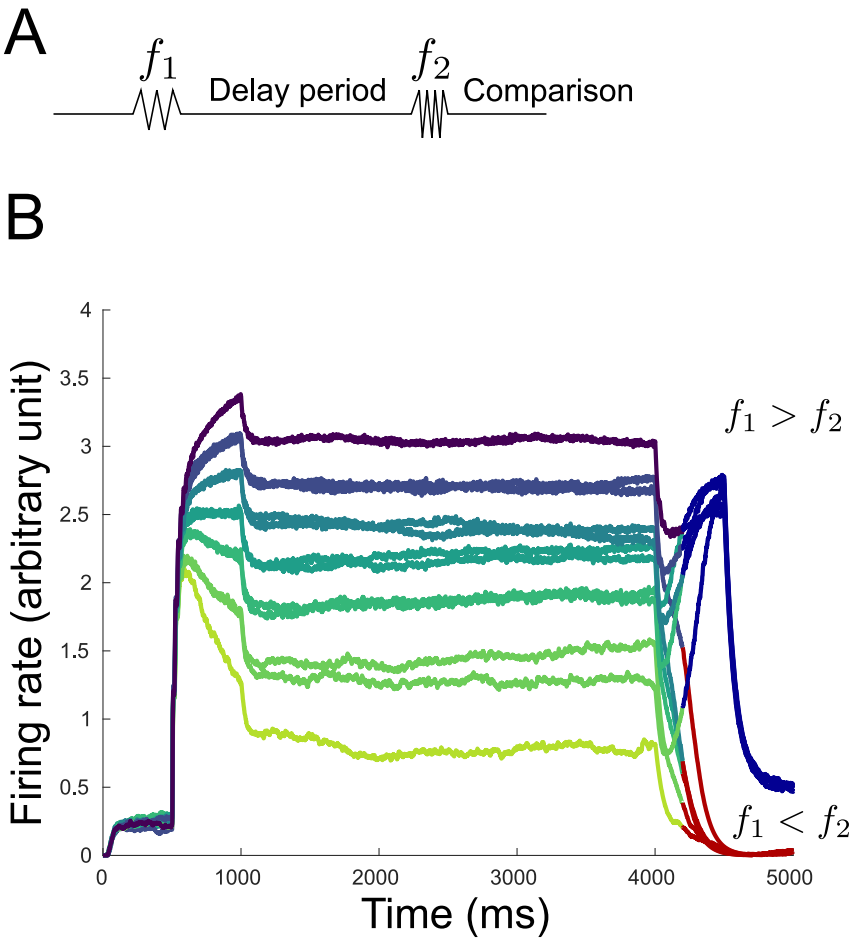


Figure 2.2: **Vibrotactile discrimination task.** (A) Schematic diagram of the task. Two vibrations with frequencies f_1 and f_2 are presented one after the other with a delay between them. The monkeys have to decide which one is faster than the other. (B) Simulated activity of a plus neuron (model from (Machens et al., 2005)). Red trials correspond to trials where the network decided $f_1 < f_2$, and blue trials to the opposite. Each color of firing rates corresponds to a different value of f_1. The model reproduces the result from Romo et al. (2004) (see Figure 1 of Machens et al. (2005)).

2.2.1 Drift-diffusion model

Abstract mathematical tests have been developped to decide between two probabilistic hypotheses, such as the sequential probability test (Wald and Wolfowitz, 1948). This test is optimal in the sense that it achieves a desired error rate with the minimum mean decision time. According to this test, decisions are initiated when cumulative estimates of noisy evidence variables reach a specific threshold (Ratcliff, 1978; Bogacz et al., 2006; Gold and Shadlen, 2007). The most-used model within the race framework is the drift-diffusion model (DDM). It consists in an unique integrator that accumulates the difference between the evidences for the two alternatives. The choice is made when the level of activity of the integrator exceeds a specific threshold, positive or negative depending on the alternative (Figure 2.3).

The success of the DDM framework is that it can decompose observed choice behaviors into a cognitive process. As a dynamical model, it takes into account decision accuracies

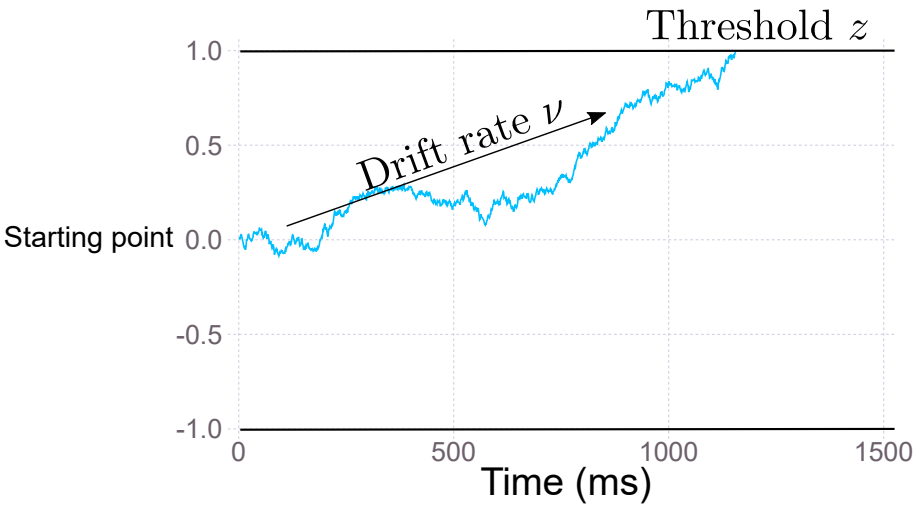


Figure 2.3: **Drift diffusion model.** Example of the dynamics of the DDM. The two black lines denote the threshold z and $-z$ corresponding to the two alternatives. The race is ended when it reaches one of the two boundaries. In this example, the decision made corresponds to a correct trials as the drift rate ν was chosen positive.

and response times and can address the speed-accuracy tradeoff. DDM will integrate for a shorter time if the evidence for the winning alternative is strong with respect to the losing alternatives. Respectively, it will take longer time to reach a decision if the difference of evidence is small. The model is based on the following parameters. The starting point a of the evidence accumulation represents a possible bias for one of the two alternative. This accumulation is performed at a certain drift rate ν and is related to the quality of information present within the stimulus. Typically, in the RDM experiment, this parameter would differ depending on the strength of the dots motion. The boundary represents the level of caution, the more this parameter is high, the less the model is sensitive to noise and the less the model makes the wrong choice. Finally, the last parameter consists in the non-decision times that are an additive lag representing the motor lag of the participants for example (Luce et al., 1986).

DDMs have been succesfully used to account for behavioral data in a wide-range of decision-making paradigms (Ratcliff and Rouder, 2000; Ratcliff et al., 2003, 2004). One should note that DDM does not only reproduce the error rates and response times, but the shape of the response times (RTs) too (Ratcliff, 1978). However, it is necessary to add across-trial variability for the different parameters in order to correctly model the RTs distributions of correct and error trials. Otherwise, due to the linearity of the model, the shapes of these distributions are strictly identical which is not what experimental studies have found (Ratcliff and Tuerlinckx, 2002).

2.2.2 Recurrent cortical circuit

There are two major criticisms that can be made about the DDM framework, when the goal is to model neural activity during decision-making. First, neural activity is non-linear (Figure 2.1.B) but the DDM is strictly linear. Secondly, DDMs have no biophysical foundations and do not explain how this integration mechanism is implemented in the brain. Different models have been proposed to account for the cortical processes of decision-making: Shadlen and Newsome (2001); Usher and McClelland (2001); Wang (2002). I will focus on the approach of Wang (2002), as the two other models consist more in an

extension of the DDM framework.

Attractor network

Neurons in LIP and prefrontal cortex have shown to display directionally tuned activity (Gnadt and Andersen, 1988; Funahashi et al., 1989). This suggests that a common mechanism could underlie working memory and decision-making. One mechanism to generate persistent activity similar to working memory is the one of strong recurrent excitation that will give rise to stimulus-selective attractor states in a local cortical circuit (Freeman, 1995; Goldman-Rakic, 1995; Brunel and Wang, 2001). Wang (2002) applied a biophysically based model to simulate the random dot motion discrimination task. In this model, the integration is achieved through a mix between feedback excitation (N-methyl-D-aspartate (NMDA) channels with relatively long time constants) and inhibitory mechanism. Each neuron consists in a leaky integrate-and-fire neuron. There are three neural populations: two excitatory ones and one inhibitory. In the two excitatory populations, some of the neurons are selective to one direction of the random dot motion. The competition between the populations is achieved through the synaptic connections with the inhibitory population. This architecture leads to a mechanism of information integration during the stimulus presentation. The working memory effect is achieved by having the network activity trapped into an attractor state when the stimulus is removed.

Later, this model has been reduced to a mean-field version that is much faster to simulate (Wong and Wang, 2006). The model consists of two neuronal pools with subpopulations of spiking neurons selective for the two choices, denoted by C_1 and C_2 (Figure 2.4.A). The two neural populations compete with each other through feedback inhibition from interneurons. Both selective neural populations receive conflicting sensory inputs, with the motion strength characterized by the quantity c (called coherence level).

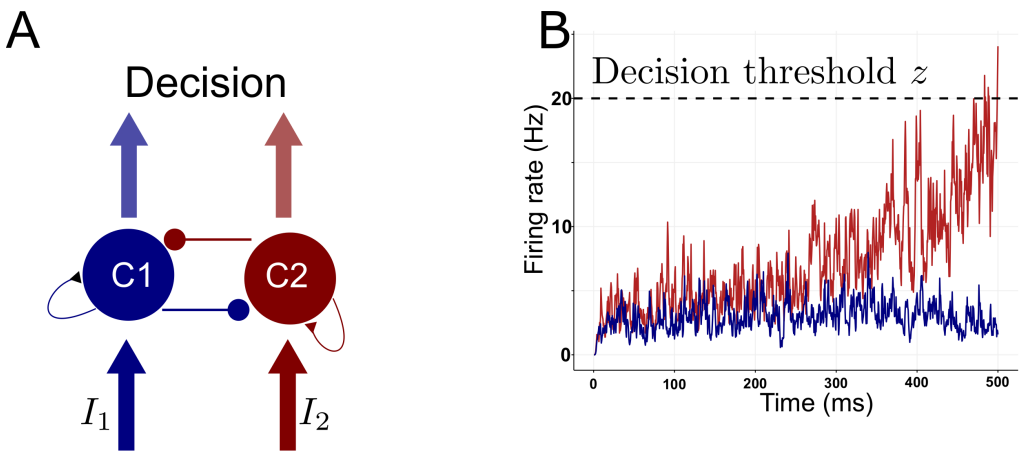


Figure 2.4: **Attractor neural network of Wong and Wang (2006).** (A) Schematic version of the local circuit of decision-making. Two neural pools (C_1 and C_2) compete with each other through lateral inhbition and are subject to recurrent excitation. This model corresponds to the mean-field version of Wang (2002). (B) Dynamics of the network during decision-making process. Each neural population shows a ramping up of activity until one of the two wins the competition and reaches a decision threshold. In this case, the winnning population is C_2 and corresponds to the choice made by the network.

Figure 2.4.B represents an example of the dynamics of this model. At the stimulus onset, the firing rates of the two populations lie together and ramp up until they diverge from each other. The divergence is due to the winner-take-all dynamics that occurs in this

network through recurrent excitation and feedback inhibition. The choice is based on which population wins the competition. In the reaction time version of the task, this is indicated by the fact that one of the two populations crosses a fixed threshold (Figure 2.4.B).

Modelling behavioral data

Attractor neural networks have been shown to account for many of the behavioral results. Figure 2.5 shows the variation of accuracies and reaction times with respect to motion strength. As expected, the stronger is the stimulus, the faster is the decision time of the model. This is due to the fact that stronger stimuli lead to a stronger ramp up of activity. As mentionned previously, to obtain the different shapes of RT distributions for correct and error trials within the DDM framework, one needs to implement a variability of the parameters across trials. One key feature of attractor network is that this difference in distribution is naturally present.

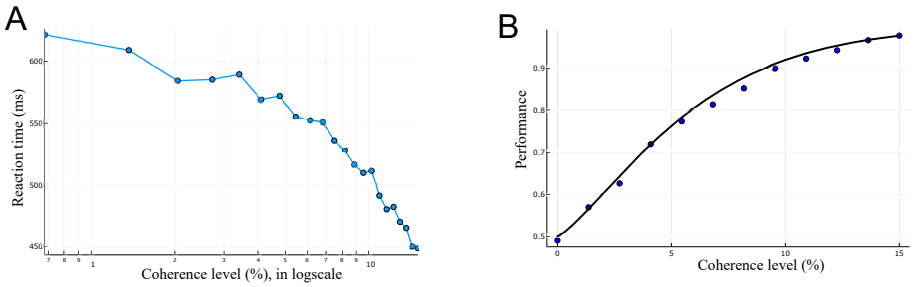


Figure 2.5: **Behavioral performances of the attractor neural network.** (A) Reaction times of the circuit with respect to coherence levels. The coherence level is the variable c and represents the difference in motion strength that arrives at both units ($I_i \propto (1 \pm c)$). (B) Performance of the model with respect to the coherence level.

Figure 2.6 represents the decision process in the phase-plane space (firing rates of population 1 against population 2). When the decision is in favor of 1, the basin of attraction corresponding to this attractor is larger, denoted by the unstable manifold. The system starts naturally in the basin of attraction of 1. To make the wrong decision, it would need to cross the boundary between the basins. Crossing a boundary between basins of attraction is a slow process, and that explains why the reaction times are slower in error trials compared to correct trials (Wong and Wang, 2006). In contrast, a neural implementation of the diffusion model would yield the opposite effect (Mazurek et al., 2003b). Another major difference between diffusion model and attractor neural network consists in the behavior at long duration of stimulus. Kiani et al. (2008) performed a motion discrimination task in monkeys but with variable stimulus duration. They found that performances reach a plateau when the duration of stimulus increase. This phenomenon is not consistent with the DDM framework as performance can improve indefinitely with stimulus time. In contrast, performances are bounded for long stimulus durations in the attractor neural network (Wang, 2002) as the ramping activity finally reaches an attractor state.

Successes of cortical circuits models

One advantage of using attractor neural network to model decision-making is that it can be used to look more closely at the neuronal process of decision-making. First, in Figure 2.1.B one can observe a drop in the firing rate of the LIP neurons before the ramping up of activity. The same behavior can be observed in attractor networks (Wong et al., 2007). In this study, the authors studied the influence of new arriving evidence during a decision

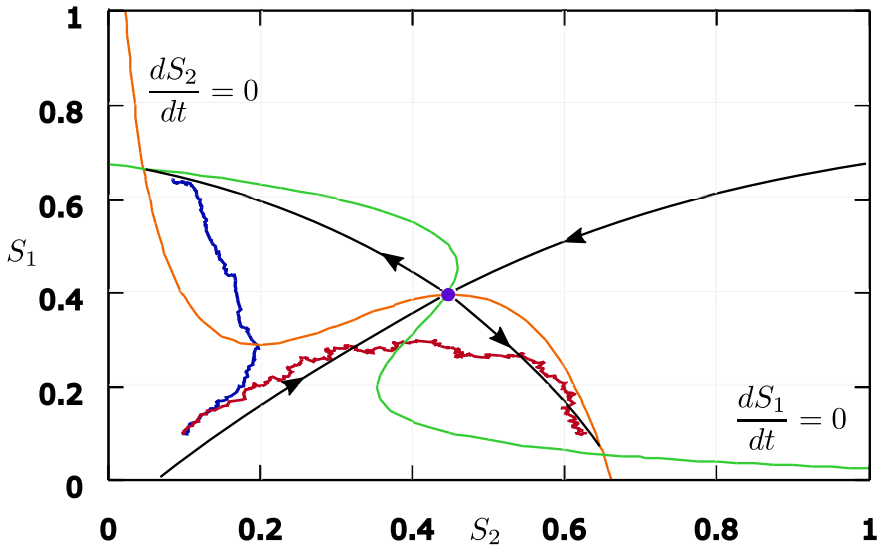


Figure 2.6: **Phase-plane analysis of the attractor network for a coherence level of 10% in favor of category 1.** The axes represent the synaptic activities in the model. The green and orange line are the nullclines. Their intersection at the middle corresponds to the unstable fixed point. The black lines show the stable and the unstable manifolds. In blue and red I have plotted the activities for two different trials.

process. They found that the influence diminishes over time, an effect that has been observed in monkey experiments too (Huk and Shadlen, 2005; Wong et al., 2007). This time-shift invariance can not be accounted for by diffusion models, as such models predict in fact the opposite effect (Wong et al., 2007).

In the attractor neural network, recurrent excitation must be balanced by feedback inhibition. This is mediated by lateral synaptic inhibition between the neural pools. In a monkey experiment, Hanks et al. (2006) performed microstimulation of LIP neurons selective to direction. They found that not only the decision towards the preferred direction was faster, but it slowed down the decision in the opposite direction too. This observation is consistent with a recurrent neural network with feedback inhibition. Attractor neural networks have not just been developped for the random dot motion discrimination task but for the somatosensory discrimination too (Machens et al., 2005). It has been shown that reciprocal inhibition between two neural pools could exhibit persistent activity as observed in prefrontal neurons (Romo et al., 2004). Moreover, this circuit can perform the discrimination computation, e.g $f_1 > f_2$, during the comparison period (Machens et al., 2005) (Figure.2.2.B).

One limit to both diffusion models and the attractor neural network I presented is the biological substrate of decision threshold. Before a saccade is made (in the case of an occolumotor task), neurons in the frontal eye field (FEF) and superior colliculus (SC) fire a burst of spikes (Hanes and Schall, 1996; Munoz et al., 2002). These neurons are selective of saccade amplitude and direction. Using the initial framework of the attractor neural network, Lo and Wang (2006) have studied an extended version of this circuit that involves the termination of the decision process through a cortico-basal ganglia system (Figure 2.7). The idea is that threshold crossing of ramping activity may be detected by neurons in a downstream motor command center, which is presumably the superior colliculus and the basal ganglia in the case of saccadic eye movements (Munoz and Schall, 2003). In this

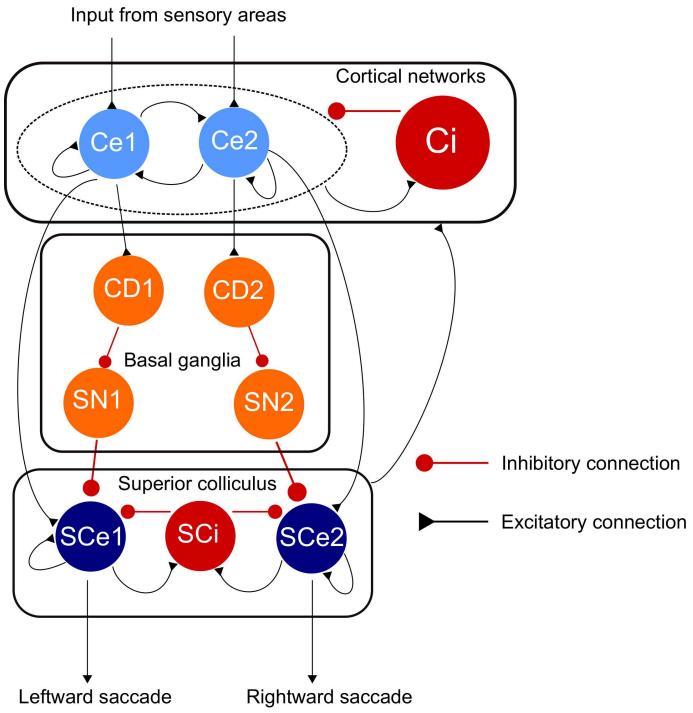


Figure 2.7: **Model architecture of Lo and Wang (2006).** Neural pools in variation of blue represent excitatory pools and in variation of red they represent inhibitory populations. Neural pools in the cortical area receive a sensory input and compete through lateral inhibition due to interneurons. They project to caudate nucleus (CD) in the basal ganglia and superior colliculus (SC). The saccade and termination of the decision-making is performed by the superior colliculus that projects back to cortical networks through inhibitory connections.

model, a threshold mechanism is achieved through an inhibitory feedback on the cortical areas. This model suggests that the regulation of a decision threshold constitutes in a specific computational mechanism that can be achieved in a network of interconnected cortex, basal ganglia and superior colliculus. Moreover, this threshold can be modulated by cortico-striatal synaptic plasticity. Therefore it can be tuned in order to minimize the decision time and the error rate.

In this first part, I address different aspects of decision-making using an attractor neural network. Many studies have argued that attractor networks were equivalent to diffusion models, see e.g. Bogacz (2007). However, this result is based on a assumption that is often forgotten: the system needs to lie at the bifurcation. Hence, it would result in a very fine tuning of the different parameters. I address key differences between attractor networks and diffusion models by studying principaly sequential effects (effects that are observed during decisions that are made in sequences). More specifically, I focus on the effect of post-error slowing (Danielmeier and Ullsperger, 2011) and confidence.

3

Sequential effects and attractor neural networks

Typical experiments on perceptual decision-making consist of series of successive trials separated by a short time interval, in which performance in identification and reaction times are measured. Several studies have demonstrated strong serial dependence in perceptual decisions between temporally close stimuli. This chapter is composed of two parts. In the first part, I adress the sequential effects that are called *choice history biases*. After a brief introduction of the different experimental observations, I show that an attractor network can perform decision-making in sequence through a relaxation dynamics and that it can explain these sequential effects. The second part consists in the study of effects called *post-error effects*. I would like to mention that the results in this part have been very surprising. When I studied the sequential effects within the attractor network I did not expect to observe post-error effects. What started as an initial step in the project in order to verify that attractor neural networks could perform sequences of decisions - turned out to last one year in order to fully understand the non-linear dynamics of the system. The more surprising result was that, without any additional or tuning of parameters, the model reproduces post-error effects with the correct order of magnitude.

3.1. Repetition biases

3.1.1 Cognitive aspect

The most studied protocol is the one of reaction time version of the two-alternative forced-choice (2AFC) task (Ratcliff, 1978; Laming, 1979; Shadlen and Newsome, 1996; Ratcliff and Smith, 2004). When subjects are instructed that the stimuli sequences are random, their reaction times (RTs) and error rates (ERs) still depend on the previous trials in a systematic way (Laming, 1968). The trials are influenced by whether or not previous choices led to positive outcomes (Rabbitt and Rodgers, 1977; Dutilh et al., 2012), the confidence in them (Desender et al., 2018a), as well as the stimulus category that was previously selected (Fründ et al., 2014; Urai et al., 2019). The influence of the previous category is called *choice history biases*. Choice history biases during perceptual decision-making are found in humans (Urai et al., 2019), monkeys (Gold et al., 2008) and rodents (Odoemene et al., 2018).

These effects can be categorized in two categories: first-order if it is caused by the previous trial, higher-order if it is due to trials earlier in the sequence. These sequential effects vary systematically with the response-stimulus interval (RSI) (Kirby, 1972; Soetens et al., 1985). For example, for shorts RSIs, repetitions of choices are faster than alternations (Figure 3.1.A). As RSI increases, this effect diminishes (Figure 3.1.B) and can sometimes leads to the opposite effect: first-order alternations become faster (Laming, 1968; Kirby, 1976).

This transition between faster repetitions and faster alternations has been observed in

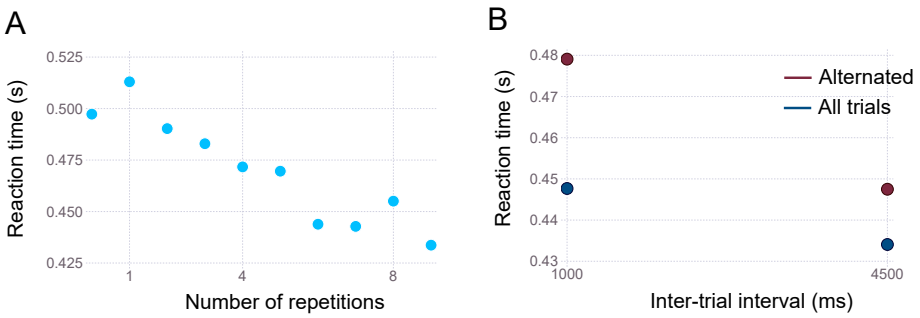


Figure 3.1: **Sequential effects in 2AFC tasks.** (A) Data from Berlemont et al. (2020). The x-axis denotes the number of previous trials where the choice have been repeated. The y-axis represents the mean reaction times of the participants after n repetitions. (B) Data from Bonaiuto et al. (2016). This panel compares the mean reaction times between alternated trials (the subject makes the opposite choice to the previous trial) and all the trials for two different inter-trial intervals. As expected, the sequential effect decreases when the time between two trials increases.

many experiments (Soetens et al., 1985; Cho et al., 2002; Jentzsch and Sommer, 2002). However, the critical response-to-stimulus interval at which it occurs is not well defined. Some studies reported an RSI around 1 second, and other shorter RSIs. Yet, for such long RSIs, higher order are more complicated. For example, after a sequence of choice A-B-A-B-A, the response of the participants is faster to B and slower to A (Bertelson, 1961; Cho et al., 2002). It seems that reaction times become faster if the presented stimulus confirms the subjects expectancy. Therefore, the difference between short RSIs and long RSIs could be due to a difference in how the task is analyzed by the subject.

3.1.2 Neural correlates of sequential effects

3.1.3 Models of sequential effects

The sequential effects have been mostly studied within the framework of statistical models of accumulation of evidence (Farrell and Ludwig, 2008; Goldfarb et al., 2012; Dutilh et al., 2012; Urai et al., 2019). However, such models do not work sequentially as the network is reset between each trials. In order to analyze the sequential effects, a drift-diffusion model is fitted independtly of the different conditions (for example between repetitions and alternations). Behavioral data can be fitted by different choices of starting points, and possibly of different thresholds (Goldfarb et al., 2012). This does not explain how the different modifications of parameters between the models could be implemented within the brain. In this section, I will present a modification of the attractor neural network that allows the model to perform sequences of decisions. I will analyze the sequential effects that are observed in the model.

A reduced recurrent network model for decision-making.

I consider a decision-making recurrent network of spiking neurons governed by local excitation and feedback inhibition (Figure 2.4), as introduced and studied in Compte et al. (2000) and Wang (2002). Since mathematical analysis is harder to perform for such complex networks, without a high level of abstraction (Miller and Katz, 2013), one must rely on simulations which, themselves, can be computationally heavy. For the analysis, I will make use of the reduced firing-rate model of Wong and Wang (2006) obtained by a systematic reduction of the detailed biophysical attractor network model. The reduction aimed at faithfully reproducing not only the behavioral behavior of the full model, but

also neural firing rate dynamics and the output synaptic gating variables. This is done within a mean-field approach, with calibrated simplified F-I curves for the neural units, and in the limit of slow NMDA gating variables motivated by neurophysiological data.

Since this model has been built to reproduce as faithfully as possible the neural activity of the full spiking neural network, it can be used as a proxy for simulating the full spiking network (Engel and Wang, 2011; Deco et al., 2013; Engel et al., 2015). Here, I mainly make use of this model to gain better insights into the understanding of the model behavior. In particular, one can conveniently represent the network dynamics in a 2-d phase plane and rigorously analyze the network dynamics (Wong and Wang, 2006).

The model consists of two competing units, each one representing an excitatory neuronal pool, selective to one of the two categories, L or R, corresponding to the dots of the RDM task going coherently leftward or rightward. The two units inhibit one another, while they are subject to self-excitation. The dynamics is described by a set of coupled equations for the synaptic activities S_L and S_R of the two units L and R:

$$i \in \{L, R\}, \quad \frac{\mathrm{d}S_i}{\mathrm{d}t} = -\frac{S_i}{\tau_S} + (1 - S_i)\,\gamma f\left(I_{i,tot}\right) \tag{3.1}$$

The synaptic drive S_i for pool $i \in \{L, R\}$ corresponds to the fraction of activated NMDA conductance, and $I_{i,tot}$ is the total synaptic input current to unit i. The function f is the effective single-cell input-output relation (Abbott and Chance, 2005), giving the firing rate as a function of the input current:

$$f\left(I_{i,tot}\right) = \frac{aI_{i,tot} - b}{1 - \exp\left[-d\left(aI_{i,tot} - b\right)\right]} \tag{3.2}$$

where a, b, d are parameters whose values are obtained through numerical fit.

The total synaptic input currents, taking into account the inhibition between populations, the self-excitation, the background current and the stimulus-selective current can be written as:

$$I_{L,tot} = J_{L,L}S_L - J_{L,R}S_R + I_{stim,L} + I_{noise,L} \tag{3.3}$$
$$I_{R,tot} = J_{R,R}S_R - J_{R,L}S_L + I_{stim,R} + I_{noise,R} \tag{3.4}$$

with $J_{i,j}$ the synaptic couplings (i and j being L or R). The minus signs in the equations make explicit the fact that the inter-units connections are inhibitory (the synaptic parameters $J_{i,j}$ being thus positive or null). The term $I_{stim,i}$ is the stimulus-selective external input. If μ_0 denotes the strength of the signal, the form of this stimulus-selective current is:

$$\begin{aligned} I_{stim,L} &= J_{A,ext}\mu_0\left(1 \pm c\right) \\ I_{stim,R} &= J_{A,ext}\mu_0\left(1 \mp c\right) \end{aligned} \tag{3.5}$$

The sign, $\pm$, is positive when the stimulus favors population L, negative in the other case. The quantity c, between 0 and 1, gives the strength of the signal bias. It quantifies the coherence level of the stimulus. For example, in the random dot motion framework, it corresponds to the fraction of dots contributing to the coherent motion. This coherence level will be given in percent. Following Wang (2002), this input forms the pooling of the activities of middle temporal neurons firing according to their preferred directions. This input current is only present during the presentation of the stimulus and is shut down once the decision is made. In the present model, in line with a large literature modeling decision

making, the input, Equation (3.5), is thus reduced to a signal parametrized by a scalar quantifying the coherence or degree of ambiguity of the stimulus.

In addition to the stimulus-selective part, each unit receives individually an extra noisy input, fluctuating around the mean effective external input I_0:

$$\tau_{noise}\frac{\mathrm{d}I_{noise,i}}{\mathrm{d}t} = -\left(I_{noise,i}(t) - I_0\right) + \eta_i(t)\sqrt{\tau_{noise}}\sigma_{noise} \tag{3.6}$$

with τ_{noise} a synaptic time constant which filters the (uncorrelated) white-noises, $\eta_i(t), i = L, R$. For the simulations, unless otherwise stated parameters values will be those of Table 3.1.

Parameter	**Value**	**Parameter**	**Value**
a	270 Hz/nA	σ_{noise}	0.02 nA
b	108 Hz	τ_{noise}	2 mS
d	0.154 s	I_0	0.3255 nA
γ	0.641	μ_0	30 Hz
τ_S	100 ms	$J_{A,ext}$	5.2×10^{-4} nA/Hz
$J_{N,LL} = J_{N,RR}$	0.2609 nA	$J_{N,LR} = J_{N,RL}$	0.0497 nA
z	20 Hz		
$I_{CD,max}$	0.035 nA	τ_{CD}	200 ms

Table 3.1: Numerical values of the model parameters: above the dashed line, as taken from Wong and Wang (2006); The last line corresponds to values of the additional parameters specific to the present model (see text).

The system has made a decision when the firing rate of one of the two units crosses a threshold z for the first time, fixed here at 20 Hz.

Corollary discharge

Studies (Roitman and Shadlen, 2002; Ganguli et al., 2008) show that, during decision tasks, neurons activity experiences a rapid decay following the responses - see e.g. Figures 7 and 9 in Roitman and Shadlen (2002). Simulations of the above model show that even when the stimulus is withdrawn at the time of decision, the decrease in activity is not sufficiently strong to account for these empirical findings. Decreasing the recurrent excitatory weights does allow for a stronger decrease in activity, as shown by Bonaiuto et al. (2016). However, both the increase and the decay of activities are too slow, and the network cannot perform sequential decisions with RSIs below 1sec. Hence the decrease in activity requires an inhibitory input at the time of the decision.

Such inhibitory mechanism has been proposed to originate from the superior colliculus (SC), controlling saccadic eye movements, and the basal ganglia-thalamic circuit (BG), which plays a fundamental role in many cognitive functions including perceptual decision-making. Indeed, the burst neurons of the SC receive inputs from the parietal cortex and project to midbrain neurons responsible for the generation of saccadic eye movements (Hall and Moschovakis, 2004; Scudder et al., 2002). Thus the threshold crossing of the cortical neural activity is believed to be detected by the SC (Saito and Isa, 2003). In turn, the SC projects feedback connections on cortical neurons (Crapse and Sommer, 2009). At the time of a saccade, SC neurons emit a corollary discharge (CD) through these feedback connections (Sommer and Wurtz, 2008). The impact of this CD as an inhibition has been

discussed in various contexts (Crapse and Sommer, 2008; Sommer and Wurtz, 2008; Yang et al., 2008). The generation of a corollary discharge resulting in an inhibitory input has been proposed and discussed in several modelling works, in the case of the modulation of the decision threshold in reaction time tasks (Lo and Wang, 2006), in the context of learning (Engel et al., 2015), and in a ring model of visual working memory (Bliss et al., 2017). In order to analyze these effects with the reduced attractor network model, after crossing the threshold, the network receives an inhibitory current, mimicking the joint effect of basal-ganglia and superior colliculus on the two neural populations (Figure 3.2.A).

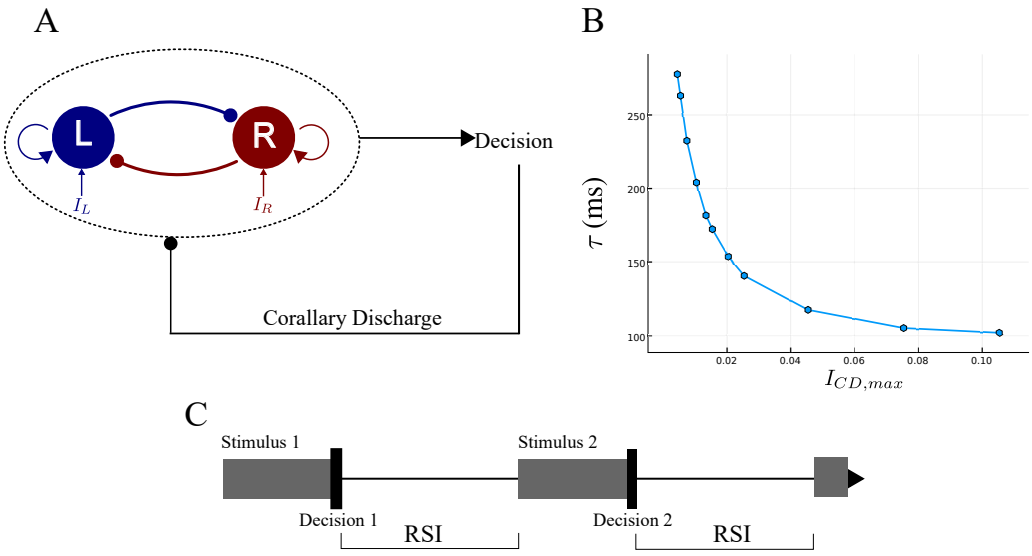


Figure 3.2: **Extended version of the reduced model with the corollary discharge.** (A) The extension consists in adding the corollary discharge originating from the basal ganglia, an inhibitory input onto both units occurring just after a decision is made. (B) Relaxation time constant of the system during the RSI (that is the relaxation dynamics towards the neutral attractor), with respect to the corollary discharge amplitude. The values are obtained by computing the largest eigenvalue λ of the fixed point of the dynamical system, Equation (7.2–3.6), when presenting a constant corollary discharge. The time constant is given by the inverse of the eigenvalue, $\tau = -1/\lambda$. (C) The time-sketch of the simulations can be decomposed into a succession of identical blocks. Each block, corresponding to one trial, consists of: the presentation of a stimulus with a randomly chosen coherence (gray box), a decision immediately followed by the removal of the stimulus, a waiting time of constant duration corresponding to the response-stimulus interval (RSI).

In the case of Engel et al. (2015), the function of the corollary discharge is to reset the neural activity in order to allow the network to learn during the next trial. For this, the form of the CD input is chosen as a constant inhibitory current for a duration of 300ms. However, such strong input leads to an abrupt reset to the neural state with no memory of the previous trial. I thus rather consider here a smooth version of this discharge, considering that the resulting inhibitory input has a standard exponential form (Finkel and Redman, 1983). The inhibitory input, $I_{CD}(t)$, is then given by:

$$I_{CD}(t) = \begin{cases} 0 & \text{during stimulus presentation} \\ -I_{CD,max} \exp\left(-(t-t_D)/\tau_{CD}\right) & \text{after the decision time, } t_D \end{cases} \tag{3.7}$$

The relaxation time constant τ_{CD} is chosen in the biological range of synaptic relaxation times and in accordance with the relaxation-times range of the network dynamics, $\tau_{CD} =$

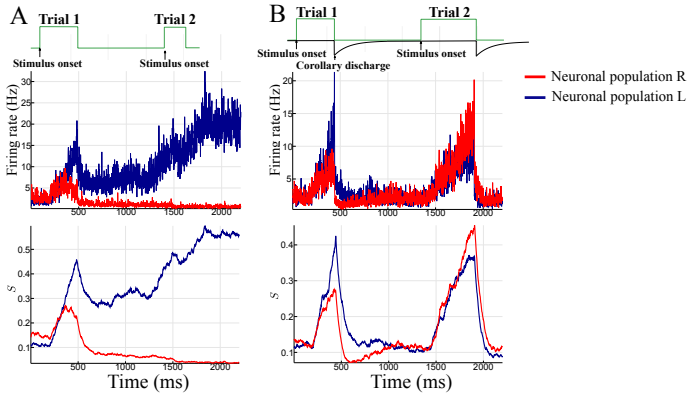


Figure 3.3: **Time course of activities during two consecutive trials.** Left panel, A: Without corollary discharge. A, Top (green plot): Time course of the stimulations. The first stimulus belongs to category L, the second to category R. A, Middle: firing rates of the L (blue) and R (red) neural pools. A, Bottom: corresponding synaptic activities. The neural activity becomes stuck in the attractor corresponding to the first decision. Right panel, B: With corollary discharge, with $I_{CD,max} = 0.035$ nA. B, Top: Time course of the stimulations (green plot, same protocol as for (A)), and time course of the inhibitory current (black curve, represented inverted for clarity of the presentation). B, Middle and Bottom: neural and synaptic activities, respectively (L pool: Blue, R pool: Red). In that case, one observes the decay of activity after a decision has been made, and the winning population is different for the two trials.

200 ms (see Figure 3.2.B).

Therefore the input currents are modified as follows:

$$I_{L,tot}(t) = J_{LL}S_L(t) - J_{L,R}S_R(t) + I_{stim,L}(t) + I_{noise,L}(t) + I_{CD}(t) \quad (3.8)$$

$$I_{R,tot}(t) = J_{RR}S_R(t) - J_{R,L}S_L(t) + I_{stim,R}(t) + I_{noise,R}(t) + I_{CD}(t). \quad (3.9)$$

With this framework it is now possible to study the dynamics of this system in a sequence of decision trials (protocol illustrated in Figure 3.2.C). I will address two issues: first, is there a parameter regime for which the network can engage in a series of trials - that is, for which the state of the dynamical system, at the end of the relaxation period (end of the RSI), is close to the neutral state (instead of being trapped in the attractor reached at the first trial) ? Second, is there a domain within this parameter regime for which one expects to see sequential effects (instead of a complete loss of the memory of the previous decision state)?

Figure 3.3 illustrates the network dynamics between two consecutive stimuli during a sequence of trials, comparing the cases with and without the corollary discharge. In the absence of the CD input, the network is not able to make a new decision different from the previous one (Figure 3.3.A). Even when the opposite stimulus is presented, the system cannot leave the attractor previously reached, unless in the presence of an unrealistic strong input bias. If however the strength $I_{CD,max}$ is strong enough, the corollary discharge makes the system escape from the previous attractor and relax towards near the neutral resting state with low firing rates. If too strong or in case of a too long RSI, at the onset of the next stimulus the neutral state has been reached and memory of past trials is lost. For an intermediate range of parameters, at the onset of the next stimulus the system has escaped from the attractor but is still on a trajectory dependent on the previous trial (Figure 3.3.B).

I have computed the time constant τ of the network during relaxation (during the RSI), with respect to the CD amplitude, $I_{CD,max}$, see Figure 3.2.B. This computation is done for a corollary discharge with a constant amplitude, $I_{CD}(t) = I_{CD,max}$. One sees that, for $I_{CD,max}$ of order $0.03 \sim 0.04$nA, the network time constant τ is four to five times smaller than the duration of the RSI. I choose the relaxation constant τ_{CD} of the corollary discharge of the same order of magnitude (as in the above simulation where $\tau_{CD} = 200$ms). With such value, at the onset of the next stimulus, the network state will still be far enough from the symmetric attractor, so that one can expect to observe sequential effects.

With the inhibitory corollary discharge, after the threshold is crossed by one of the two neural populations, there is a big drop in the neuronal activity (Figure 3.3.B), corresponding to the exit from the previous attractor state. This type of time-course is in agreement with the experimental findings of Roitman and Shadlen (2002); Ganguli et al. (2008), who measured the activity of LIP neurons during a decision task. They showed that neurons that accumulate evidence during decision tasks experience rapid decay, or inhibitory suppression, of activity following responses, similar to Figure 3.3.B (but see Lo and Wang (2006) for a related modeling study with spiking neurons, or Gao et al. (2009) for rapid decay of neural activity with an other type of attractor network).

Now, I derive the conditions on I_{CD} under which the network is able to make a sequence of trials. To this end, I analyze the dynamics after a decision has been made, during the RSI (hence during the period with no external excitatory inputs). The results are illustrated in Figure 3.4 on which I represent a sketch of the phase plane dynamics and a bifurcation diagram.

Consider first what would happen under a scenario of a constant, time independent, inhibitory input during all the RSI (Figure 3.4.A-B-C-D) (formally, this corresponds to setting $\tau_{CD} = +\infty$ in Equation 3.7). At small values of the inhibitory current, the attractor landscape is qualitatively the same as in the absence of inhibitory current: in the absence of noise there are three fixed points, one associated with each one of both categories and the neutral one (Figure 3.3.B). At some critical value, of about 0.0215 nA, there is a bifurcation (Figure 3.4.D): for larger values of the inhibitory current, only one fixed point remains, the neutral one (Figure 3.4.D). As a result, applying a constant CD would either have no effect on the attractor landscape - current amplitude below the critical value - so that the dynamics remains within the basin of attraction of the attractor reached at the previous trial; or it would reset the activity at the neutral state (current amplitude above the critical value), loosing all memory of the previous decision.

Now in the case of a CD with a value decreasing with time (Figure 3.4.E-F-G-H, scenario of an exponential decay), the network behavior will depend on where the dynamics lies at the time of the onset of the next stimulus. The dynamics, starting from a decision state (e.g. near the blue attractor in Figure 3.4.F-G), is more easily understood by considering the limit of slow relaxation (large time constant τ_{CD}). Between times t and $t + \Delta t$, with Δt small compared to τ, the dynamics is similar to what it would be with a constant CD with amplitude $I_{CD}(t)$. Hence if $I_{CD}(t)$ is larger than the critical value discussed above, the dynamics 'sees' a unique attractor, the neutral state, and is driven towards it. When $I_{CD}(t)$ becomes smaller than the critical value, the system 'sees' again three attractors, and finds itself within the basin of attraction of either the initial fixed point (corresponding to the previous decision, Figure 3.4.F), or of the neutral fixed point (Figure 3.4.G). In the latter case, the network is able to engage in a new decision task.

In order to have the network performing sequential decision tasks, one needs $I_{CD,max}$ to be larger than the critical value (about $I_{CD} = 0.0215$ nA, Figure 3.4.H), and, for a given value of $I_{CD,max}$, to have a time constant τ_{CD} large enough compared to the RSI for the

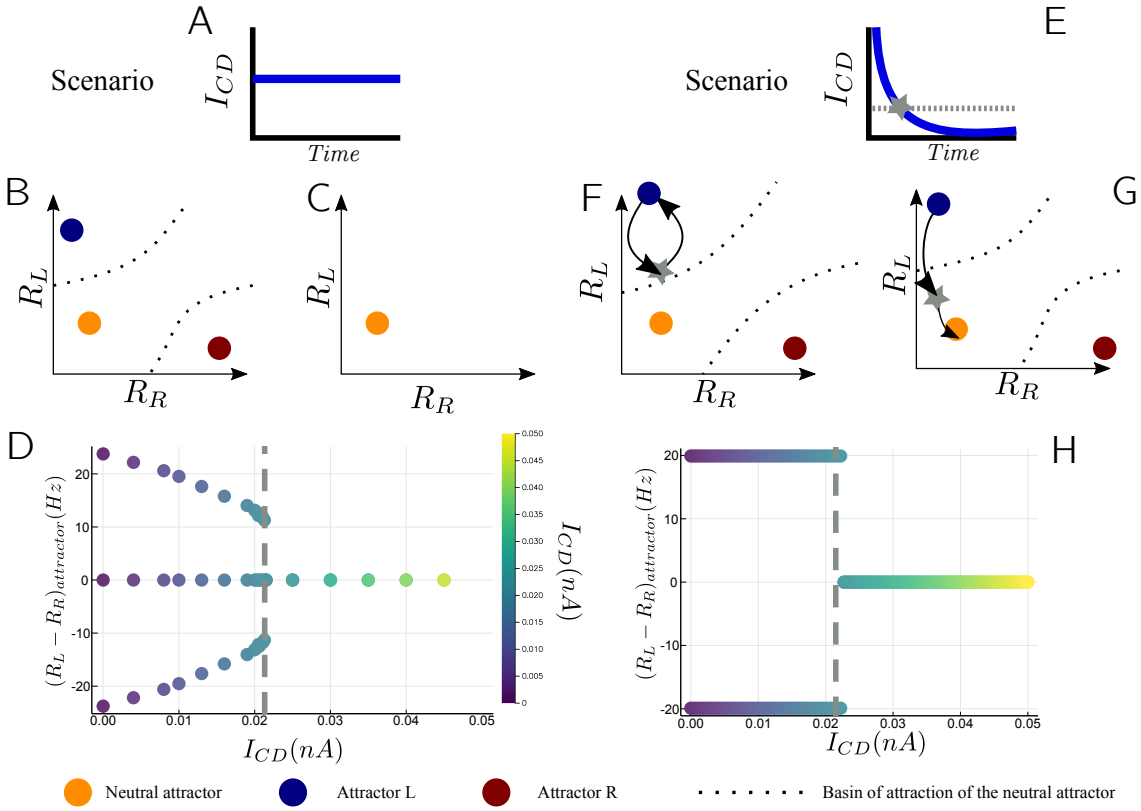


Figure 3.4: **Bifurcation diagram of sequential decision making, for two scenario of I_{CD}.** (A) Scenario with a constant value of the inhibitory current for the left part of the figure, panels B, C and D. (B) Phase plane representation of the attractors at low I_{CD} (below the critical value). (C) Phase plane representation of the attractor landscape at high I_{CD} (above the critical value). Only the neutral attractor exists, corresponding to the right side of panel D. (D) Attractors state (as the difference in firing rates, $R_L - R_R$) with respect to I_{CD}. The gray line, at $I_{CD} = 0.0215$ nA, represents the bifurcation point. On the left side three attractors exist, on the right side only the neutral one exists. The case without inhibitory current corresponds to $I_{CD} = 0$ nA. (E) Scenario with an inhibitory current decreasing exponentially with time, for the right part of the figure, panels F, G and H. The dashed line corresponds to $I_{CD} = 0.0215$ nA, value at which the bifurcation at constant I_{CD} occurs (see panel D). The time at which the current amplitude crosses this value is denoted by the gray star in panels E and F. (F) Schematic phase-plan dynamics corresponding to the left side of (H). The blue attractor corresponds to the starting point and the black arrow represents the dynamics. At the time I_{CD} becomes lower than 0.0215 nA (gray star), the system is still within the basin of attraction of the initial attractor. Hence, it goes back to the initial attractor. (G) Schematic phase-plan dynamics corresponding to the right side of panel H. At the time I_{CD} becomes lower than 0.0215 nA, the system lies within the basin of attraction of the neutral attractor. Hence, the dynamics continues towards the neutral attractor. Those conditions are the ones needed for sequential decision-making. (H) Attractors that can be reached when starting from a decision state, for the relaxation dynamics under the scenario represented on panel E. On the left side of the dashed gray line, the value of $I_{CD,max}$ is too weak and the network remains locked to the attractor corresponding to the previous decision state. On the right side the network dynamics lies within the basin of attraction of the neutral attractor, allowing the network to engage in a new decision task.

system to relax close enough to the neutral attractor at the onset of the next stimulus. However, sequential effects may exist only if the current decreases sufficiently rapidly, so that the trajectory is still significantly dependent on the state at the previous decision. This justifies the choice of exponential decrease of the inhibitory current, Equation 3.7, and the numerical value of $\tau_{CD} = 200$ms. Moreover, recording from relay neurons, Sommer and Wurtz (2002) show that the signal corresponding to the corollary discharge lasts several hundred of milliseconds. This time scale falls precisely in the range of values of the relaxation time constant of the model (see Figure 3.2.B), and corresponds to values for which the model shows sequential effects.

Sequential effects

The dynamical properties described above give that, for the appropriate parameter regime, the RSI relaxation leads to a state which is between the previous decision state and the neutral attractor. If it is still within the basin of attraction of the previous decision state at the onset of the next stimulus, one expects sequential biases. This mechanism is similar to the one discussed by Bonaiuto et al. (2016). However, the relaxation mechanisms are different. This results in different quantitative properties, notably and quite importantly in the time scale of the relaxation, which is here more in agreement with experimental findings (Cho et al., 2002). After running simulations of the network dynamics with the protocol of Figure 3.2.C, I analyze the effects of response repetition by separating the trials into two groups, the *Repeated* and *Alternated* cases. The repeated (respectively alternated) trials are those for which the decision is identical to (resp. different from) the decision at the previous trial. I do not consider whether the *stimulus* category is repeated or alternated: the analysis is based on whether the *decision* is identical or different between two consecutive trials (Fleming et al., 2010; Padoa-Schioppa, 2013). Such analysis is appropriate, since the effects under consideration depend on the levels of activity specific to the previous decision. I run a simulation of 1000 consecutive trials, each of them with a coherence value randomly chosen between 20 values in the range $[-0.512, 0.512]$. I do so for two values of the corollary discharge amplitude, $I_{CD,max} = 0.035$ nA and $I_{CD,max} = 0.08$ nA, with a RSI of 1 s, the other parameters being given on Table 3.1.

The distributions of coherence values are identical for the two groups, for both values of $I_{CD,max}$ (Anderson-Darling test, $p = 0.75$ and $p = 0.84$ respectively). I study the reaction times separately for the two groups, and present the results in Figure 3.5. Figure 3.5.C represents the so called energy distance (Székely and Rizzo, 2013; Rizzo and Székely, 2016) between the repeated and alternated reaction times distribution. As one can observe, the distance decreases, hence the sequential effect diminishes, as the corollary discharge amplitude $I_{CD,max}$ increases. For the specific case of Figures 3.5.A and B, the corresponding E-statistic for testing equal distributions leads to the conclusion that in the case $I_{CD,max} = 0.035$ nA, the two reaction-time distributions are different ($p = 0.0019$). This implies that the behavior of the network is influenced by the previous trial. There is a faster mean reaction time (around 55 ms) when the choice is repeated (Figure 3.5A), with identical shape of the reaction times distributions. The difference in means is of the same order as found by Cho et al. (2002) in experiments on 2AFC tasks. On the contrary, for $I_{CD,max} = 0.08$ nA (Figure 3.5.B), the two histograms cannot be distinguished (E-statistic test, $p = 0.25$).

I have checked that increasing the RSI has a similar effect to increasing the corollary discharge amplitude. Sequential effects can be observed for RSI values in the range 0.5 to 5 seconds, in accordance with two-choices decision-making experiments, where such effects are observed for RSI less than 5 seconds (Rabbitt and Rodgers, 1977; Laming, 1979; Soetens et al., 1985).

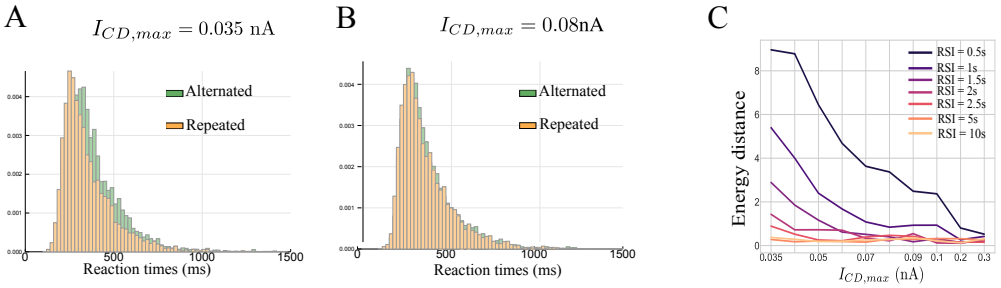


Figure 3.5: **Histogram of the reaction times.** Simulations run at, (A): $I_{CD,max} = 0.035$ nA, and (B): $I_{CD,max} = 0.08$ nA, with a RSI of 1 second. The green histogram corresponds to the Alternated case, that is when the decisions made at the nth and nth $+1$ trials are different. The orange histogram corresponds to the Repeated case, that is when the decisions made at the nth and nth $+1$ trials are identical. (C): Energy distance between the repeated and alternated histograms. The x-axis represents the strength of the corollary discharge, and the color codes the duration of the RSI in seconds.

Neural correlates: Dynamics analysis

With the relaxation of the activities induced by the corollary discharge, the state of the network at the onset of the next stimulus lies in-between the attractor state corresponding to the previous decision, and the neutral attractor state. When averaging separately over repeated and alternated trials, I find that this relaxation dynamics has different behaviors depending on whether the next decision is identical or different from the previous one. This is a statistical effect which can only be seen by averaging over a very large number of trials.

In Figure 3.6, I compare two examples of network activity, one with an alternated choice, and one with a repeated choice, by plotting the dynamics during two consecutive trials. Figure 3.6.A represents the alternated case, and shows that, previous to the onset of the second stimulus (light blue rectangle), the activities of the two populations are at very similar levels. In contrast, for the case of a repeated choice, Figure 3.6.C, the activities are well separated, with higher firing rates.

Figure 3.6.B shows a classical phase-plane representation of the network dynamics during two consecutive trials, with the axes as the synaptic activities of the wining versus loosing neuronal populations in the first trial. One sees a trajectory starting from the neutral state, going to the vicinity of the attractor corresponding to the first decision, and then relaxing to the vicinity of the neutral state (as illustrated in Figure 3.4.G). Then the trajectory goes towards the attractor corresponding to the next decision, different from the first one. This aspect of the dynamics is similar to what is obtained in Gao et al. (2009) with another type of attractor network. In Figure 3.6.D, I show the phase-plane dynamics in the case of a repeated choice (trajectory in blue). On this same panel, for comparison I reproduce in light red the dynamics, shown in Figure 3.6.B, during the first trial in the alternated case. As can be seen in Figure 3.6.D, the network states at the time of decision are different depending on whether the network makes a decision identical to or different from the one made at the previous trial.

In order to check the statistical significance of these observations, I represent in Figure 3.7 the mean activities during the RSI, obtained by averaging the dynamics over all trials, separately for the alternated and repeated groups. As expected, for small values of $I_{CD,max}$ (0.035 nA), the two dynamics are clearly different. This difference diminishes during relaxation. However at the onset of the next stimulus one can still observe some residues,

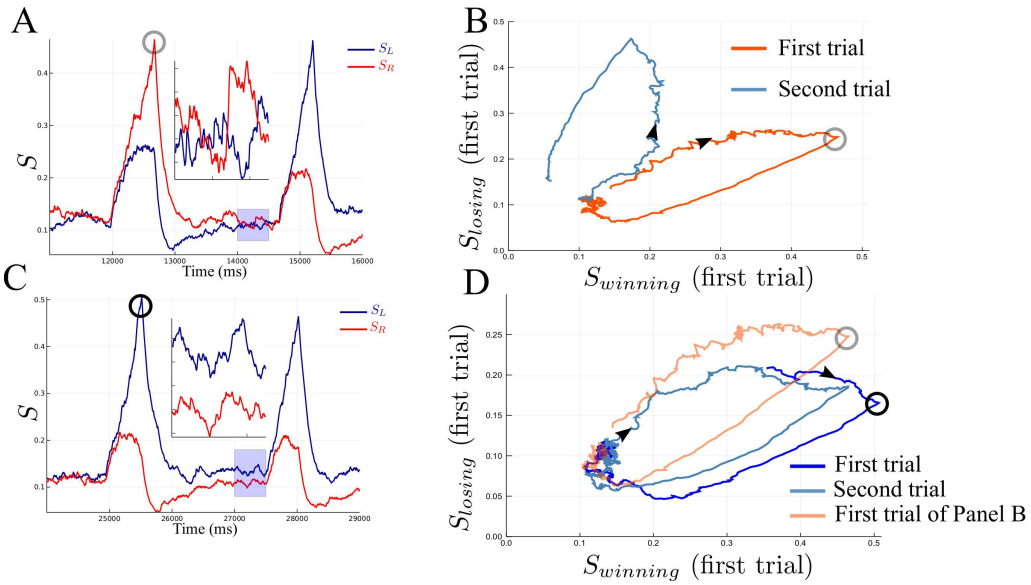


Figure 3.6: **Network activity during two consecutive trials.** Panels (A) and (B) represent the alternated case where the decision made is R then L, and panels (C) and (D) represent the case where decision L is made and repeated. Panels (A) and (C) plot the time course activities of the network. The light blue zone is zoomed in order to better see the dynamics just before the onset of the second stimulus. The red and blue curves correspond to the activities of, respectively, the R and L network units. Panels (B) and (D) represent, respectively, the (A) and (C) dynamics in the phase-plane coordinates. On panel (B) the dynamics evolves from dark red (first trial) to light blue (2nd trial), and on panel (D) from dark blue (first trial) toward light blue (2nd trial). The gray – respectively black – circles identify the same specific point during the dynamics in panels (A) and (B) – resp. (C) and (D). The circles are not at the exact same value because the decision threshold is on the firing rates and not for the synaptic activities. In order to compare the alternated and repeated cases, (A,B) and (C, D), the dark red curve of panel (B), is reproduced on panel (D) as light orange curve.

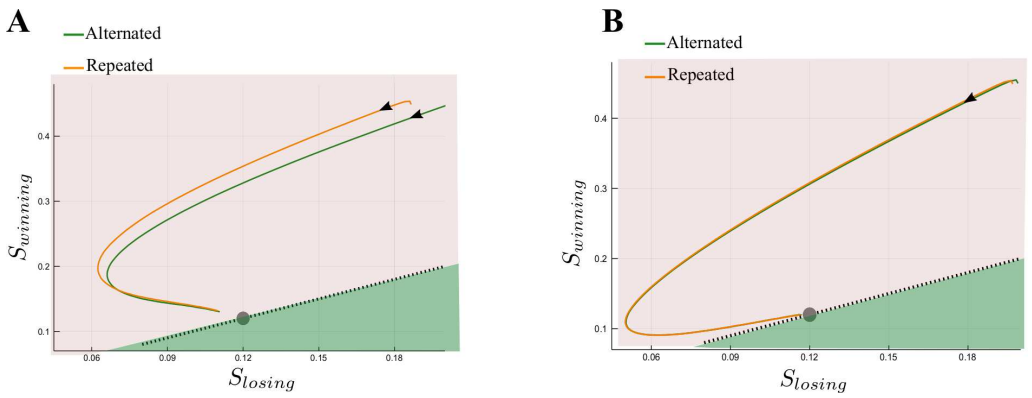


Figure 3.7: **Phase plane dynamics.** Dynamics of the decaying activity between two successive trials, (A) for $I_{CD,max} = 0.035$ nA, and (B) for $I_{CD,max} = 0.08$ nA. The synaptic activity is averaged over all trials separately for each one of the two groups: alternated (green) and repeated (orange). The axes are $S_{winning}$ and S_{losing} (not S_R and S_L) corresponding to the mean synaptic activity of, respectively, the winning and the losing populations for this trial. Note the difference in scale of the two axes. The time evolution along each curve follows the black arrow. The dashed black line denotes the symmetric states ($S_L = S_R$) of the network, and the gray circle the neutral attractor. The shadow areas represent the basins of attraction (at low coherence levels) for the repeated and alternated trials, respectively pink and green.

statistically significant according to an Anderson-Darling test done on the 500 ms prior to the next stimulus (between winning population, $p = 0.0034$, between losing population $p = 3.2 \times 10^{-8}$).

Looking at Figure 3.7.A, one can note that the ending points of the alternated and repeated relaxations are biased with respect to the symmetric state. At the beginning of the next stimulus the network is already in the basin of attraction of the repeated case. Hence, it will be harder to reach the alternated attractor stated (in the green region). When increasing $I_{CD,max}$ (Figure 3.7.B), the ending state of the relaxation lies closer to the attractor state. Hence, the biases in sequential effects disappear because at the beginning of the next stimuli the network starts from the symmetric (neutral) state. The same analysis holds for longer RSI, the dynamics are almost identical (Anderson-Darling test: between winning population, $p = 0.25$, between losing population $p = 0.4$), and both relaxations end near the neutral attractor state. The bias depending on the next stimulus is not observed anymore, and the sequential effect on reaction time hence disappears.

I would like to precise that the sequential effects only depend on whether or not the states at the end of the relaxation lie on the basin boundary. However, the effects can also be observed at the level of the relaxation dynamics, since the trajectories for alternated and repeated cases are identical when there is no effect, and different in the case of sequential effects.

The analysis of the dynamics also leads to expectations for what concerns the bias in accuracy towards the previous decision. Indeed, this can be deduced from Figure 3.7. If the choice at the previous trial was R (respectively L), then, at the end of the relaxation, the network lies closer to the basin of attraction of attractor R (respectively L). Thus when presenting the next stimulus, the decision will be biased towards the previous state, so that the probability of making the same choice will be greater than the one of making the opposite choice. Otherwise stated, given the stimulus presented at the current trial,

the probability to make the choice R will be greater when the previous choice was also R, than when the previous choice was L. Numerical simulations confirm this analysis, as illustrated in Figure 3.8. The RSI dependency is statistically significant (Generalized Linear Model: $r = -3.9$, $p < 0.0001$). For small RSI (500 ms), the decision is biased towards the previous one, and for RSI of several seconds this effect disappears. These results are in agreement with experimental findings of Bonaiuto et al. (2016). The authors studied response repetition biases in humans with RSIs of at least 1.5 seconds. In these experiments, they measured the Left-Right indecision point, that is the level of coherence resulting in chance selection. Compared to the repeated case, they found out that the indecision point for the alternated case is at a higher coherence level, and this shift decreases as the RSI increases.

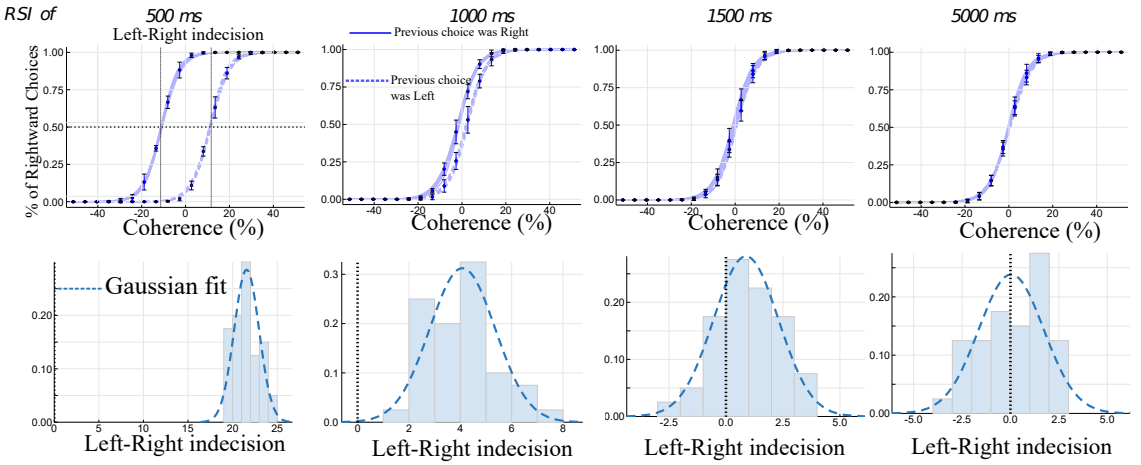


Figure 3.8: **Repetition biases for several RSI values.** Upper panel: percentage of Right choices, with respect to coherence level, depending on the previous choice (Left or Right). The blue points represent the mean accuracy (on 24 simulations) and the confidence interval at 95% (bootstrap method). The blue lines denote the fit (of all simulations) by a logistic regression of all (plain: previous choice was Right, dashed: previous choice was Left). Bottom panel: histogram of the Left-Right indecision point (on 24 simulations to stay in the experimental range). It characterizes the fact that the positive shift in the indecision point is increased for small RSI. The mean of the indecision point shift decreases with longer response-stimulus intervals.

To conclude this section, at the time of decision, the winning population has a firing rate higher than the losing population. After relaxation, at the onset of the next stimulus, the two neural pools have more similar activities, but are still sufficiently different, that is the dynamics is still significantly away from the neutral attractor. At the onset of the next stimulus, the systems finds itself in the basin of attraction of the attractor associated to the same decision as the previous one. This results in a dynamical bias in favor of the previous decision. The probability to make the same choice as the previous one is then larger than the one of the other choice, and the reaction time, for making the same choice (repeated case), is shorter than for making the opposite choice (alternated case). In accordance with these results, studies on the LIP, superior colliculus and basal ganglia have found that the baseline activities before the onset of the stimuli can reflect the probabilities of making the saccade, under specific conditions (Lauwereyns et al., 2002; Ding and Hikosaka, 2006; Rao et al., 2012). The model shows that these modulations of the baseline activities can be understood as resulting from the across-trial dynamics of the decision process.

3.2. Post-error effects

3.2.1 Experimental evidences

During sequential decision-making people show post-error adaptations. Different types of behavioral post-error adjustments have been observed such as post-error slowing (PES) and post-error improvement in accuracy (PIA). The most studied effect consists in PES (Laming, 1979), and see Danielmeier and Ullsperger (2011) for a review. It consists of prolonged reaction times in trials following an error, compared to reaction times following a correct trial. This effect has been observed in a variety of tasks: categorization (Jentzsch and Dudschig, 2009), flanker (Debener et al., 2005), Stroop (Gehring and Fencsik, 2001) tasks. Jentzsch and Dudschig (2009) and Danielmeier and Ullsperger (2011) found that the PES effect depends on the response-stimulus interval. The amplitude of this effect, defined as the difference between the mean reaction times of post-error and post-correct trials, decreases as one increases the RSI, with values going from several dozens of milliseconds to zero. For RSI longer than $750-1500$ ms, PES is not observed anymore. Moreover, this effect is automatic and involuntary (Rabbitt and Rodgers, 1977), and is independent of error detection and correction process which involve other cortical areas (Rodriguez-Fornells et al., 2002). This suggests a rather low level processing origin in line with the present model.

Figure 3.9 presents different behavioral results from Danielmeier and Ullsperger (2011) in a 2AFC task. One can see that different participants show different behaviors of post-error adjustments. Some of them show PIA and PES, PIA and post-error quickening (PEQ) or just PIA. Moreover, the PES effect is strongly impacted by the RSI as when the interval between two trials is of the order of the second, there is no effect anymore (Figure 3.9.B).

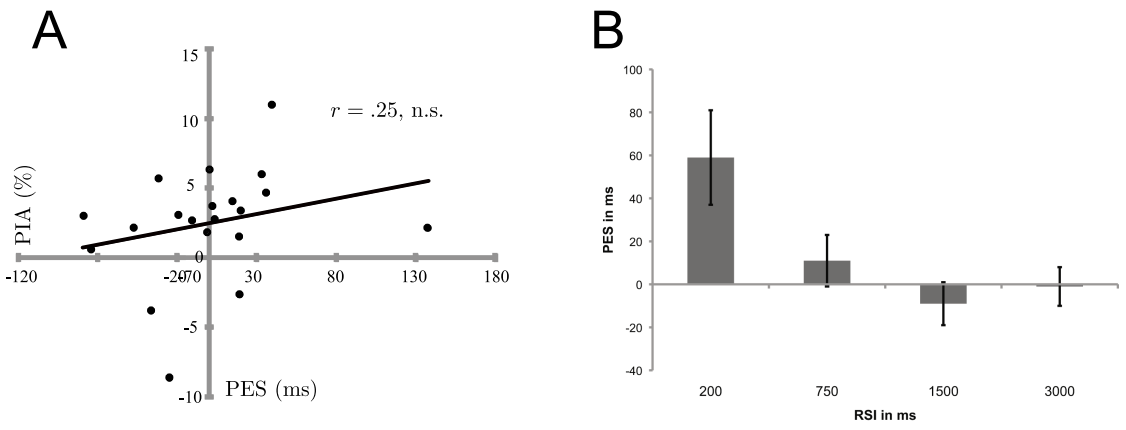


Figure 3.9: **Post-error adjustments (adapted from Danielmeier and Ullsperger (2011)).** The results are from the experiment of Danielmeier et al. (2011). (A) The participants showed PES and PIA, or PIA without PES. Some participants showed the opposite effect of PES, post-error quickening (faster answers after an error). (B) Mean post-error slowing with respect to the RSI in the experiment of Danielmeier et al. (2009). When RSI increases, post-error slowing becomes smaller. For more details about the results, see Danielmeier and Ullsperger (2011).

Remarkably, the PES effect is reported in cases where the subject does not receive information on the correctness of the decision (Jentzsch and Dudschig, 2009; Danielmeier et al., 2011; Danielmeier and Ullsperger, 2011). The classical framework used to analyze the post-error effects is the one of DDMs (Dutilh et al., 2012). These studies have shown that

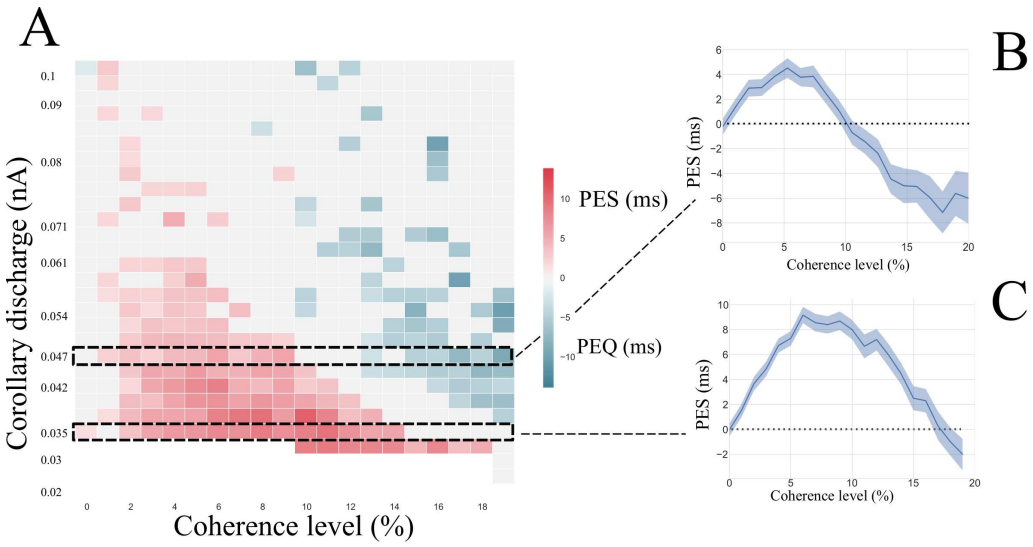


Figure 3.10: **Post-error slowing in the simulated network at a RSI of 500 ms.** (A) Phase diagram of the PES effect at RSI of 500 ms. The bottom white zone corresponds to parameters where sequential decision-making is impossible as the network is unable to leave the attractor state during the RSI. The red square corresponds to regions where PES is observed, and the blue ones where PEQ is observed (the darker the color, the stronger the effect). The black dashed squares correspond to specific regions where Panels B and C zoom. (B) PES effect (ms) with respect to the coherence level at $I_{CD,max} = 0.047$ nA. The light blue zone corresponds to the bootstrapped (Efron and Tibshirani, 1994) confidence interval at 95%. (C) PES effect (ms) with respect to the coherence level at $I_{CD,max} = 0.035$ nA. The light blue zone corresponds to the bootstrapped confidence interval at 95%.

post-error and post-correct trials can be fitted by DDMs with different set of parameters values for post-error and post-correct trials. Dutilh et al. (2012) argue that the modification of the decision threshold within the DDM framework would correspond to the hypothesis of increased response caution, the decision becoming more cautious after an error. Yet, the neural correlates, which would determine the threshold or the starting point, remain obscure, especially in the absence of feedback on the correctness of the trial.

3.2.2 Post-error adjustments in attractor neural networks

I will present an analysis of the post-error adjustments in the attractor neural network that has been modified to perform sequences of decisions. In this framework, the network does not receive any feedback about the correctness of its decision. Therefore it is well suited to study post-error adjustments in the case where no feedback to the participants is given (Dutilh et al., 2012).

Post-error slowing

I studied the occurrence of the PES effect in the model with respect to the coherence level and $I_{CD,max}$, at an intermediate RSI value of 500 ms, leading to the phase diagram in Figure 3.10.A. I find that a large domain in parameter space shows PES effect (in red in the figure). Figure 3.10.B zooms on a value of $I_{CD,max}$ for which PES occurs ($I_{CD,max} = 0.035$ nA). The magnitude of the PES effect goes from zero to ten milliseconds at $c = 10\%$, hence remaining within the range of behavioral data (Jentzsch and Dudschig, 2009; Danielmeier and Ullsperger, 2011) ($10 - 15$ ms for a RSI of $0.5 - 1$ second). In these experiments (a flanker task with stimuli belonging to one of two opposite categories, Left

or Right directions), the ambiguity level is not quantified. However, the observed error rates are found around 10% which, within our model, corresponds to a coherence level of about $c = 10\%$. On the phase diagram, one can observe the variation of the PES effect with respect to the coherence level. In the region where I observe a PES effect, I find that it is enhanced under conditions when errors are infrequent. However, for large values of the coherence level, this effect cannot be observed anymore due to the absence of any error in the successive trials (almost 100% of correct trials). This occurrence of PES, principally at low error rates, has been found in experiments of Notebaert et al. (2009); núñez Castellar et al. (2010), for which the authors observe PES when errors are infrequent, but not when errors are frequent. These experiments are with 4-AFC tasks, but the same type of properties is expected for 2AFC tasks – and the model could easily be adapted to such cases with a neural pool specific to each one of the four categories.

The phase diagram, Figure 3.10.A, also shows parameter values with no effect at all (in gray), and a domain with the opposite effect, that is with reaction times faster after an error than after a correct trial (in blue). I propose to call this effect *post-error quickening* (PEQ), as opposed to post-error slowing. As shown in Figure 3.10.C, for a given value of $I_{CD,max}$, one can have PES at low coherence level, and PEQ at high coherence level.

This PEQ effect, although much less studied, has been observed in various AFC experiments, either without feedback (Rabbitt and Rodgers, 1977; Notebaert et al., 2009; King et al., 2010) or with feedback (Purcell and Kiani, 2016), notably for fast-response regimes (Notebaert et al., 2009; King et al., 2010). The conditions for observing PEQ remain however not well established, with some contradictory results. With Go/no-go protocols (which are similar to AFC protocols in many respects), Hester et al. (2005) reported post-error decrease in reaction times for aware errors, but not for unaware errors, whereas Cohen et al. (2009) on the contrary reported no PEQ effect, but a larger PES effect for aware errors than for unaware errors. The fact that the model predicts PEQ in TAFC tasks at high coherence levels is more in line with the results of the fMRI experiments of Hester et al. (2005). Indeed, at high coherence levels, responses are fast and most often correct. In the rare case of an error, the subject is likely to become aware that an error has been made (Yeung and Summerfield, 2012). This thus may lead to a correlation (without causal links) between aware errors and PEQ.

I also studied the RSI dependency of the PES effects by plotting the phase diagram at $I_{CD,max} = 0.045$ nA with respect to the RSI (Figure 3.11). In behavioral experiments the PES effect depends strongly on the RSI. For RSI longer than $1000 - 1500$ ms the observation or not of PES depends specifically on the decision task (Jentzsch and Dudschig, 2009; King et al., 2010). However, a common observation is that, whenever PES is observed, if one keeps increasing the RSI, the PES effect eventually disappears. In Figure 3.11, for parameters where PES is observed at a RSI of 500 ms, increasing the RSI leads to the weakening of the post-error slowing effect until its disappearance. At a RSI of $1000 - 1500$ ms this effect is not present anymore, in agreement with experimental results (Jentzsch and Dudschig, 2009).

The variation of PEQ with respect to RSI has not been experimentally studied, as this effect is more controversial. However, the model shows that the dependence on RSI is similar to the one of PES, and predicts that, when both effects exist at a same RSI value (for different coherence levels), increasing the RSI leads to the disappearance of both of them.

The set of phase diagrams that I present in this work on the various effects for post-error trials, Figures 3.10 to 3.13, provides testable behavioral predictions. As just discussed in the particular case of PES and PEQ , they predict how the effects on reaction times are or

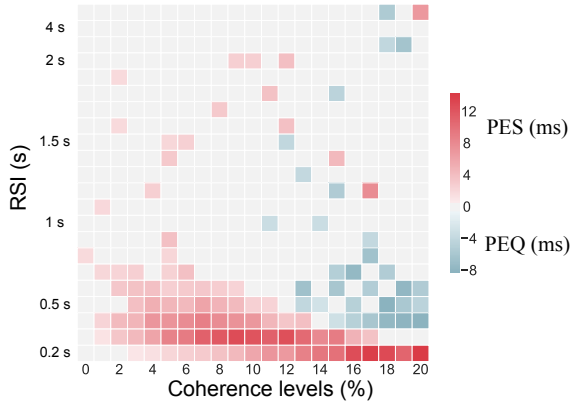


Figure 3.11: **Post-error slowing depending on RSI.** Phase diagram of the PES effect at $I_{CD,max} = 0.045$ nA. The red squares correspond to regions where PES is observed, and the blue ones where PEQ is observed (the darker the color, the stronger the effect). I used a bootstrapped confidence interval in order to decide whether or not PES/PEQ is observed.

are not correlated, and in particular how they qualitatively depend on and co-vary with the coherence level or the duration of the RSI.

Post-error improvement in accuracy

Post-error improvement in accuracy (PIA) is another sequential effect reported in experiments (Laming, 1979; Marco-Pallarés et al., 2008; Danielmeier and Ullsperger, 2011). PIA has been observed on different time-scales: long-term learning effects following error (Hester et al., 2005) and trial-to-trial adjustments directly after commission of error responses. I will only consider this latter type of PIA. The specific conditions under which PIA can be observed in behavioral experiments have not been totally understood. I investigate this effect in the specific context of the attractor neural network, considering that the strength of the effect is linked to the difference in error rates between post-error and post-correct trials.

Figure 3.12 represents the phase diagram of the PIA effect with respect to coherence levels (x-axis) and corollary discharge amplitude (y-axis). One can observe a large region of parameters for which PIA is present. I find a magnitude of the PIA effect of about $2-4\%$, which is of the same order of magnitude as in the experiments where, for RSIs in the range $500-1000$ ms, it is found that post-error accuracy is improved by approximatively 3% (Jentzsch and Dudschig, 2009).

Looking at Figure 3.12, one sees that the PIA and PES effects append in the same region of parameters. However, zooming in on specific regions (Figure 3.12.B and C), one can notice some differences in the variation of these effects. The black dashed rectangular regions correspond to the same parameters as in Figure 3.10. PIA is also observed in these regions. However, there was a decrease of PES at very large coherence (Figure 3.10.B), but not of PIA (Figure 3.12.B). Moreover the decrease of the PIA effect in Figure 3.10.C does not occur at the same values of parameters as for the PES one. It would be tempting to interpret PIA as a better accuracy resulting from taking more time for making the decision. This is not the case, since PIA does not appear uniquely in the PES region, but in the PEQ one too. In agreement with these model predictions, Danielmeier et al. (2011), in a TAFC task with color-based categories, observeed that PIA can occur in the absence of

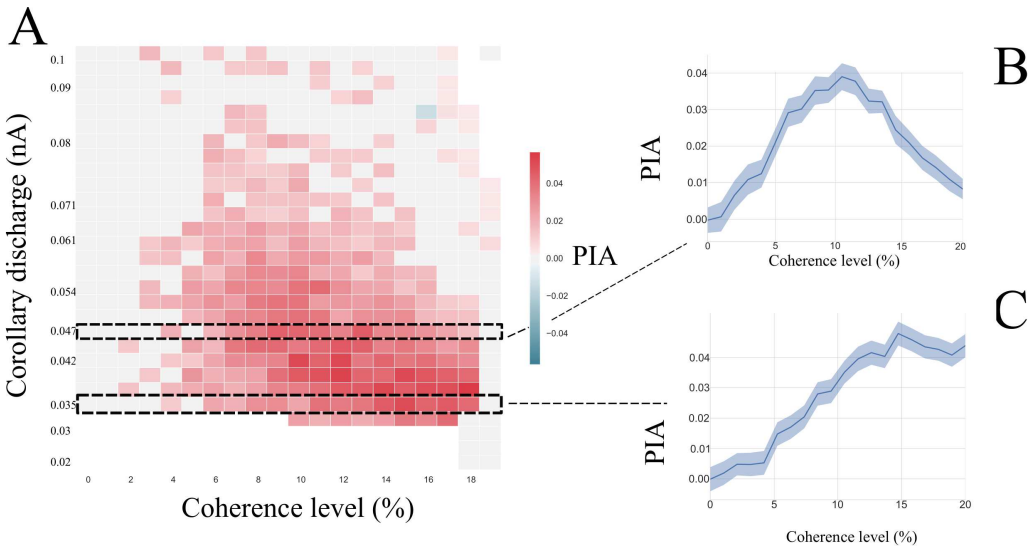


Figure 3.12: **Post-error improvement in accuracy at a RSI of 500 ms.** (A) Phase diagram of the PIA effect at RSI of 500 ms. The bottom white zone corresponds to parameters where sequential decision-making is impossible. The red squares correspond to regions where PIA is observed, whereas the black dashed squares correspond to specific regions where panels B and C zoom. (B) PIA effect with respect to the coherence level at $I_{CD,max} = 0.047$ nA. The light blue zone corresponds to the bootstrapped confidence interval at 95%. (C) PIA effect with respect to the coherence level at $I_{CD,max} = 0.035$ nA. The light blue zone corresponds to the bootstrapped confidence interval at 95%.

PES, but that the occurrence of PES is always associated with PIA (except for one subject among 20, results reported in Figure 3.9 (Danielmeier and Ullsperger, 2011)).

In EEG experiments, Marco-Pallarés et al. (2008) found that time courses of PES and PIA seem to be dissociable as they observed post-error improvements in accuracy with longer inter-trial intervals (up to 2250 ms) than post-error slowing. The authors considered protocols with and without stop-signals. I investigate the variation with respect to the RSI of PIA in our model (Figure 3.13). For long RSIs, the PIA effect is not observed anymore. However as observed in Marco-Pallarés et al. (2008), the PIA effect occurs for longer RSIs than the PES effect (Figure 3.12.A). In the same way, PIA is more robust with respect to the intensity of the corollary discharge. This is corroborated by Figure 3.14-A-B, which represents PES and PIA effect for a larger relaxation time, $\tau_{CD} = 500$ ms, hence with a stronger corollary discharge. However, all the regimes previously observed are present, for slightly different parameter ranges. This shows that the global picture illustrated by the phase diagrams, Figures 3.10, 3.12, is not specific to a narrow range of $I_{CD,max}$ and τ_{CD} values.

Verguts et al. (2011) found that PIA and PES seem to happen independently, suggesting that at least two post-error processes take place in parallel. An important outcome of my analysis is to show that PIA and PES effects can both result from the same underlying dynamics. In addition, in the parameters domain where they both occur, the variations of these effects with respect to the coherence levels are indeed uncorrelated (Pearson correlation test: RSI of 500 ms and $I_{CD} = 0.035$ nA, $p = 0.58$, $I_{CD} = 0.05$, $p = 0.79$ and $I_{CD} = 0.1$ nA, $p = 0.25$; RSI of 2000 ms and $I_{CD} = 0.035$ nA, $p = 0.37$). This non-correlation highlights the complexity of such post-error adjustments, as explored in Verguts et al. (2011).

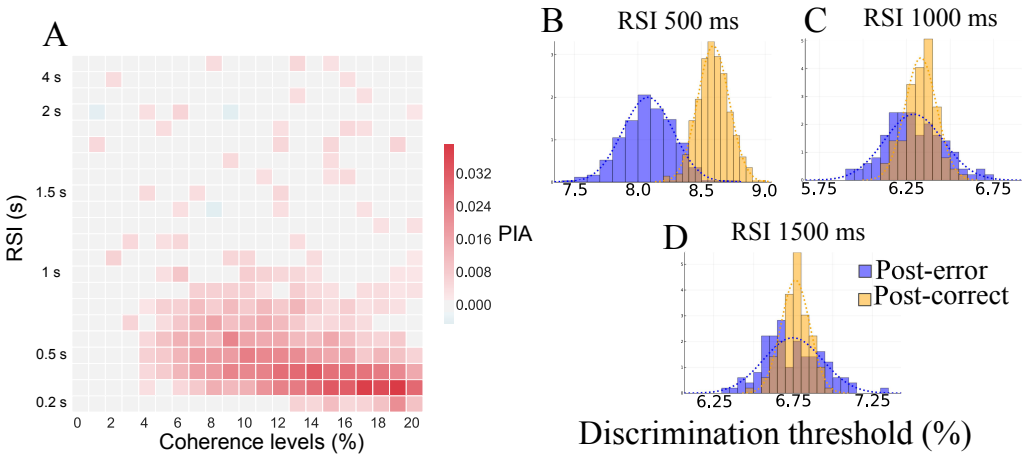


Figure 3.13: **Post-error improvement in accuracy depending on RSI.** (A) Phase diagram of the PIA effect at $I_{CD,max} = 0.045$ nA. The red squares correspond to regions where PIA is observed. (B)-(C)-(D) Distribution of the discrimination threshold for three values of RSI (500, 1000, 1500 ms respectively). In yellow I represent the histogram of the post-correct trials, and in blue the post-error ones. The dashed curves of the corresponding color correspond to the cumulative functions of these distributions. The corollary discharge is $I_{CD,max} = 0.035$ nA.

In order to gain more insights into the PIA effect, I study the discrimination threshold following an error or a success, with respect to the RSI (Figure 3.13.B-D). Figure 3.13.B represents the distribution of the discrimination threshold for $I_{CD,max} = 0.035$ nA and a RSI of 500 ms. For these parameters, the distributions for the post-error and post-success cases are highly different (Smirnov-Kolmogorov test: $p < 10^{-20}$). If one increases the RSI (1000 ms for Figure 3.13.C and 1500 ms for Figure 3.13.D), this difference disappears (Smirnov-Kolmogorov test: $p = 0.038$ and $p = 0.4$ respectively). However, the model predicts a wider distribution of the discrimination threshold after an error than after a correct trial, independently of the presence of the PIA effect. This might result from the wider distribution in the neural (or synaptic) activities after an error, which I discuss in the next section. To my knowledge, this effect has not been studied in behavioral experiments.

3.2.3 Non-linear dynamics

In this section I analyze the PES and PEQ effects in terms of neural dynamics. First of all, I will represent and discuss the dynamics on individual trials for the three regions of parameters: with neither PES nor PEQ effects, with PES effect, and with PEQ effect (Figure 3.15). One can observe the dynamics for post-error and post-correct trials during the relaxation period following a decision and during the presentation of the next stimulus. One can already notice differences between the regions on individual trials. Figure 3.15.A represents a trial in the region without PES or PEQ. The post-error/correct dynamics are indistinguishable. Hence one does not observe any differences in the reaction times. Looking at a trial in the PEQ region (Figure 3.15.B), the population L (the winning one for the second stimulus) for the post-error case seems a bit higher in activity than for the post-correct case. This leads to the post-error quickening effect, as the post-error (orange) curve will reach the threshold sooner than the post-correct (blue) one. Finally, Figure 3.15.C represents individual trials for parameters in the PES region. In the phase diagram (Figure 3.10) the effect was more pronounced than PEQ, thus it is more pronounced on the dynamics too. During the relaxation, and the presentation of the next stimulus,

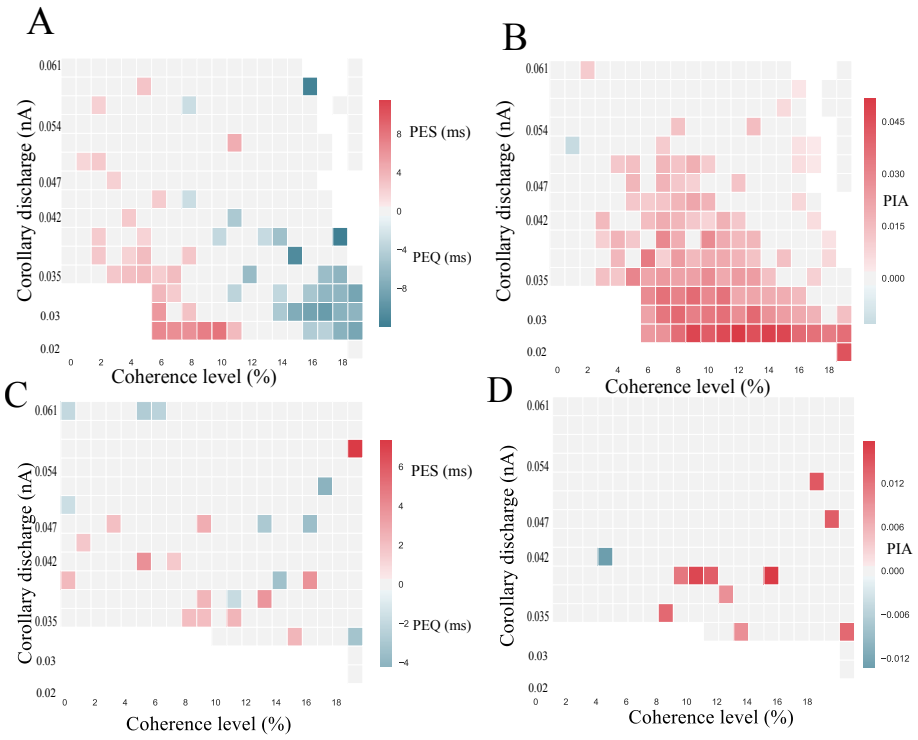


Figure 3.14: **Post-error adjustments at $\tau_{CD} = 500$ ms (panels A and B), and second order post-error adjustments (panels C and D).** (A) Phase diagram of the PES effect. I used a bootstrapped confidence interval in order to decide whether or not PES (or PEQ) is observed. (B) Phase diagram of the PIA effect. The observation of post-error adjustments is highly impacted with the value of τ_{CD}, as I do not observe PES for the same range of parameters. (C) Phase diagram of the PES effect at the $n+2$ trial. (D) Phase diagram of the PIA effect at the $n+2$ trial. One sees rare isolated red squares, indicating the absence of any systematic effect. For all panels: Simulations with a RSI of 500 ms, other parameters as in Table 1. Color code as in Figure 9.

the post-correct dynamics (blue curve) for population L (the winning one for the second stimulus) is higher than the post-error one. This leads to a faster decision time for the post-correct trial than for the post-error one.

I will now show that the dynamics explains the three effects PES, PEQ and PIA. Figures 3.17, 3.18 and 3.19 provide a semi qualitative and semi quantitative analysis of the dynamics of the synaptic activities in the phase plane of the system, for several parameters regimes. Here again, the analysis is easier working on the synaptic activities. This can be seen by considering Figure 3.16 on which the mean firing rate and synaptic activity of the winning population in the PES case are represented. Due to the range of variation of the firing rates, and the intrinsic noise of the system, it is hard to observe a difference between the neural activities. However, the difference (sub-panel of Figure 3.16) can still give some insight about what happens. At the beginning of the next trial, the difference between the post-error and post-correct firing rates is significantly below zero, hence the reaction time will be shorter for post-correct than for post-error trials. The same behavior for the synaptic activities (Figure 3.16.B) is observed, but much less noisy.

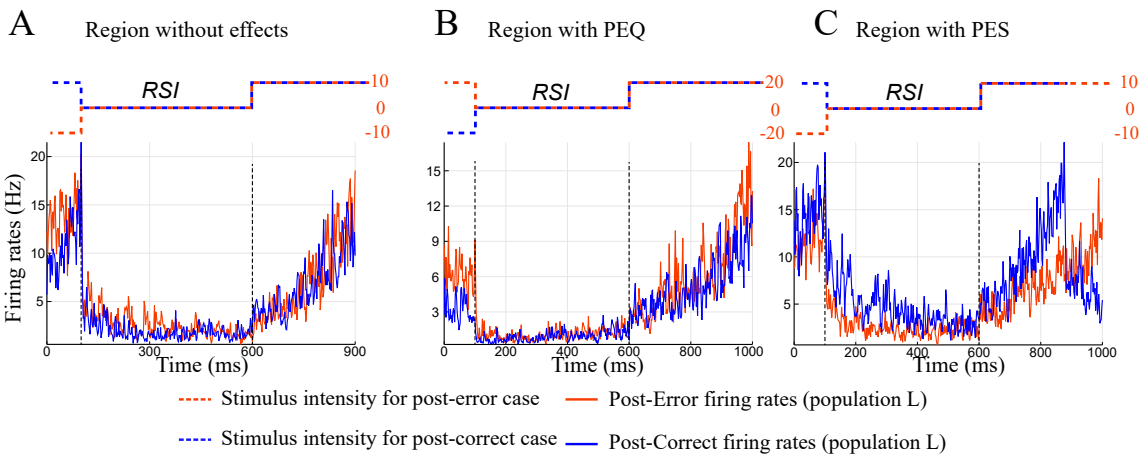


Figure 3.15: **Neural activities of individual trials.** (A) Dynamics for individual trials for the winning populations of the next trial: in blue the post-correct case and in red the post-error one. The dashed lines represent the coherence of the stimuli with respect to time. In blue I represent the post-correct case, and in red the post-error one. The parameters are set to a region without PES or PEQ effects ($I_{CD} = 0.047$ nA and $c = \pm 10\%$). (B) This panel represents the dynamics in the region of PEQ ($I_{CD} = 0.047$ nA and $c = \pm 20\%$). On this trial I can notice that the post-error dynamics is faster than the post-correct one. (C) The parameters are set to the PES region ($I_{CD} = 0.035$ nA and $c = \pm 10\%$). The post-correct dynamics (in blue) reaches the threshold sooner than the post-error one (in red).

PES effect.

Now, I will detail the analysis of the PES effect (and of the concomitant PIA effect) based in Figure 3.17. Each panel is done the following way. Without loss of generality, I assume that the last decision made is R. Repeated and Alternated cases thus correspond respectively to next trial decisions R and L. The x and y axis are the synaptic activities S_L and S_R, respectively – hence, the losing and wining populations for the first trial.

On the left panels, I represent with dashed lines the average dynamics during the relaxation period, that is from the decision time for the previous stimulus to the onset of the next stimulus. This allows to identify clearly the typical neural states at the end of the relaxation period. The average is done over post-error (resp. post-correct) trajectories sharing a same state at the time of the last decision. The choice of these two initial states is based on the following remark. A typical trial with a correct decision will lead, at the time of decision, to losing and wining populations with highly different activity rates, hence a neural activity, and thus a synaptic activity S_L, far from the threshold value. On the contrary, a typical error trial will show a losing activity not far from the threshold – this can also be observed in Figure 4B in Wong et al. (2007). I can thus represent post-correct trials, respectively post-error trials, by dynamics with initial states having a rather small, respectively large, value of S_L (and in both cases the first trial winning population S_R at threshold value).

Then I represent with a continuous line the average trajectory following the onset of the next stimulus. This dynamics is observed during the same time for post-error and post-correct cases – as if there were no decision threshold – in order to compare the dynamics of post-error and post-correct cases for the same duration of time. Decision actually occurs when the trajectory crosses the decision line (dashed gray line) – this is approximate: because of the noise, there is no one to one correspondence between a neural

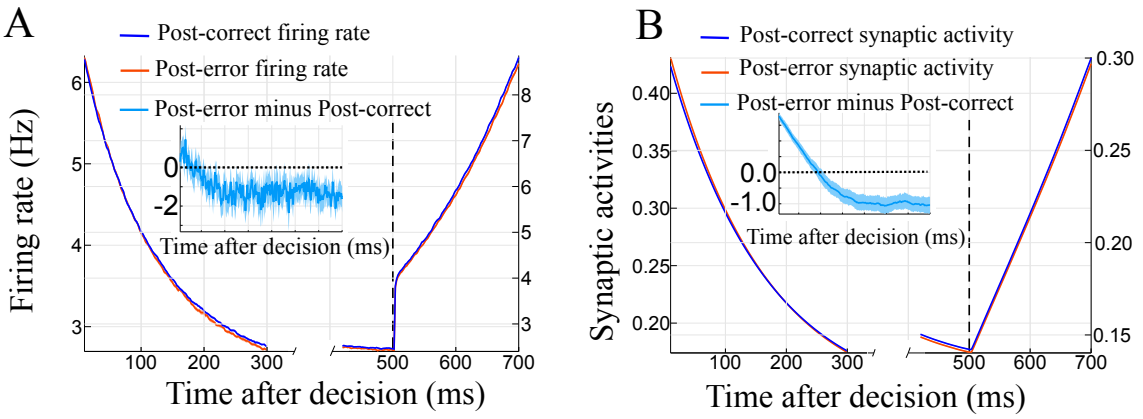


Figure 3.16: **Mean firing rates of the winning population.** (A) Mean firing rates for the winning populations of the next trial: in blue the post-correct case and in orange the post-error one. The ribbons represent the 95% confidence interval on 25 simulations(bootstrap method). The left-axis represents the relaxation of the dynamics. The right-axis is for the beginning of the next stimulus. The parameters are set to a region with PES effects ($I_{CD} = 0.035$ nA and $c = \pm 10$). The sub-panel with the light blue curve is the difference between the post-error firing rates and the post-correct with respect to time (in percent). The ribbon stands for the 95% confidence interval. As expected, this difference is negative. Hence the post-correct dynamics is faster and crosses sooner the threshold. This leads to the PES effect. (B) Mean synaptic activities for the winning populations of the next trial: in blue the post-correct case and in orange the post-error one. The sub-panel with the light blue curve is the difference between the post-error synaptic activities and the post-correct with respect to time (in percent).

activity reaching the decision threshold and a particular value of the associated synaptic activity. Having all the trajectories plotted for the same duration (and not only until the decision time) allows to visually compare the associated reaction times.

On the right panels, I represent typical trajectories during the presentation of the next stimulus. The black dot on every panel gives the location of the neutral attractor that exists during the relaxation dynamics. The basins of attractions that are represented are the one associated with the attractors L, R, of the dynamics induced by the onset of the next stimulus. It shall be reminded that these attractors are different from the ones associated to the dynamics during the relaxation period.

I can now analyze the dynamics. In the repeated case (Figure 3.17.A and B), at the end of the relaxation (that is at the onset of the next stimulus), both post-correct and post-error trials lie into the correct basin of attraction. Hence, the error rates for these trials are similar. However, the neural states reached at the end of the relaxations are different. Compared to the post-error trial, the post-correct state is closer to the boundary of the new attractor associated to decision R, and the corresponding decision will thus be faster. In the alternated case (Figure 3.17.C and D), the states reached at the end of the relaxation period do not lie within the correct basin of attraction. During the decision-making dynamics, the trajectory needs to cross the boundary between the two basins of attraction. The post-correct trials leading to an alternate decision have a rather straight dynamics across the boundary, leading to relatively fast decision times. In contrast, the states at the onset of the stimulus of the post-error trials are closer to the boundary so that the corresponding trajectories cross with a smaller angle with respect to the basin boundary. This leads to longer reaction times, hence the PES effect. It would be interesting

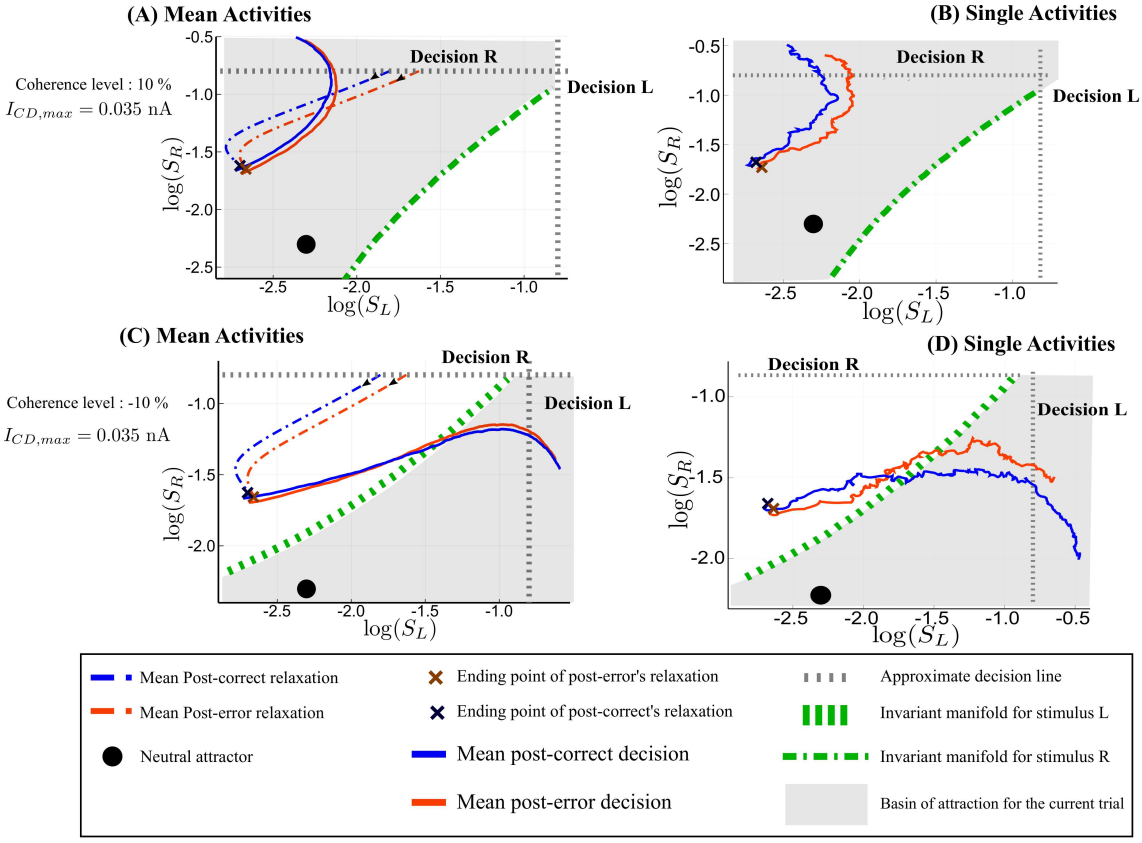


Figure 3.17: **Analysis of the post-error trajectories for the PES regime.** Phase-plane trajectories (in log-log plot, for ease of viewing) of the post-correct and post-error trials. I consider that the previous decision was decision R. The black filled circle shows the neutral attractor state (during the relaxation period). During the presentation of the next stimulus, the attractors and basins of attraction change (represented by the gray area and the green dashed lines). Panels (A) and (B): PES and PIA regime ($c = 10\%$ and $I_{CD,max} = 0.035$ nA) in the repeated case. The blue color codes for post-correct trials, and the red one for post-error. Panel (A): average dynamics; Panel (B): single trajectories during the next trial. Panels (C) and (D): regime with PES and PIA in the alternated case ($c = -10\%$ and $I_{CD,max} = 0.035$ nA). The dynamics after the relaxation is followed during 400ms for repeated and 800 ms for alternated case, as if there were no decision threshold. The actual decision occurs at the crossing of the dashed gray line, indicating the threshold.

to have electro-physiological data with which the model predictions could be directly compared. However, in a typical experiment on monkeys, a feedback on the correctness of the decision is given, since the animal learns the task thanks to a reward-based protocol. Nevertheless, I note that, in the random dot experiments on monkeys of Purcell and Kiani (2016), the authors find a higher buildup rate of the neural activity for post-correct trials than for post-error trials (see Figure 6 in Purcell and Kiani (2016)). Within the framework of attractor neural networks, this can be understood as trajectories that cross the basin boundary quicker for post-correct trials, in accordance with the model's predictions. This suggests that the observed difference in buildup rates may not result from some mechanism making use of the information on the correctness of the decision, but rather from the nonlinear dynamics discussed here.

The PIA is understood from the same analysis as for the PES effect. For specific realizations of the noise that lead to error trials, the post-error trials dynamics is closer to the boundary. Thus it has a higher probability to fall on the other side of the basin of attraction. Hence, the error rates are lower for post-error trials than post-correct trials.

PEQ effect.

The PEQ effect can be understood from the same kind of analysis, based here in Figure 3.18 (analogous for the PEQ effect to Figure 3.17 for the PES effect). As seen previously, the PEQ effect occurs mostly at high level of coherence. I consider first the repeated case (Figure 3.18.A and B). Since the coherence level is high, at the end of the relaxation period, both post-correct and post-error trials lie within the correct basin of attraction, far from the basin boundary. The reaction times and error rates of post-correct and post-error trials for repeated decisions are thus similar.

In contrast, the alternated case (Figure 3.18.C and D) exhibits both the PIA and the PEQ effects. The post-error's end of relaxation now is inside the basin of attraction of the alternated choice. Hence, the error rate will be lower than when the ending point is outside this region (post-correct trials begin at the boundary of the basin of attraction). Moreover, the post-correct trials dynamics have to cross the boundary. Hence they are closer to the manifold, which leads to slower dynamics, whereas the post-error dynamics can directly reach the new attractor state. This analysis explains why the decreasing of PES and PIA does not occur at the same coherence level too. Indeed the decreasing of PIA occurs when the ending point of the post-error relaxation crosses the boundary, whereas the post-correct ending point remains into the same basin of attraction. For the PES effect to decrease, the dynamics for both cases just need to be closer to the boundary and not necessarily on the opposite side. Hence the decrease of the PES effect occurs at lower coherence than the PIA one.

Here I have highlighted the fact that the occurrence of the PEQ effect depends on some very specific and fragile feature, the crossing or not of a basin boundary. The conditions for observing the effect are thus likely to vary from individual to individual, and from experiment to experiment. This may explain why the experimental results about the PEQ effect remain controversial.

In Figure 3.19 A and B I investigate the parameter regime, at low coherence level, for which there is no effect – neither PES, nor PEQ or PIA. The post-error and post-correct dynamics are very similar and lead to the same relaxation ending point, far from the basin boundary. Finally, in Figure 3.19 C and D I consider the parameter regime, at high coherence level, with only the PIA effect. Here the relaxations of post-error and post-correct trials are different. However, as for the PEQ effect, at high coherence level both dynamics will be fast. For alternated trials, none of the two ending points are in the correct basin of attraction.

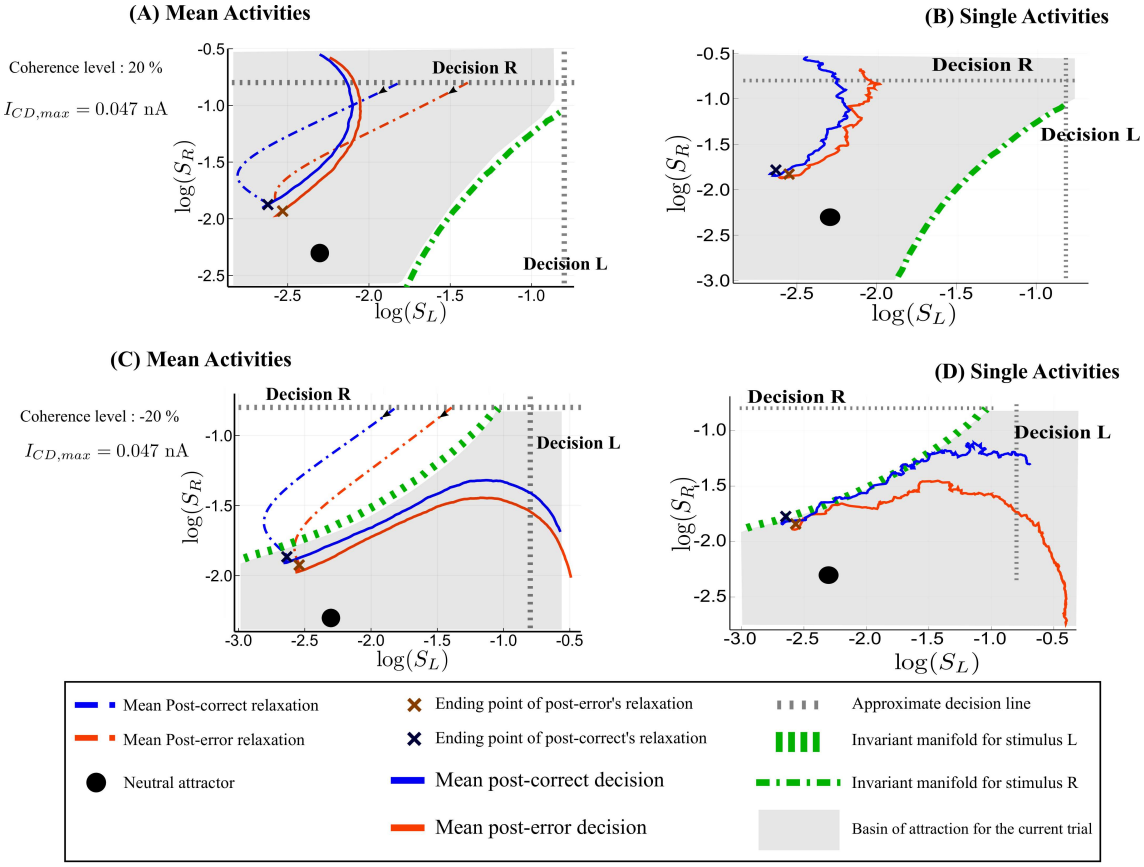


Figure 3.18: **Analysis of the post-error trajectories in the PEQ regime.** Phase-plane trajectories (in log-log plot, for ease of viewing) of the post-correct and post-error trials (same as Figure 3.17 in the PEQ case). I consider that the previous decision was R. The black filled circle shows the neutral attractor state (during the relaxation period). During the presentation of the next stimulus, the attractors and basins of attraction change (represented by the gray area and the green dashed line). Panels (A) and (B): PEQ and PIA regime ($c = 20\%$ and $I_{CD,max} = 0.047$ nA). The blue color codes for post-correct trials, and the red one for post-error. The plain lines represent mean dynamics for (A) or single dynamics (B). Panels (C) and (D): regime with PEQ and PIA in the alternated case ($c = -20\%$ and $I_{CD,max} = 0.047$ nA).The post-error relaxation already lies within the alternated basin of attraction. For alternated trials, the dynamics needs to cross the invariant manifold (green dashed line), which denotes the boundary between the basins of attraction. The dynamics is followed during 400ms for repeated and 800 ms for alternated case, as if there were no decision threshold. The actual decision occurs at the crossing of the dashed gray line, indicating the threshold.

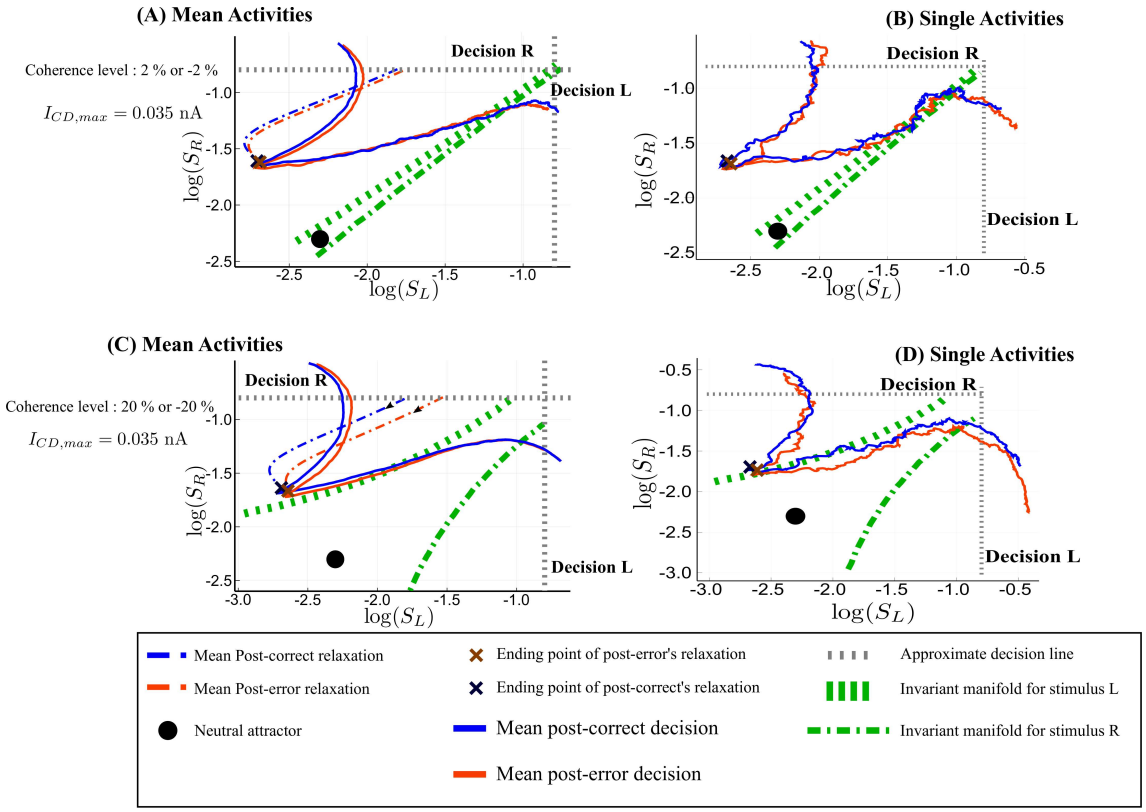


Figure 3.19: **Analysis of the post-error trajectories in the regime with neither PES nor PEQ effect** Phase-plane trajectories (in log-log plot, for ease of viewing) of the post-correct and post-error trials. I consider that the previous decision was decision R. The black filled circle shows the neutral attractor state (during the relaxation period). During the presentation of the next stimulus, the attractors and basins of attraction change (represented by the gray area and the green dashed lines). Panels (A) (mean dynamics) and (B) (single dynamics): regime without PES or PIA ($c = \pm 2\%$ and $I_{CD,max} = 0.035$ nA). I show both the alternated and the repeated case, with the corresponding basins of attraction. The blue color codes for post-correct trials, and the red one for post-error. For alternated trials, the dynamics needs to cross the invariant manifold (green dashed line), which denotes the boundary between the basins of attraction. Panels (C) (mean dynamics) and (D) (single dynamics): regime with PIA but without PES ($c = \pm 20\%$ and $I_{CD,max} = 0.035$ nA). The dynamics is followed during 400ms for repeated and 800 ms for alternated case, as if there were no decision threshold. The actual decision occurs at the crossing of the dashed gray line, indicating the threshold.

As discussed for the PES effect, electro-physiological data only exist for experiments with feedback on the correctness of the decision. In experiments on monkeys, Purcell and Kiani (2016) obtained puzzling results for what concerns the PEQ effect. They observed an important difference in baseline activities for post-correct and post-error trials, which is not well accounted for either by their DDM analysis or by the model. However, in terms of neural dynamics, this observed difference in the level of neural activities obviously implies that the dynamical states are different at the time of the onset of the stimulus, a fact in agreement with our model's predictions. One may wonder if the separation in baseline activities, and not just in starting points, could be a consequence of the feedback.

Correlating post-error effects with the activity distributions at the previous decision.

To go beyond the above analysis on the post-error adjustments (PES, PEQ and PIA effects), I analyze the respective influence of the winning and losing population levels of activity at the time of the previous decision, onto the decision at the next trial. This will first confirm the previous analysis, but also provide more insights on the the specificity of the two opposite effects, PES and PEQ.

The mean activity, at the time of the decision, of the winning population is indistinguishable between post-correct and post-error trials (Unequal Variance (Welch) test: Fail to reject, $p = 0.16$ at RSI of 500ms and Fail to reject, $p = 0.87$ at RSI of 2000ms). However, for short RSIs (corresponding to PES regime) the mean synaptic activities, at the time of the decision, of the losing population are different for post-correct and post-error trials (Unequal Variance (Welch) test: Reject, $p = 2.7 \times 10^{-20}$ at RSI of 500ms and Fail to reject, $p = 0.57$ at RSI of 2000ms).

To get more insights, Figure 3.20 represents the amplitude of the PES effect with respect to the inter-percentile range of the distribution of the synaptic activities of the winning and losing populations at the time of the previous decision. When PES occurs, the higher the activity of the losing population at the time of decision, the stronger this effect will be. The influence of the winning population is observed, although in an opposite way. When PES occurs these effects are correlated (Dark Blue: Pearson correlation: $r = -0.98$ and $p = 2.6 \times 10^{-7}$, Medium Blue: $r = -0.98$ and $p = 9.5 \times 10^{-7}$), in the sense that the variations with respect to the inter-percentile of the winning and losing population are correlated. These observations are consistent with the analysis of the PES phase-plane trajectories. Indeed, the higher the losing population activity is, the closer to the invariant manifold the state at the end of the relaxation period will be. Hence, the effect will be stronger as it becomes easier (more likely) to cross the boundary.

However, Figure 3.20, panels A and C, shows a different behavior for the PEQ effect: there is an almost constant value of the PEQ effect with respect to the inter-percentiles of the distributions of the winning and losing populations activities. This is explained by the fact that, at the end of the relaxation, if the category of the next stimulus is the opposite of the previous decision, the network state finds itself within the (correct) associated basin of attraction, but very close to the boundary. This is true whatever the correctness of the previous decision. However, the post-correct case will lead to an even closer location from the basin boundary. The nonlinearity of the dynamics near the basin boundary will strongly amplify the small difference between post-correct and post-error ending point. The PEQ effect will thus not be correlated with the size of this difference.

For what concerns the PIA effect, Figure 3.20.C-D shows a similar dependency in the synaptic activities as for the PES effect, with a stronger effect for high activities of the losing population. This corroborates the above phase plane analysis of the trajectories (Figure 3.14). Indeed, the PES and PIA effects both depend on the position of the relaxation in the phase plane. Being closer to the boundary (high activity of the losing population) leads to a smaller error rate in the next trial.

From the above analysis, a prediction of the model is that, whenever there are PES or PIA effects, the mean activity of the *losing* population is different for correct and error decisions. Moreover, this level of activity is correlated with the amplitude of the post-error adjustment effect. This can be seen in Figure 3.20, panels A, B. In this figure, I present the quantiles of the synaptic activities. The results would be similar, but much more noisy, for the firing rates. This prediction can be tested in experiments by measuring the correlation between the amplitude of the PES (or PIA) effect, and the difference in mean activities of

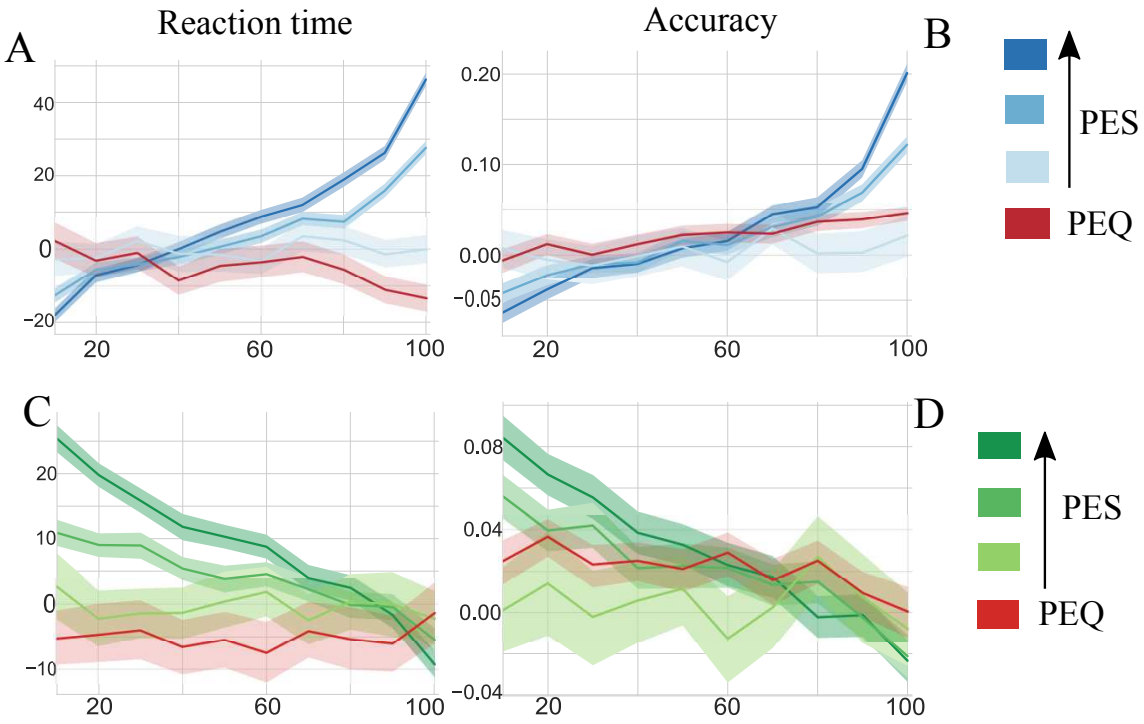


Figure 3.20: **Influence of the losing and winning population on the post-error adjustments.** Panels (A) and (B) represent respectively the reaction time (PES effect) and accuracy (PIA effect) with respect to inter-percentiles range of the losing population synaptic activity distribution, at a RSI of 500 ms. The red curve corresponds to $c = 18$ and $I_{CD,max} = 0.047$, where I observe PEQ. The dark blue one corresponds to strong PES effect ($c = 10$, $I_{CD,max} = 0.035$), the medium blue one to medium PES effect ($c = 5$, $I_{CD,max} = 0.05$). The light blue one corresponds to no effect at all ($c = 10$, $I_{CD,max} = 0.035$), for a RSI of 2 seconds. Panels (C) and (D) represent the same curves for the winning population, with the same color code. The shadow area represents the 95% bootstrapped confidence intervals of the corresponding effect.

the losing neural population (difference between post-error and post-correct trials).

4

A cognitive experiment to study confidence

Confidence judgments in one's decision are considered a central example of metacognition. How can we access someone else's subjective sense of confidence? In this chapter, I review different behavioral measures that have been developped in order to obtain this sense of confidence in humans and in animals. In a second part, I present the experiment I performed, in collaboration with Jean-Rémy Martin and Jérôme Sackur at ENS, during my PhD.

4.1. How to measure confidence experimentally ?

4.1.1 In humans

Explicit reports of confidence

Reports of confidence in humans can be made explicit through verbal communications for example. Thus, the most straightforward paradigm to measure confidence is to ask the subjects to assign at each trial a numerical rating corresponding to how sure they are of their answer (Zizlsperger et al., 2014; Fleming and Lau, 2014). The participants provide a subjective probability on the correctness of their response as a confidence judgment. One important aspect of such reports is that performance accuracy and response times are well-correlated with self-confidence reports (Baranski and Petrusic, 1994; Pleskac and Busemeyer, 2010). This phenomena occurs for a variety of tasks such as general knowledge tests (Perfect et al., 1993), perceptual decisions (Fleming et al., 2010) and reasoning tasks (Stankov, 2000). In Figure 4.1, I represent an example of such relation between these behavioral variables. The experiment consisted in a categorization task (with Gabor patches) followed by a confidence judgment on a four-points scale (Adler and Ma, 2018). The data are publicly shared through the *confidence database* project (Rahnev et al., 2020). The most striking effect that appears from having access to direct reports of confidence is the *under/over-confidence* (Lichtenstein and Fischhoff, 1977). These deviations occur systematically with overconfidence when the decisions are difficult and underconfidence for easy decisions (Kepecs and Mainen, 2012). For instance, in Figure 4.1, the orange participant clearly shows under-confidence and the blue ones over-confidence. However, it is worth noting that these biases in confidence vary greatly with the type of judgments that are asked, and across participants (Klayman et al., 1999).

Other measures of confidence

In the recent years, some concerns have arised when reporting confidence on a scale. For example, should the scale be discrete or continuous ? (Lisi et al., 2018) Will some participants only use one end of the scale? (Morgan et al., 1997) To try to address these concerns, various paradigms have been proposed. For example, one can give the option to the participants to opt-out of the trial if they feel too uncertain (Gherman and Philiastides,

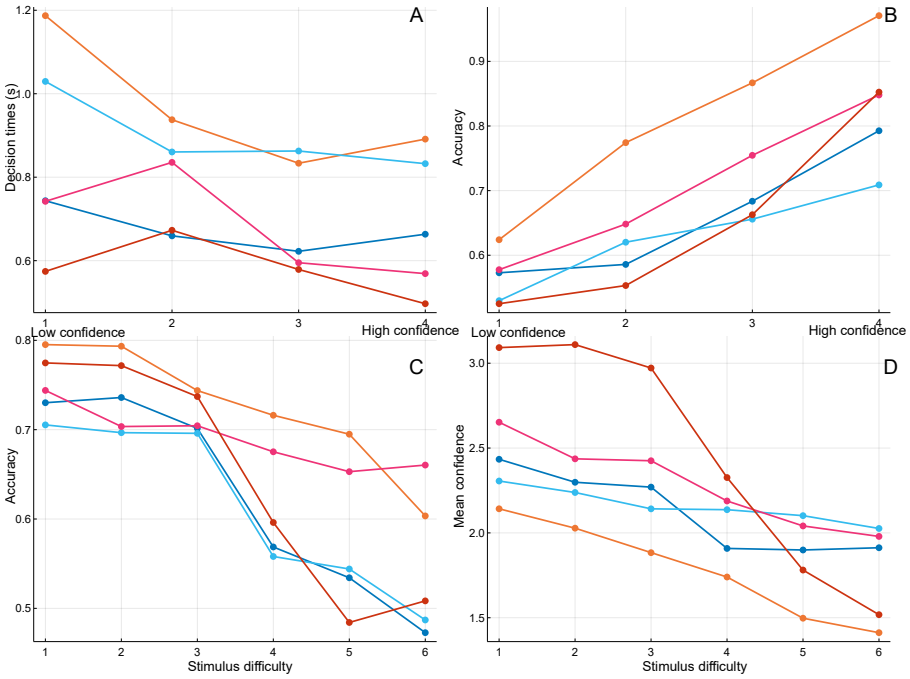


Figure 4.1: **Behavioral results (data from Adler and Ma (2018)).** Each color represents a different participant. (A) Decision times of the participants with respect to their self-reported confidence. (B) Accuracy of the participants with respect to their self-reported confidence. (C) Accuracy with respect to the stimulus difficulty. (D) Mean confidence with respect to the stimulus difficulty.

2015). However, this kind of task could be interpreted as a three-choices tasks instead of a confidence report. Postdecision wagering has been proposed to improve confidence rating (Persaud et al., 2007). In this paradigm participants must bet on the outcome of their decisions. The participants will have a tendency to bet more for trials with an higher feeling of confidence. Finally, I would like to mention a last paradigm used to measure confidence in humans. Two stimuli are shown to the participants on each trial. They must report for which one they feel they are more likely to be correct. By comparing the performances between the trials chosen and the ones that were not, one can identify if the sense of confidence of the participant is correct (de Gardelle and Mamassian, 2014). This method does not rely on an explicit report of confidence by the participants and is therefore not affected by the participants' bias of under/over-confidence.

4.1.2 In animals

For animals, one can not simply ask them to explicitly report their confidence. Therefore, more sophisticated tasks have been employed to elicit a report of confidence in animals.

Uncertain option task

One of the most common task used for this purpose is the *uncertain option* task. It extends the classical two-choices paradigm that has been mentioned in this thesis. In addition to the two available responses, the animal is offered a third choice, that will correspond to a small but certain reward. This framework has been used in many species such as monkeys, dolphins and rats (Smith et al., 1995, 1997; Hampton, 2001; Kiani and Shadlen, 2009). Interestingly, when compared to human performances on this type of task, dolphins and

monkeys showed qualitative similar strategies and response distributions. However, one can address the following criticism to this kind of task: there is the possibility that this task is considered as a three-choices task by the animal. The animal could be learning the association between the uncertain reward and a difficult task.

Opt-out task

To address this problematic, Hampton (2001) developed a modified version of the uncertain option task. It consists in a memory task in which monkeys perform a delayed-maching-to-sample task. At the end of the delay period, monkeys were presented with the option of declining or accepting the discrimination test. Moreover, Hampton (2001) imposed that on some trials the monkeys have no choice but to make the discrimination test. This is to ensure that the monkey is not learning to associate longer delay with *opt-out option*. They found that the performances of the monkeys on freely chosen trials are greater than the ones in forced choice trials. This opt-out task has been used in macaque monkeys too (Kiani and Shadlen, 2009; Komura et al., 2013), but for a binary categorization of visual motion task. They found that the frequency of choosing the opt-out option increased with stimulus difficulty, as well as greater performances on freely chosen trials. Interestingly, this task has been tested in pigeons and rats but the researchers did not find a change in performances between forced-choices and free-choices. These results raised the question of whether these animals could perform confidence judgments or not.

Decision restart and leaving decision tasks

In order to address the criticisms of two previous paradigms, such as the fact that, for these paradigms, either a confidence report or a decision report is collected at each trial, Kepecs et al. (2008) proposed a new paradigm in rodents. The rats were trained to perform a 2AFC olfactory discrimination task. Depending on the dominant part of the odour mixture, the rats needed to choose the left or the right. A variable delay was imposed after a correct trial and the animal could restart the trial at will (Figure 5.8). There was no feedback on error trials which allowed to measure the rats confidence in these trials. For some of the correct trials, the reward was omitted which allowed to measure confidence for correct trials too. It has been found that the waiting time increased with respect to odour contrast for correct trials, but decreased for error trials. Moreover, accuracy was an increasing function of the waiting time. This suggests that the waiting time of each trial consists in a robust proxy for confidence.

Together these different paradigms have shown that humans, primates and rodents have access to a sense of confidence and can use it to optimize the rewards.

4.2. A 2AFC task in humans

In this section I will describe the cognitive experiment that I have performed to study decision-making and confidence in humans. Previously to the design of this experiment, I briefly analyzed the results of an experiment of Jean-Rémy Martin and Jérôme Saclur within the framework of attractor neural network. The preliminary results led to the design of an experiment whose goal was to study the impact of confidence on decision-making and various sequential effects.

4.2.1 Experimental set-up

The experiment was performed at the Laboratoire de Psychologie cognitive et de Psycholinguistique's database (LSCP, DEC, ENS-EHESS-CNRS, PSL, Paris, France). The experiment followed the ethics requirements of the Declaration of Helsinki (2008) and has been approved by the local Ethics Committee. It consists in a direction categorization

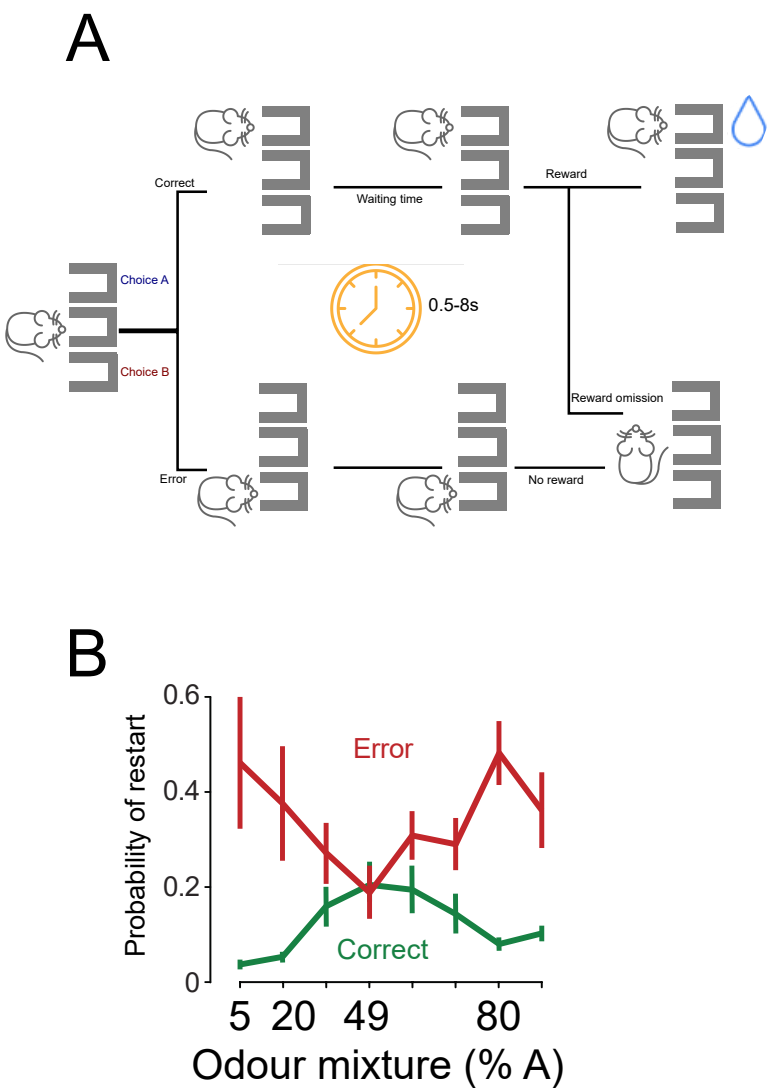


Figure 4.2: **Waiting task in rodents.** (A) The animal makes a perceptual decision and then needs to wait for the reward (on most correct trials). The rat can choose to wait for the reward or to start a new trial. (B) Figure adapted from Kepecs et al. (2008). Probability of reinitiation for a single rat plotted as a function of odour stimulus and trial outcome.

task, where participants classify Gabor patches as clockwise or anti-clockwise. In some of the trials, the decision was followed by an auditory feedback or a confidence evaluation.

The stimuli were generated using Matlab along with the Psychophysics toolbox (Brainard, 1997). They were displayed on a monitor at 57.3 cm in front of the participants' head. The participants performed the experiment in a quiet and darkened experimental room. Their heads were stabilized thanks to a chin-rest. The instructions given to the participants (translated from french) were the following (the emphazised sentences correspond to additional information not provided to the participants):

- In each trial, you will see very briefly (200 *ms*) a black dot at the center of the screen that you will need to look at (Figure 4.3.A). Right after the dot disappears (*200 ms*), you will see a circular grating at the center of the screen like the one in Figure 4.3.B.

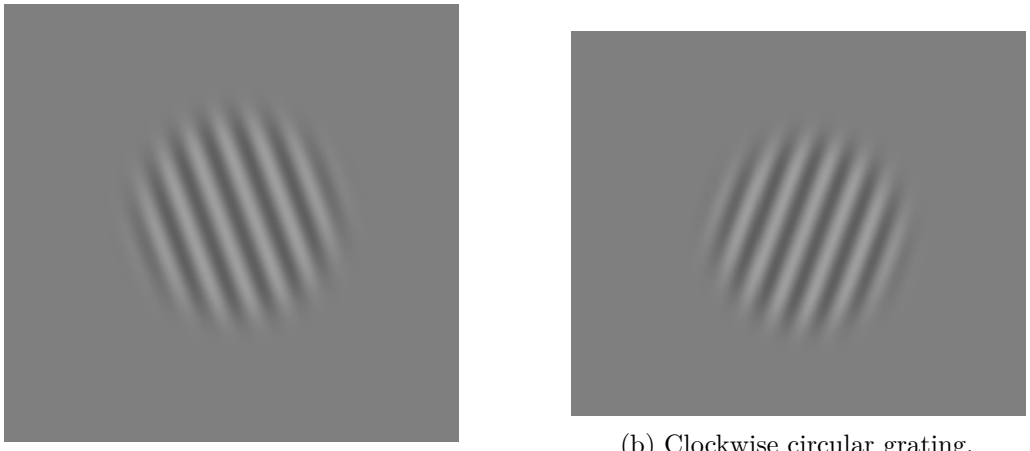
The parameters of the circular grating are diameter = 4°, Tukey window, 2 cycles per degree, Michelson contrast = 89%, duration = 100 ms, phase randomly selected at each trial.

(a) Fixation dot

(b) Example of the circular grating.

Figure 4.3

- In each trial you will need to indicate if the grating was oriented clockwise or anti-clockwise (Figure 4.4).
- As soon as the disk disappears, if you think it was anti-clockwise oriented you will press the left directionnal arrow. If it was clockwise oriented you will press the right directionnal arrow.



(a) Anti-clockwise circular grating.

(b) Clockwise circular grating.

Figure 4.4

- In the case where you do not know at all which direction it was, you will still press one of the two keys by following your intuition. When this happens, do not always press the same key.
- You need to answer fast but not at the expense of accuracy. After 1.5 second, you will see a message at the screen: "Please, answer". Ideally, you will answer before this text appears at the screen.
- You will have 3 bocks of trials:

– In the first block, once you have answered to a trial, the software will automatically run the next trial.
– In a second block, you will receive a feedback on your answer at each trial. If the answer was correct, you will hear a pitched tone. If your answer was wrong you will hear a deep tone.
– In a third block, you will need to evaluate your confidence level in your answer using the scale that will appear on the screen (Figure 4.5). You will move the black slider towards right or left using the left and right stickers on the keyboard (q and e keys). The scale is composed of 10 levels. The case at the left corresponds to the "Pure guessing" case: you choose randomly the orientation of the grating. At the right, you have the "Certain to be sure" case: you are absolutely certain to be right, there is no possible doubt. Between these two cases, you have access to intermediate levels of confidence. The choice of the confidence level is performed using the space bar.

Figure 4.5: Confidence-scale.

- Be aware that, in the confidence block, the black slider will appear randomly on the scale. Do not be biased by the initial position of the cursor and move it to the level that reflects your level of confidence.
- During the confidence block, in the case you wanted to choose the right arrow but you chose the left one (or inversely), do not answer to the confidence scale. Just press the keyboard key with a red sticker on it, it will go to the next trial.
- The software will choose the order of the block and you will know at the beginning of each block which one it is. Before the main experiment starts, you will have a small training.

Nine participants (7 Females, Mean Age = 27.3, SD = 5.14) have been recruited from the Laboratoire de Psychologie cognitive et de Psycholinguistique's database (LSCP, DEC, ENS-EHESS-CNRS, PSL, Paris, France). Every subject had normal or corrected-to-normal vision. The participants performed three sessions on three distinct days in the same week for a total duration of about 2h15. Three participants were excluded. Two of the excluded participants did not complete correctly the experiment and one exhibited substantially asymmetric performance (98% of correct responses for an angle of 0.2°, but 18% at -0.2°degree). As a result, I analyzed data from 6 participants. I obtained written informed consent from every participant who received a compensation of 15 euros for their participation. Participants performed three sessions on three distinct days. Each session (45 min) consisted in three runs, each run being composed of one exemplar of each of the three types of block, in a random order.

4.2.2 Behavioral measures

The experimental procedure is shown in Figure 4.6. The waiting time between each trial was deliberately chosen to be short and similar between blocks, in order to study the sequential effects.

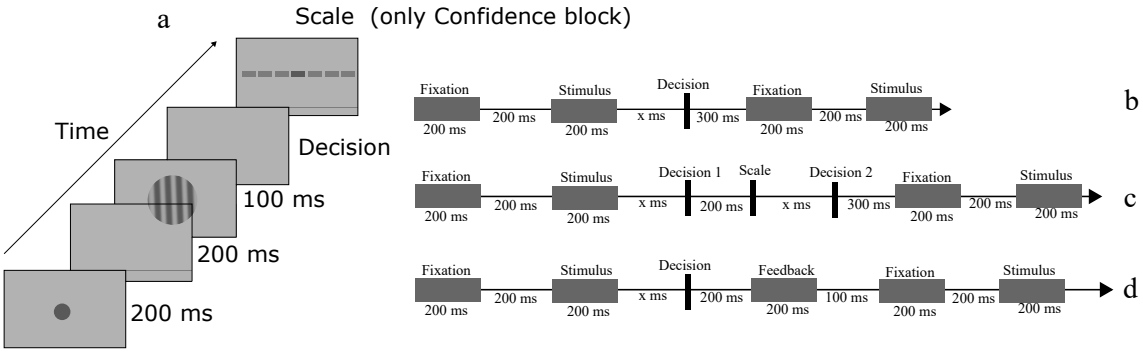


Figure 4.6: **Procedure of the discrimination task, for the three blocks.** (A) Structure of a trial: Following a fixation period, the circular grating (Gabor patch, oriented clockwise, C, or counterclockwise, AC) appears and participants make the decision (C or AC). In confidence blocks, after a delay, participants report their confidence with respect to their choice, on a discrete scale with 10 levels. (B) Time course of a *pure* block trial. (C) Time course of a *confidence* block trial. (D) Time course of a *feedback* block trial.

Pure block

In this block, participants waited 300 ms after each decision, before the black fixation point appeared. The stimulus appeared 200 ms after this fixation point. The eight possible orientations for the circular grating were [-1.6°, -0.8°, -0.5°, -0.2°, 0.2°, 0.5°, 0.8°, 1.6°] and a stimulus was chosen randomly among them with the following weights: [0.05, 0.1, 0.15, 0.2, 0.2, 0.15, 0.1, 0.05].

Feedback block

In this block, 200 ms after the decision, the participants received an auditory feedback (during 200 ms) about the correctness of the decision they just made. The black fixation dot appeared 100 ms after this feedback and then a new trial began. The orientations of the circular gratings were chosen randomly from [-1.6°, -0.8°, -0.2°, 0.2°, 0.8°, 1.6°] with the following weights [0.12, 0.18, 0.2, 0.2, 0.18, 0.12].

Confidence block

In the confidence block, participants had to evaluate the confidence on the orientation task 200 ms after the decision. After the choice of confidence, the participants had to wait 300 ms before the black fixation dot appeared. The stimulus appeared 200 ms after the fixation dot. The orientations of the circular gratings were the same as in the feedback block.

I will present different results from this experiment, without the framework of attractor network that will be discussed in the next chapter of this manuscript. First, one can see that the participants show increase of accuracy and decrease in response times with respect to the stimuli difficulty (Figure 4.7). Despite the variability across participants, the trends are globally the same. The goal of the experiment was to study sequential effects due to confidence and post-error adjustments. It came as a surprise that, when comparing the performances using a linear mixed model in the pure block and in the feedback block, there was no difference (Figure 4.7). However, it is worth noting that the participants were highly trained in the orientation discrimination which could explain the absence of differences between the two blocks. On the contrary, Figures 4.7.C and D show the data of the pure block and of the confidence block. The fact that the participants had to give their confidence on their decisions had an impact on the behavioral performances: accuracy was

higher and response times slower. This effect was already observed by Jean-Rémy and Jérôme in a previous experiment.

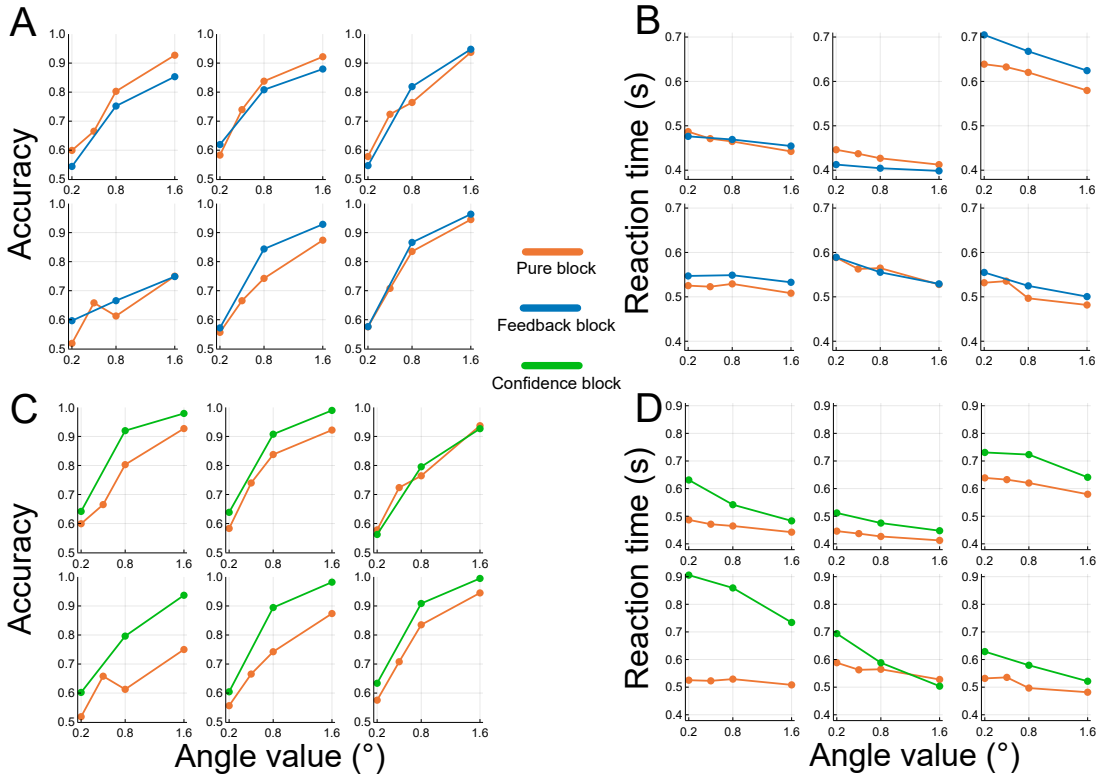


Figure 4.7: **Behavioral results of the experiment.** Eah subpanel represents a different participant. (A) and (C) represent the accuracy of the participants in the different blocks with respect to the orientaiton of the stimuli. (B) and (D) show the response times of each participant.

To test this effect more specifically I ran a binomial regression of responses with fixed factors of orientation, type of block (pure or confidence), the interaction between these factors and a random participant intercept. The orientation coefficient was 2.15 (SD $= 0.17$, $z = 12.44$ and $p < 10^{-16}$); there was no effect of block type ($p = 0.385$). There was a significant orientation by block type interaction (value of 0.55, SD $= 0.08$, $z = 6.97$ and $p = 3 \cdot 10^{-12}$), indicating that participants were more accurate in confidence blocks than in no-confidence blocks. In a similar way, I tested the effect on response times by using a mixed effect regression with the same factors and intercept as for the accuracy (only on the absolute value of the orientation). The orientation coefficient (value of -0.08, SD $= 0.013$ and $p = 0.0006$) and the block type coefficient (value of 0.095, SD $= 0.028$ and $p = 0.011$) were significant, meaning that participants are slower in the confidence block. Moreover, the slope by block type interaction with orientation was also significant (value of -0.028, SD $= 0.010$ and $p = 0.031$), meaning that the difference between the two types of blocks is more important at low orientation.

5

Confidence in perceptual decision-making

In daily life, we usually ask ourselves whether it was the right or the wrong decision, after we make one. For example, someone wants to cross the road at a busy intersection. They will need to assess the speed of the vehicles as well as the size of the gaps between them. To judge whether it is safe to cross the road or not, one should also estimate how trustworthy one's inferences are about speed and distance. Thus, this ability to estimate the accuracy of a decision is critical in everyday life. In this chapter, I first adress different models of confidence in decision-making. In a second part, I focus on the impact of confidence on sequential effects during perceptual decision-making.

5.1. Models of confidence in perceptual decision-making

Since the early work on decision-making, confidence judgments have been recorded alongside with decisions. In Peirce and Jastrow (1884), participants were asked to report their confidence in the perceptual decision they just made. In this work, the task was to discriminate between pressures applied to their fingers and to report their confidence on a four-ratings scale analog to the one described in the preceeding chapter. Surprisingly, the confidence ratings of the participants could be described by the following formula:

$$w = c \log \frac{p}{1-p} \tag{5.1}$$

with w the degree of confidence, p the probability of being right and c a constant called the index of confidence. However, this type of measure is sensible to many undesirable effects such as the fact that participants interpret differently the confidence scale thus displaying over- or under-confident behavior (Fleming and Lau, 2014). For this reason, current confidence ratings' analysis requires more detailed analyses and I will describe the main contempory frameworks of confidence.

5.1.1 Signal detection theory (SDT) framework

In SDT, the observer receives an observation of evidence e (Green et al., 1966). This observation is caused by a stimulus and differs depending on the category the stimulus belongs to. However, the evidence e is corrupted by noise, thus the likehoods for the stimulus to belong to one category or the other are overlapping. For example, let us consider two categories C_1 and C_2 that follow normal distribution with means ($\mu_{C_1} = -0.5$ and $\mu_{C_2} = 0.5$) and equal variance ($\sigma^2_{C_1} = \sigma^2_{C_2} = 1$) (Figure 5.1.A). The task of the observer is to infer the posterior probability $P(C = C_1|e)$. Using Bayes' rule, this posterior probability can be rewritten as:

$$P(C = C_1|e) = \alpha P(e|C_1) \cdot P(C = C_1) \tag{5.2}$$

with α a normalization factor computed through the relation $P(e|C_1) = P(e|C_2) = 1$. In the following, I will suppose that $P(C = C_1) = P(C = C_2) = 0.5$ (Figure 5.1.B).

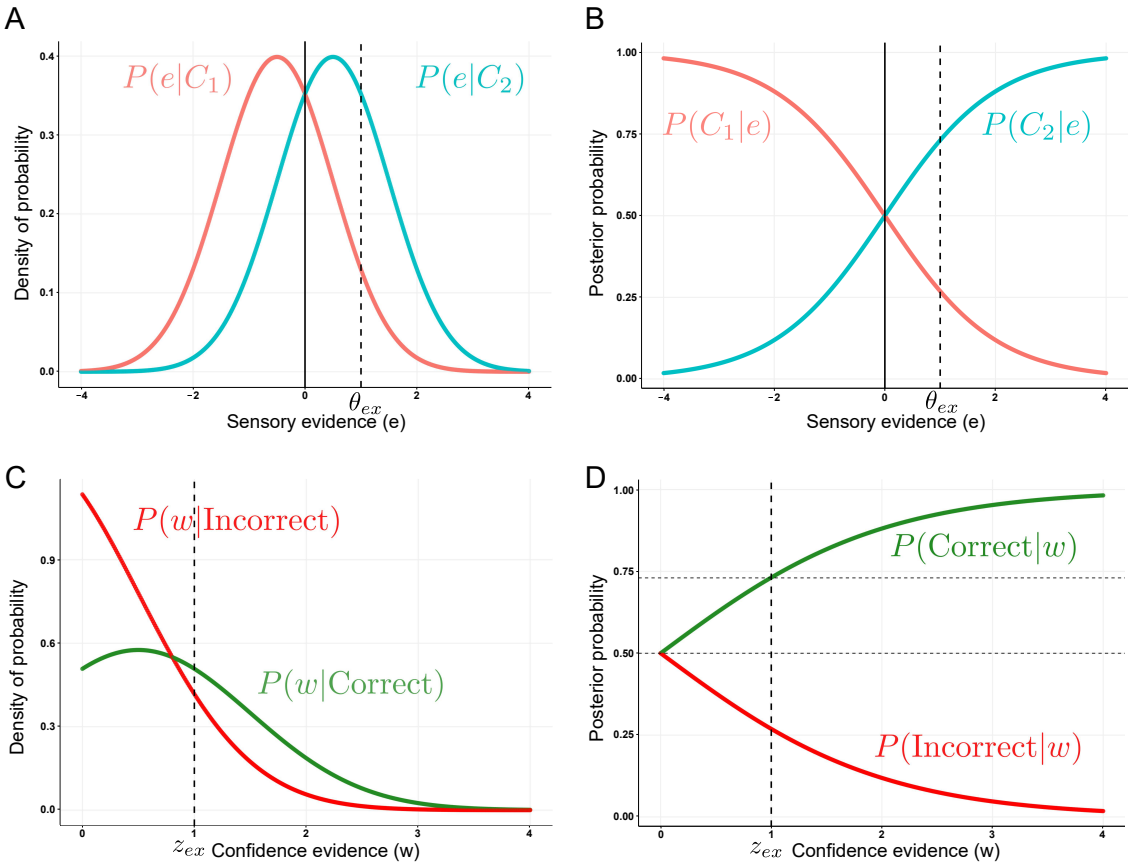


Figure 5.1: **Signal detection theory framework.** (A) Density of probability of the two categories C_1 and C_2 as defined in the text. The line at 0 denotes the value of the optimal criterion. The dashed line at θ_{ex} corresponds to a criterion slightly biased towards category C_1. (B) Posterior probability of the categories with respect to the sensory evidence e. (C) Density of probability of the confidence evidence for correct (green) and incorrect (red) trials. The dashed line at z_{ex} corresponds to an example of a confidence criterion that is slightly underconfident. (D) Posterior probability of the correctness of the trial with respect to the confidence evidence.

To achieve such estimation, the observer places a criterion θ along the sensory evidence axis (Figure 5.1.A). The chosen category will then be C_2 if the sensory evidence e exceeds this criterion, and C_1 otherwise. In the setup of Figure 5.1, the symmetry of the problem indicates that the optimal criterion is $\theta = 0$ (Figure 5.1.A), but the observer could choose another value, for example to model biases in decisions.

The observer's evidence about the correctness of their decisions varies along a confidence axis (w). There are many possibilities to model the confidence evidence within the signal detection theory framework. For instance, one could assume that the confidence evidence is the distance between the sensory evidence and the criterion θ (Clarke et al., 1959) or that it is the likelihood ratio of the sensory evidence given that the perceptual evidence was correct or incorrect (Galvin et al., 2003). In the example I am considering, I will assume that the confidence evidence is directly the distance between the sensory evidence and the

criterion θ:

$$w = |e - \theta| \tag{5.3}$$

It is important to distinguish between the confidence evidence for correct and incorrect decisions. Indeed, in this analysis of confidence evaluation, the categories "I am correct" and "I am wrong" will play the role of the categories C_1 and C_2. The likehood of confidence evidence for correct decisions is:

$$P(w|\text{Correct}) = P(|e-\theta||\text{Correct}) = P(e-\theta|C_2) + P(\theta - e|C_1) \tag{5.4}$$

This likelihood is illustrated in Figure 5.1.C in the case of an optimal criterion $\theta = 0$. Finally using the Bayes' rule, one can obtain the posterior probability of being correct (or incorrect) with respect to the confidence evidence (Figure 5.1.D).

$$P(\text{Correct}|w) = k \cdot P(w|\text{Correct}) \cdot P(\text{Correct}) \tag{5.5}$$

$$P(\text{Incorrect}|w) = k \cdot P(w|\text{Incorrect}) \cdot (1 - P(\text{Correct})) \tag{5.6}$$

with k a normalization constant.

To decide between high and low confidence trials, the observer places a second criterion z along the confidence evidence axis (Figure 5.1.C). If the evidence w is higher than z, the trial is favored towards the high confidence hypothesis. However, it is important to note that, in this case, there is an optimal location for the criterion z too. This location corresponds to the intersection of the likelihood of confidence evidence (Figure 5.1.C). If the criterion is placed after this value, the observer will show risk-aversion. In contrast if the criterion is placed closer to the origin, it will increase the risk of assigning high confidence to incorrect trials thus being overconfident. To conclude, this framework can show different behaviors observed in the experiments such as risk-aversion and overconfidence. However, SDT models assume that the decision is made with a fixed amount of time due to the fact that the sensory evidence is drawn from an unique sample. Thus, this model cannot describe the relationships between confidence, response time and stimulus difficulty in the case of the two-alternatives forced choice task.

5.1.2 Accumulation of evidence framework

To model the dynamics of the decision process as well as the confidence in one's decision, one solution is to use evidence-accumulation models. I focus here on decision models where the decision is made when the accumulation process reaches a specific bound z. Multiples variations of accumulations can be found in the literature, such as the drift-diffusion model (Ratcliff, 1978; Bogacz et al., 2006) or the independent race model (Raab, 1962; Vickers, 1970; Merkle and Van Zandt, 2006). The IRM being slightly more general than the DDM, I will be using the IRM framework to present the modelisation of confidence, in the case of a two-alternative forced choice task.

When an observer O is presented with a stimulus S, it initiates two simultaneous races representing the evidence of favor of both options (Figure 5.2.A). In order to model confidence with this model, Vickers (2014) proposed to define the confidence as the balance of evidence (Figure 5.2.A) at the time of the decision. One should note that this definition means that the losing race plays a role in confidence evaluation, even if it does not play a role in the decision. When the two races are close at the time of the decision, the balance of evidence is small. This means that the confidence in this decision is going to be low, as a small perturbation in the races would have led to the opposite decision. On the contrary,

if the races are far apart, the balance of evidence is high and the confidence will be high too. The equations for the IRM are the following:

$$\dot{x}_1 = \mu_1 + \sigma\eta_1(t) \tag{5.7}$$
$$\dot{x}_2 = \mu_2 + \sigma\eta_2(t) \tag{5.8}$$

with x_1, x_2 the decision variables, η_1, η_2 the white noises corresponding to each alternative and μ_1, μ_2 the mean drift of each races.

I recall that the decision is made once one of the races has reached a fixed threshold. This means that the balance of evidence can be characterized just by the state of the losing race at the time of the decision. Without loss of generality, the winning race corresponds to x_1, and I define $\Delta x = z - x_2$ the balance of evidence at the time of decision. The decision is correct if, indeed, $\mu_1 > \mu_2$. Using Fokker-Planck equation, one can find that:

$$P(\mu_i \,|\, x_i, t) = \frac{1}{\sqrt{2\pi\sigma^2/t}} \exp\left(-\frac{(\mu_i - x_i/t)^2}{2\sigma^2 t}\right) \tag{5.9}$$

The confidence in the decision corresponding to race 1 is defined by $P(\mu_1 > \mu_2 \,|\, x_2, t, x_1 = \theta)$ (Moreno-Bote, 2010). The probability of having chosen the right race corresponds to the probability that a Brownian motion finishes at state Δx or lower, meaning that the noise was not strong enough to elicit the wrong decision. This leads to:

$$P(\mu_1 > \mu_2 \,|\, x_2, t, x_1 = \theta) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\Delta x/(\sigma\sqrt{t})} \exp\left(-z^2/2\right) dz \tag{5.10}$$

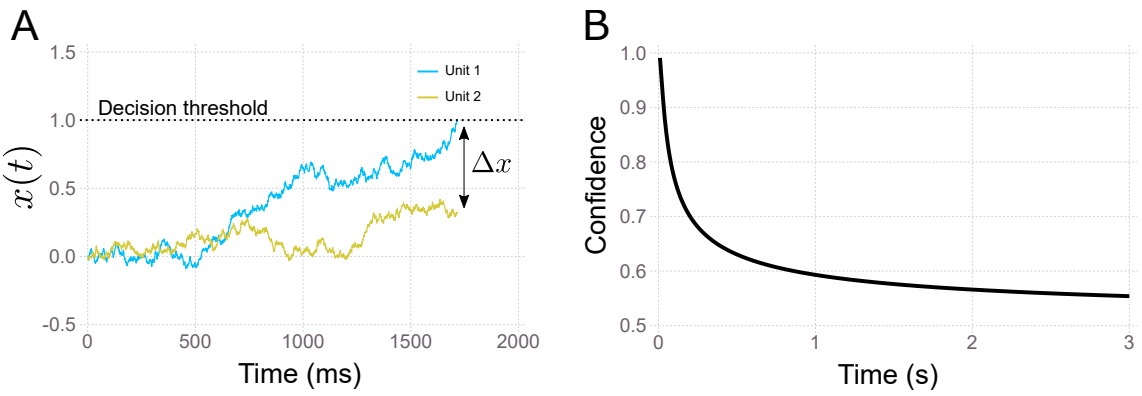


Figure 5.2: **Independent Race Model.** (A): Example of the dynamics for two independent races (denoted by each color). The parameters are $\theta = 1$, $\sigma = 1.0$, $\mu_1 = 0.07$ and $\mu_2 = 0.05$. The balance of evidence corresponds to the arrow Δx at the time of the decision. (B): Confidence with respect to the decision time (Equation 5.10). The parameters are $\sigma = 3$, $\theta = 1$ and $x_2 = 0.5$ (the state of the losing race at the time of the decision).

This framework gives a relation between response times and confidence that is represented in Figure 5.2.B. Confidence is a monotonic function of decision time, that decays from 1 to 0.5. This decay is explained by the fact that fast trials are more likely to correspond to higher drift rates, thus higher performances. Experimental measurements have found such decay of confidence with respect to decision times (Vickers, 2014).

Other models have been proposed to model confidence within an evidence-accumulation framework. Indeed, in the case of DDMs there is no losing race. To use the notion of

balance of evidence, it has been proposed that the accumulation process continues after the decision has been reached (Pleskac and Busemeyer, 2010; Navajas et al., 2016) and the final location of the process is used as a proxy for confidence. Other models propose to use different DDMs, each of them corresponding to a specific confidence level, the winning DDM determines the confidence level of the trial (Ratcliff and Starns, 2013). In any case, all these models have in common the fact that they attempt to characterize decision times, accuracy and confidence, within a perceptual task, using a mechanism based on the accumulation of evidence.

5.1.3 Neural substrates of decision confidence

Recent experimental studies have focused on understanding the representation of confidence, seen as the subjective probability the decision chosen is correct, within the brain. The first study addressing this question was the one of Kiani and Shadlen (2009). The authors recorded neurons in the lateral intraparietal cortex (LIP) of behaving monkeys. The animals were performing a specific version of the random dot motion task (RDM) specifically designed to study confidence (opt-out task). In addition to the two choices corresponding to the dots moving towards left or right, the monkeys could choose a third target corresponding to a small but certain reward. Interestingly, the monkeys chose the sure option in a way that is correlated with the chance of making the correct decision. They chose the sure target more frequently when the probability of making the correct decision was small (weaker motion strengths and smaller decision times). By analyzing the single-cell recordings of the LIP neurons, Kiani and Shadlen (2009) showed that intermediate firing rates at the time of the presentation of the sure target correspond to a higher probability of choosing this target. This experiment linked the mechanisms of decision formation with the establishment of a degree of confidence.

In the previous study, the link between the LIP neurons and confidence was implicit. Experimental studies have found neurons that represent confidence explicitly: orbitofrontal cortex in rats (Kepecs et al., 2008) and pulvinar neurons in monkeys (Komura et al., 2013). In these studies, a single firing rate trace is informative of the confidence level. In the opt-out task with the monkeys, the pulvinar neurons predict successfully the upcoming behavior of the monkey. Moreover, inactivation of the pulvinar neurons did not lead to modification of the performances in the categorization task but rather increased the probability of choosing the sure target. These results indicate that confidence does not necessarily need *metacognition* but can be computed using the decision variables (Figure 5.3).

If these results are in accordance with the balance of evidence framework in evidence-accumulation model, one effect should be noted. It has been shown that confidence in correct choices is stronger than confidence in incorrect choices, even when the choice difficulty is controlled (Fetsch et al., 2014). In linear models such as DDM or IRM, this effect can not be produced as the diffusion model parameters are invariant across correct and incorrect trials. One solution to solve this problem is to consider diffusion models that use post-decision evidence to estimate confidence. Another solution is to consider attractor neural networks to model the decision-making process (Wang, 2002).

Various models have been proposed to model subjective confidence using attractor neural networks (Rolls et al., 2010; Wei and Wang, 2015; Paz et al., 2016; Berlemont et al., 2020). The most successful one consists in defining a relation between the confidence and the difference of neural activity of the two neural populations at the time of the decision (equivalent to the balance of evidence) (Wei and Wang, 2015; Berlemont et al., 2020). Wei and Wang (2015) simulated the opt-out paradigm of Kiani and Shadlen (2009) using a continuous attractor neural network. They showed that such network can reproduce the

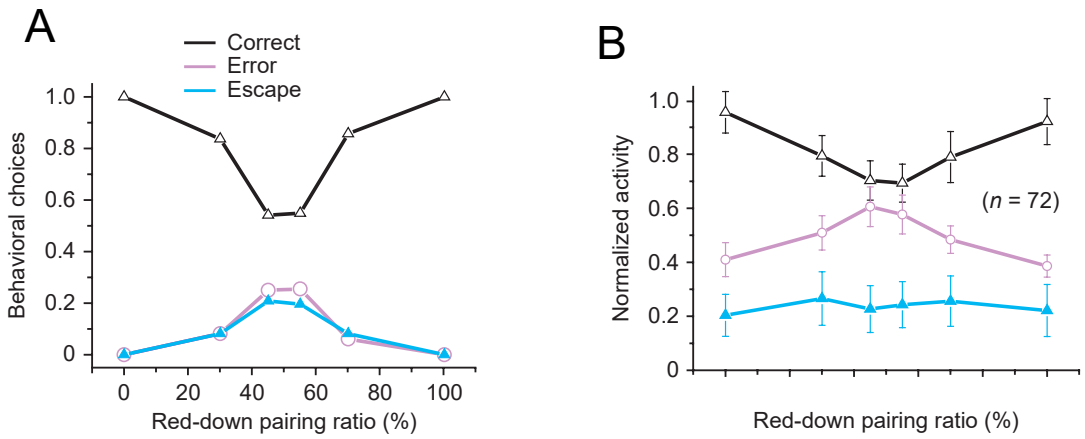


Figure 5.3: **Oup-out task and pulvinar activites (adapted from Komura et al. (2013)) .** (A) Opt-out task performances as a function of the stimulus. (B) Normalized firing rate as a function of the stimulus and the monkeys' choice.

behavioral results and the single-neuron activity data from the experiment in Kiani and Shadlen (2009). Moreover, the probability of choosing the sure target for the network is closely linked to the balance of evidence between the neural activities in this model. More recently, Jaramillo et al. (2019) have proposed a model of pulvino-cortical interactions that can compute the absolute difference between two excitatory populations. They showed that pulvinar responses in the model reflect decision confidence, and that a lesion to the pulvinar leads to an increase number of *escape* trial, as in the experiment of Komura et al. (2013). In the following section I will investigate closely the relation between behavioral expression of confidence and a model of confidence using attractor neural networks.

5.2. Confidence reports and attractor neural networks

Among the different models of perceptual decision-making, attractor neural networks are the ones which are not quantitatively, but only qualitatively, compared to behavioral data. This is due to the complexity of fitting noisy non-linear systems with various parameters. However, this is necessary if one wants to analyze specific behavioral effects and to compare them with diffusion models. In this section, I present a method to fit an attractor neural network on the behavioral task of the previous chapter and I analyze the representation of confidence in the network. The equations of the mean-field version of the network are the ones from Chapter 3.

5.2.1 Fitting an attractor network to behavioral data

When making a decision, the response times between the presentation of the stimulus and the decision can be decomposed into two terms: a decision and a non-decision time. The non-decision time is considered to be due to encoding and motor execution (Luce et al., 1986). The first step to model behavioral data is to model this non-decision time as it is not present in standard decision-making models such as attractor neural networks. Diffusion models assume that the non-decision time is a constant across trials that can be adjusted during the fit of the parameters. This assumption is based on the fact that human studies of decision-making commonly report right-skewed response times (Ratcliff, 1978; Luce et al., 1986), and that the long right tails are well captured by drift-diffusion

models (Ratcliff and Rouder, 1998).

However, with trained subjects, the right-skewed is less pronounced and the response times can be accurately reproduced by a Gaussian distribution (Peirce, 1873). Moreover, experiments in monkeys do not show such long right tails in response times histograms (Ditterich, 2006). Verdonck and Tuerlinckx (2016) proposed a mathematical method to fit a non-parametrical non-decision time distribution. Analyzing various humans experimental data with this method within the framework of drift-diffusion models, they find that strongly right skewed non-decision time distributions are common. These findings suggest that the assumption of a constant non-decision time is quite arbitrary and that, on the contrary, the non-decision times are indeed a right-skewed distribution.

In contrast to diffusion models, when assuming a constant value for the non-decision time, attractor neural network models cannot account for the right-skewed distributions, but accurately reproduce the shape of the distributions in monkeys experiments (Wang, 2008). For the range of parameters I will consider, the decision-time distribution can be approximated by a Gaussian distribution. To estimate the non-decision time distribution in attractor neural networks I propose the following procedure.

As discussed previously, I consider that the non-decision time (NDT) distribution is an exponentially modified Gaussian (EMG) distribution:

$$\rho_{NDT}(t) = \frac{\lambda_{NDT}}{2} \exp\left(\frac{\lambda_{NDT}}{2}(2\mu + \lambda_{NDT}\sigma_{NDT}^2 - 2t)\right) \operatorname{erfc}\left(\frac{\mu_{NDT} + \lambda_{NDT}\sigma_{NDT}^2 - t}{\sqrt{2}\sigma_{NDT}}\right) \tag{5.11}$$

with erfc the complementary error function. The NDT distribution is thus fully described by the three parameters λ_{NDT}, μ_{NDT} and σ_{NDT}. Assuming that there are no correlations between the decision and the non-decision time distribution:

$$\rho_{model}(t) = \rho_{decision}(t) * \rho_{NDT}(t) = \int_0^t \rho_{decision}(t-u)\rho_{NDT}(u)\,du \tag{5.12}$$

with ρ_{model} the response time distribution of the full model and $*$ standing for the convolution operation. Under the assumption that the decision time distribution is a Gaussian distribution, ρ_{model} is an EMG distribution. Taking the characteristic function of the distributions, Equation 5.12 can be rewritten in:

$$\begin{aligned}&\left(1 - \frac{it}{\lambda_{model}}\right)^{-1} \exp\left(i\mu_{model}t - \frac{1}{2}\sigma_{model}^2 t^2\right) = \\ &\exp\left(i\mu_{decision}t - \frac{1}{2}\sigma_{decision}^2 t^2\right)\left(1 - \frac{it}{\lambda_{NDT}}\right)^{-1} \exp\left(i\mu_{NDT}t - \frac{1}{2}\sigma_{NDT}^2 t^2\right)\end{aligned} \tag{5.13}$$

The goal is to fit the behavioral data of the experiment, hence one can identify the different parameters in the previous equation:

$$\lambda_{NDT} = \lambda_{data}, \tag{5.14}$$

$$\langle NDT\rangle = \mu_{NDT} + \frac{1}{\lambda_{NDT}} = \langle RT\rangle_{data} - \langle RT\rangle_{decision} \tag{5.15}$$

and

$$\sigma_{NDT}^2 = \sigma_{data}^2 - \sigma_{decision}^2\,. \tag{5.16}$$

The parameters of the NDT distribution are thus defined using the parameters of the decision time distribution. The calibration of the attractor network is performed separately for each participant and each block (see Appendix 1). The cost function is based on the mean response times and accuracy of the participants for each stimulus difficulty.

The model reproduces faithfully the mean response times and accuracy of the participants across the different blocks (Figure 5.4). Secondly, the values of the parameters obtained for the pure and confidence blocks are different. Participants have higher decision threshold (Signed Rank test Wilcoxon (1945) $p = 0.03$), higher stimulus strength level by angle (Signed Rank test, $p = 0.031$) and higher mean non-decision times (Signed Rank test $p = 0.03$). As mentioned earlier, the fitting procedure allows to estimate the NDT distribution.

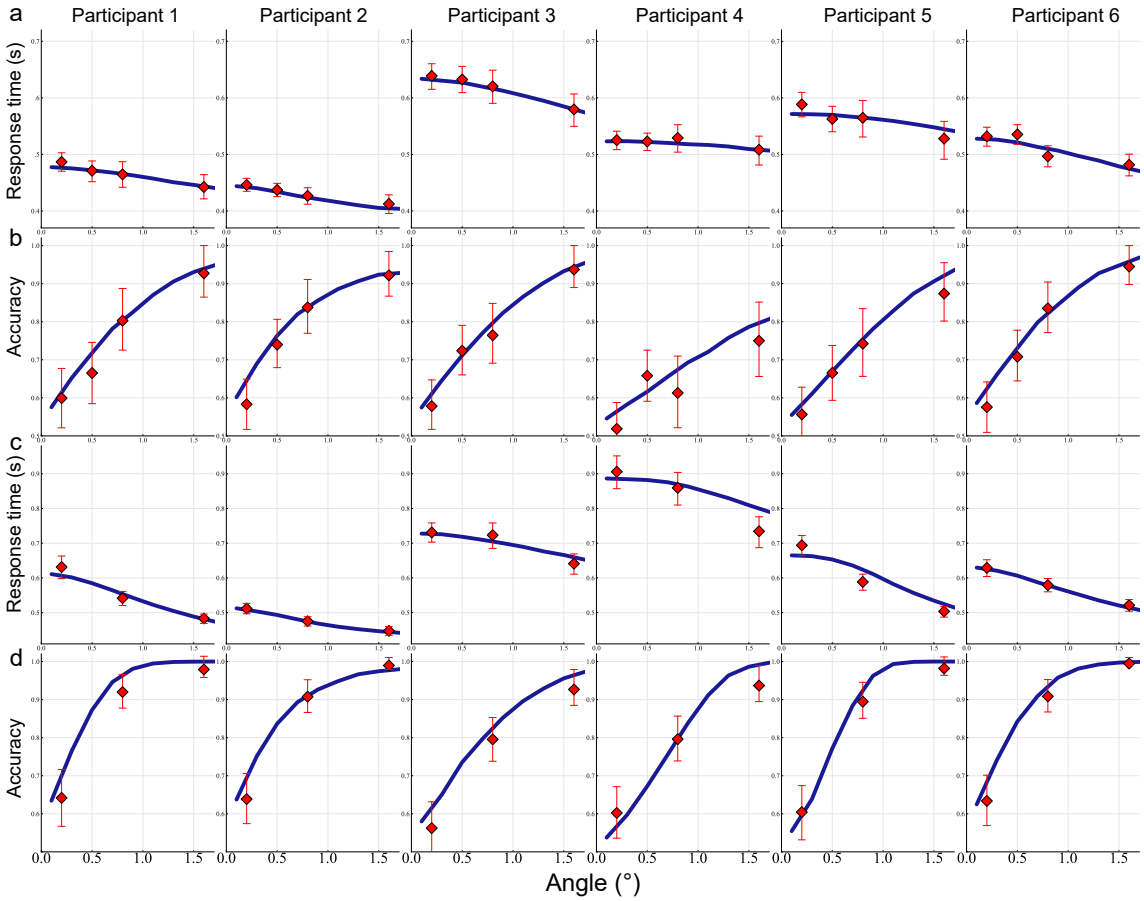


Figure 5.4: **Mean response times (A,C) and accuracies (B,D) as a function of the absolute value of stimulus orientation**, in the *pure* (A and B) and *confidence* (C and D) blocks. For each subject I represent the behavioral data (red dots) and the associated fitted model (blue line). Error bars are 95% confidence interval using the bootstrap method.

In Figure 5.5, I show the histogram of the response times across participants for the pure and confidence blocks. The red curve shows the distribution of non-decision times in the model, and the black curve the response times distribution of the model. One should note that, with a fit only based on the mean response times and accuracies, the model also accurately accounts for the distributions of response times. I find that the minimum value of non-decision time is 75 ms for the pure block, and 100 ms for the confidence block, and the average non-decision times are within the order of magnitude of saccadic latency (Luce et al., 1986; Mazurek et al., 2003a). Finally, the NDT distributions clearly show a right skew for several participants, in agreement with Verdonck and Tuerlinckx (2016). This justifies the modelling of non-decision times with an exponentially modified

Gaussian distribution (EMG), instead of simply adding a constant non-decision time to every decision time.

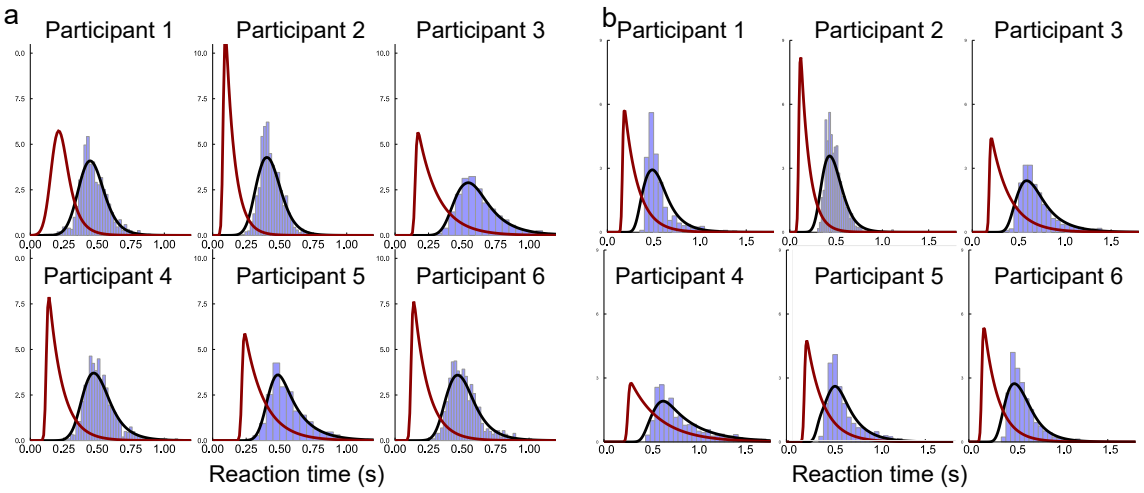


Figure 5.5: **Distributions of RTs for each subject** (A) *Pure* block data, (B) *confidence block* data. For both panels: In blue, participants' histograms of the response times; Black curve: density of response times of the simulated network model; Red curve: the associated non decision response times distribution.

5.2.2 Confidence model

To model confidence within the attractor neural network, I made the hypothesis that the confidence in the decision is based on the difference Δr between the neural activities of the winning and losing neural pools, measured at the time of the decision (balance of evidence) (see Figure 5.6 and Figure 5.7). In the experiment from chapter 4, the measure of confidence is the one reported by the subjects on a discrete scale, and it is this reported confidence level that I want to model. I used the parameters obtained by the fitting procedure of the previous section, and the simulation protocol is similar to the experimental procedure. Within this framework, I quantitatively link this empirical confidence to the neural difference Δr by matching the distribution of the neural evidence balance with the empirical histogram of the confidence levels. This is done by using a procedure called *h*istogram matching (Gonzalez et al., 2002) between the distribution of the neural balance of evidence and the discrete distribution of the reported confidence.

One important point of this analysis is that the shape of the mapping is not chosen a priori but is non-parametrically inferred from the experimental data. This is in contrast with previous studies in which the sigmoidal shape is imposed (Beck et al., 2008a; Kepecs et al., 2008; Kepecs and Mainen, 2012; Wei and Wang, 2015). However, I find that, for each participant, the mapping is well-approximated by a sigmoidal function of the type $1/(1+\exp(-\beta(\Delta r-\kappa)))$, with participants specific parameters κ and β. The similarity of my findings thus suggests that the human reported confidence can be understood as a discretization of a probabilistic function.

Studies have shown that confidence ratings are closely linked to response times (Baranski and Petrusic, 1994; Desender et al., 2018a) and choice accuracy (Peirce and Jastrow, 1884; Baranski and Petrusic, 1994; Sanders et al., 2016; Desender et al., 2018a). Response times decrease and accuracy increases with confidence. In what follows, I study whether the neural balance of evidence can account for the link between the behavioral data: response times, accuracies and confidence reports. Figure 5.8 represents the response times

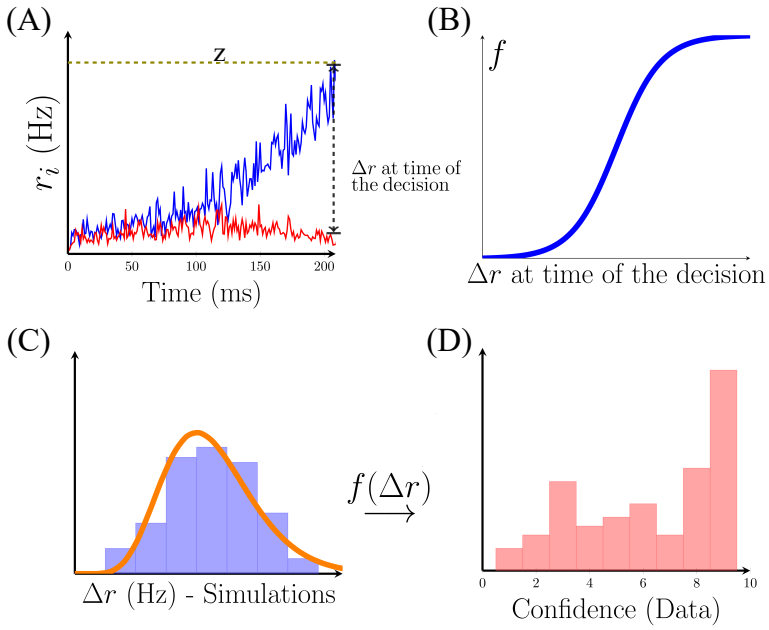


Figure 5.6: **Confidence matching procedure** (A) Example of the dynamics of a decision. (B) Function f that performs the matching between Δr and the confidence (C) Schematic version of the histogram of Δr (D) Histogram of the confidence of the participants.

(Figure 5.8.A) and choice accuracy (Figure 5.8.B) with respect to the reported confidence level for each participant. The data points show the experimental results (with the error bars as the bootstrapped 95 % confidence interval), and the colored line the result of the simulation (with the light colored area the bootstrapped 95 % confidence interval). I find a monotonic dependency between response times and confidence, and between accuracy and confidence, but with specific shapes for each participant. One should note that some values of confidence are only observed for a few trials, resulting then in large error bars especially for accuracy as it consists in the mean of a binary variable. For the numerical simulations, the relatively large size of the confidence interval is due to the limited number of trials, since the simulation protocol is the same as the experimental one (same number of trials). These results show that an attractor neural network can correctly reproduce the psychometric and chronometric functions with respect to confidence for each participant, despite the important difference of response times between participants.

5.2.3 Sequential effects and confidence

In chapter 3, I showed that an attractor neural network, using a simple relaxation dynamics, could reproduce various sequential effects observed in perceptual decision-making experiments such as history biases and post-error slowing. Very recently, the effects of confidence on the history biases have been experimentally investigated (Braun et al., 2018; Samaha et al., 2018; Desender et al., 2018b). One main finding is that decisions with high confidence confer stronger biases upon the following trials. Here, I investigate the influence of confidence upon the next trial in the empirical data, and I will show that the results are well reproduced by the behavior of the dynamical neural model.

First, I performed a statistical analysis of the effect of history biases on response times in the experimental data. To perform this analysis, I transformed the response times of each participant using the z-score (Kreyszig, 1979). This allows us to study all participants

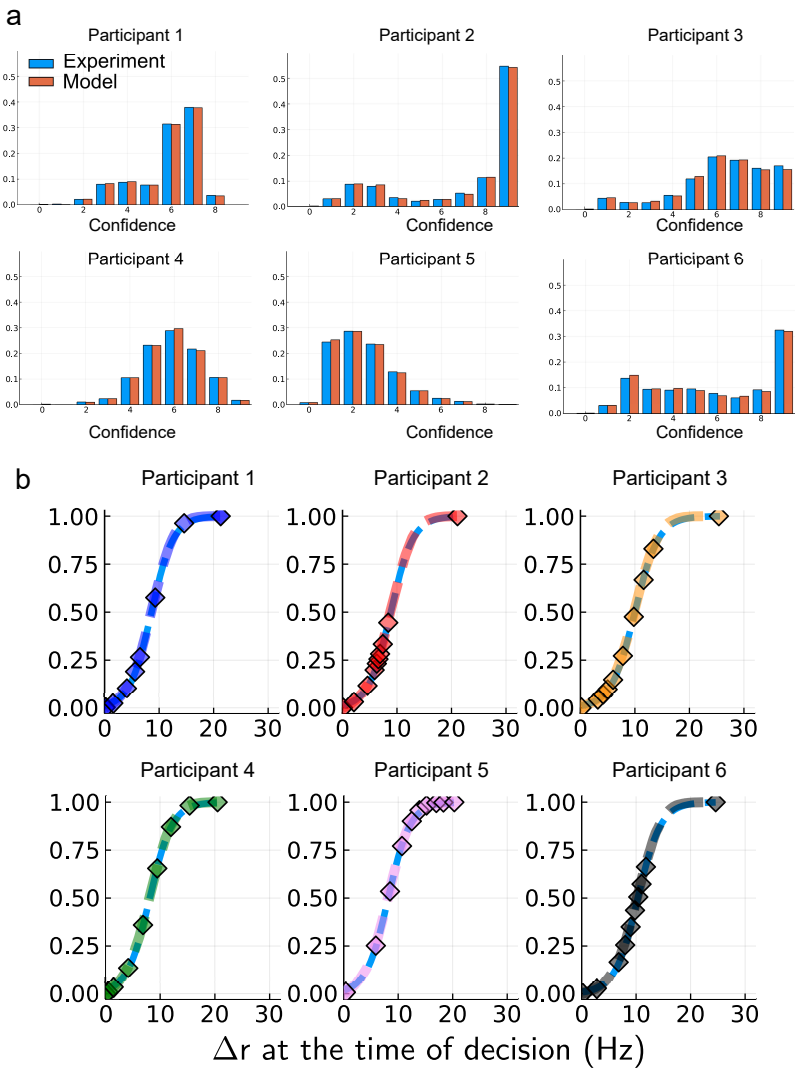


Figure 5.7: **Matching network confidence measure to empirical behavioral confidence.** (A) Confidence histograms. The x-axis gives the value of the confidence on a discrete scale from 0 to 9. Each sub-panel corresponds to a different participant with, in blue, the histogram of the reported confidence, and in orange, the one from the model. For clarity I plot the blue and orange bars side by side, but the bins of the histograms are, by construction, identical. (B) Transfer function F for each participant. The x-axis denotes the difference in neural pools activities Δr at the time of the decision, and the y-axis the cumulative distribution of Δr. Each point represents the levels of Δr delimiting the level of confidence (from left to right, confidence level 0 to confidence level 9). The dashed colored curve is the cumulative distribution function (CDF) and the light blue dashed curve is the fit of the CDF by a sigmoid.

together as the response times are now normalized. I used RStudio (RStudio Team, 2015) with the package *lme4* (Bates et al., 2015) to perform a linear mixed effects analysis Gelman and Hill (2007) of the history biases of the reaction times. The linear mixed effects model (LMM) I consider assumes that the logarithm of the response time at trial n, RT_n, is a linear combination of factors as follows:

$$\ln(RT_n) = a_{0,p} + a_{1,p}|\theta| + a_2 x_{\text{repetition}} + a_{3,p}\ln(RT_{n-1}) + a_4 \text{Conf}_{n-1} \tag{5.17}$$

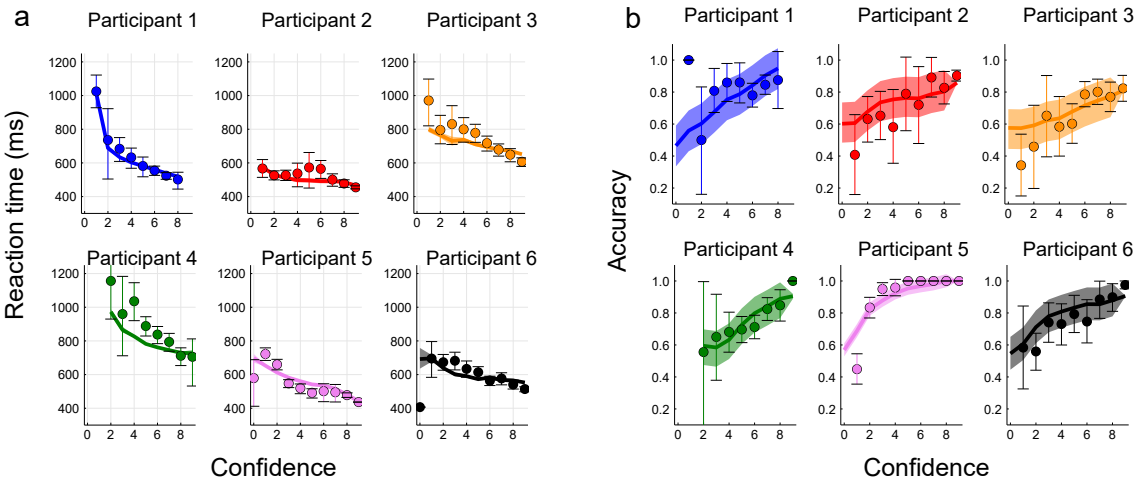


Figure 5.8: **Response times and Accuracy as a function of confidence.** (A) Response times, (B) Accuracy. For both panels: each sub-panel represents a different participant. Dots are experimental data with 95% bootstrapped confidence interval as error bars. Lines are averages over 20 simulations of the attractor neural network model. The shaded area represents the 95% bootstrapped confidence interval on the mean.

with $x_{\text{repetition}}$ a binary variable taking the value 1 if the correct choice for the current trial is a repetition of the previous choice (and 0 otherwise), θ the orientation of the Gabor (in degree), RT_{n-1} the response times of the previous trials, and Conf_{n-1} the confidence of the previous trial coded as 0 for *low* and 1 for *high*. The subscript p in a coefficient (e.g $a_{0,p}$) indicates that for this parameter a random slope per participant is allowed. For each participant, a trial is considered low confidence (resp. high confidence) if the reported confidence is below (resp. above) the participant's median.

To ensure that the chosen model is the preferable one I compared it to other ones that do not include all the combinations of factors. This comparison is done using the *ANOVA* function (with the lme4 package Bates et al. (2015)) that performs model comparison based on the Akaike and Bayesian Information Criteria (AIC and BIC) (Bates et al., 2014). Table 5.1 presents the results of this comparison and shows that our LMM is preferable in all cases.

	Df	AIC	BIC	LogLik.	p value
$a_{0,p} + a_{1,p}\lvert\theta\rvert + a_2 x_{\text{repetition}} + a_{3,p}\ln(RT_{n-1}) + a_4 \text{Conf}_{n-1}$	12	-335	-254	180	
$a_{0,p} + a_{1,p}\lvert\theta\rvert + a_2 x_{\text{repetition}} + a_{3,p}\ln(RT_{n-1})$	11	-324	-249	173	0.0003
$a_{0,p} + a_{1,p}\lvert\theta\rvert + a_2 x_{\text{repetition}}$	7	-4	-42	9	<2e-16
$a_0 + a_1\lvert\theta\rvert + a_2 x_{\text{repetition}} + a_3\ln(RT_{n-1}) + a_4\ln(\text{Conf}_{n-1})$	7	-225	-177	119	<2e-16
a_0	3	-475	-495	-234	<2e-16

Table 5.1: **LMM tests on experimental data, models comparison.** The first row gives the tests for the LMM from Eq. 5.17. The p-values are for the tests based on BIC and AIC (Bates et al., 2014) between the LMM from Eq. 5.17 and the one of the corresponding row.

The results of the analysis of the experimental data are presented in Table 5.2. In line with previous work, higher orientations lead to faster response times and the repetition biases on response times (Cho et al., 2002). Moreover, high confidence has the effect of

speeding up the following trial. Finally, I find that the previous response time has an effect on the subsequent one, meaning that the participants have the tendency to show sequences of fast (or slow) response times.

	Estimate	Std. Error	df	t-value	Pr	
$a_{0,p}$	5.428	1.466e-01	9.0	37.038	$6.92 \cdot 10^{-11}$	***
$a_{1,p}$	-0.1027	0.02390	9.0	-4.296	0.002001	**
a_2	$-3.402 \cdot 10^{-2}$	$2.840 \cdot 10^{-3}$	$8.472 \cdot 10^3$	-11.978	$< 2 \cdot 10^{-16}$	***
$a_{3,p}$	$1.517 \cdot 10^{-1}$	$1.651 \cdot 10^{-2}$	7.0	9.187	$3.23 \cdot 10^{-5}$	***
a_4	$-2.063 \cdot 10^{-2}$	$5.969 \cdot 10^{-3}$	$5.537 \cdot 10^3$	-3.456	0.000553	***

Table 5.2: **Results of the application of the LMM from Eq. 5.17 on the experimental data.** $**$ stands for $p < 0.005$ and $***$ for $p < 0.001$.

The next step is to investigate whether the model can reproduce these various sequential effects or not, even being calibrated only on the mean response times, without the serial dependencies. The results are shown in Table 5.3 and I will summarize them below. The model captures the variation of response times with respect to angle orientation, as expected from Wong and Wang (2006) and observed in the experiment. The repetition bias is correctly reproduced too. Quite remarkably, the model shows an effect of confidence on response times, with a negative slope, as in the experiment.

	Estimate	Std. Error	df	t value	Pr	
$a_{0,p}$	5.999	0.08032	4.229	74.690	$9.22 \cdot 10^{-8}$	***
$a_{1,p}$	-0.01744	$5.551 \cdot 10^{-4}$	2.886	-31.420	$9.47 \cdot 10^{-5}$	***
a_2	-0.1814	$8.133 \cdot 10^{-3}$	$4.822 \cdot 10^3$	-22.301	$< 2 \cdot 10^{-16}$	***
$a_{3,p}$	-0.02075	$1.545 \cdot 10^{-2}$	4.628	-1.343	0.24139	
a_4	-0.02324	$8.336 \cdot 10^{-3}$	$4.847 \cdot 10^3$	-2.788	0.00533	**

Table 5.3: **Results of the application of the LMM from Eq. 5.17 on the data from the neural network simulations.** $**$ stands for $p < 0.005$ and $***$ for $p < 0.001$.

Underlying neural dynamics.

The analysis of the dynamics performed to understand how the neural dynamics leads to these confidence-specific effects is similar to the one done in chapter 3. Figure 5.9 presents the result of this analysis. On each panel, I compare the mean neural dynamics for post-low and post-high confidence trials (respectively red and blue lines). Without loss of generality, one can assume that the previous decision was a C grating (clockwise). The relaxation dynamics between two consecutive trials are different, resulting in different starting points for the next trial, from post-low and post-high confidence trials. Panel (A) corresponds to the case where the new stimulus is also C oriented ("repeated" case), at low strength level. The ending points of the relaxations fall into the correct basin of attraction. Because the post-high confidence relaxation lies deeper into the basin of attraction than the one of post-low trials, the subsequent dynamics will be faster for post-high confidence trials in this case. Panel (B) represents the case, still at low stimulus strength, where the stimulus orientation of the new stimulus is the opposite ("alternated" case) to the one corresponding to the previous decision (hence an AC grating). Both dynamics lie close to the basin boundary of the two attractors, thus the dynamics are slow and there is no significant difference between post-low and post-high confidence trials. In panels (C) and (D) I represent the same situations as panels (A) and (B), respectively, but for high strength

levels (easy trials). The ending points of the relaxations are far from the boundary of the basins of attraction, whatever the grating presented. The response times for post-high and post-low confidence trials are thus similar. This analysis shows that the non-linearity of the network dynamics is responsible for the considered sequential effect. Indeed, in the absence of non-linearity, the repeated and alternated cases would compensate each other and there would be no specific effect related to the basin boundaries.

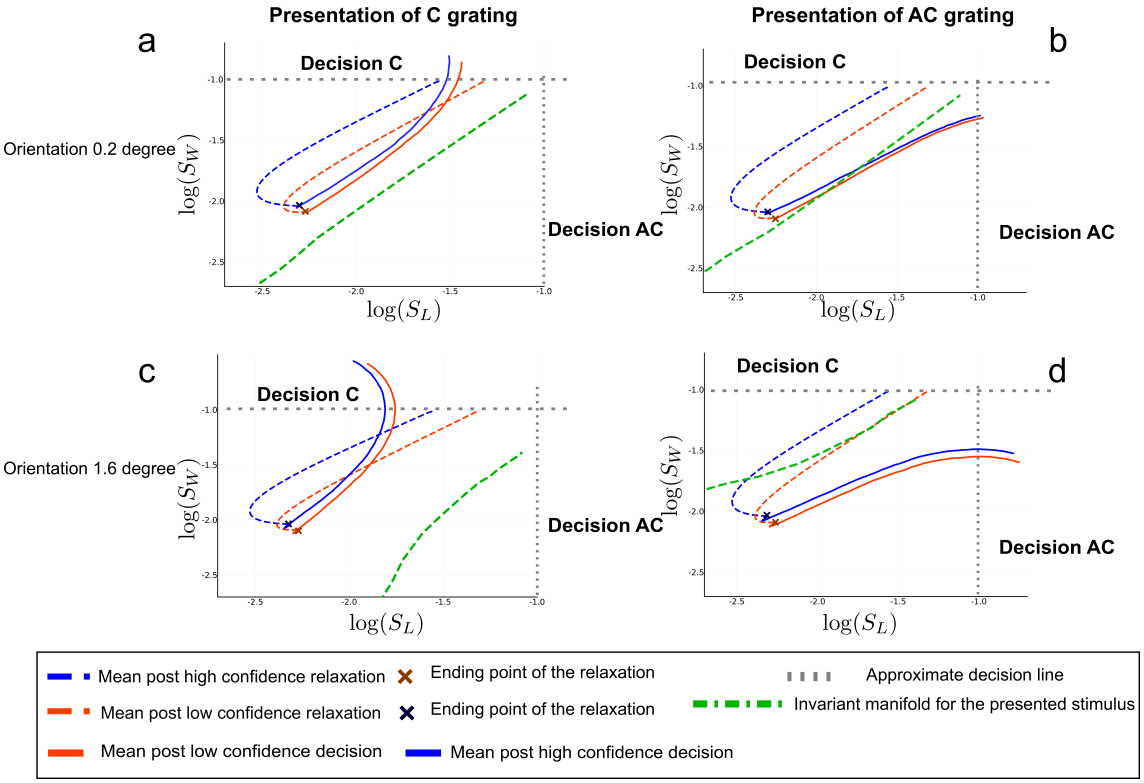


Figure 5.9: **Non linear dynamics in post-low and post-high confidence trials.** Phase-plane trajectories (in log-log plot, for ease of viewing) of the post low and high confidence trials. I assume that the previous decision was decision C. The axes represent the losing neural pool S_L and the winning neural pool S_W at the previous trial. The blue color codes for post-high confidence trials, and the red one for post-low confidence. Panels (A) and (B): Repeated and alternated case for low orientation stimuli; Panels (C) and (D): Repeated and alternated case for high orientation stimuli. In order to compare the decision times, the dynamics starting at the onset of the next stimulus is followed during 200ms, as if there were no decision threshold. The actual decision occurs at the crossing of the dashed gray line, indicating the threshold.

The next step is to compare this analysis of the dynamics of the model with the experimental data. To do so, I regrouped the response times of the experiment into the same categories: high and low stimulus strength, repeated or alterned trials. Within each of these four categories, I compare the post-high and low confidence trials, using a t-test (Fay and Proschan, 2010). The results are the following: mean response times between post-low and high confidence trials are different in the low orientation and repeated case (t-test, $p = 0.044$), but are identical in the three other cases. This is an important point, as the dynamics of the model does not only reproduce the global effect of confidence on the next trial, but also difference between high and low orientations trials.

5.2.4 Attractor neural networks vs. other models

As discussed previously, various models have been proposed in order to model confidence during decision-making. Here, I will compare our model with other models that have been proposed. Previous studies found that, during a perceptual task, reported confidence increases with stimulus strength for correct trials, but decreases for incorrect trials (Kepecs et al., 2008; Sanders et al., 2016; Desender et al., 2018b). This effect is in accordance with a prediction of statistical confidence, defined as the Bayesian posterior probability that the decision-maker is correct (Griffin and Tversky, 1992; Sanders et al., 2016; Navajas et al., 2017). I investigated this effect within my framework. Figure 5.10 represents the mean confidence as a function of stimulus strength, for correct and error trials. One can note that, indeed, the participants exhibit this behavior but that the attractor neural network is able to capture this tendency. Thus, an attractor network model of decision-making reproduces a key feature of statistical confidence.

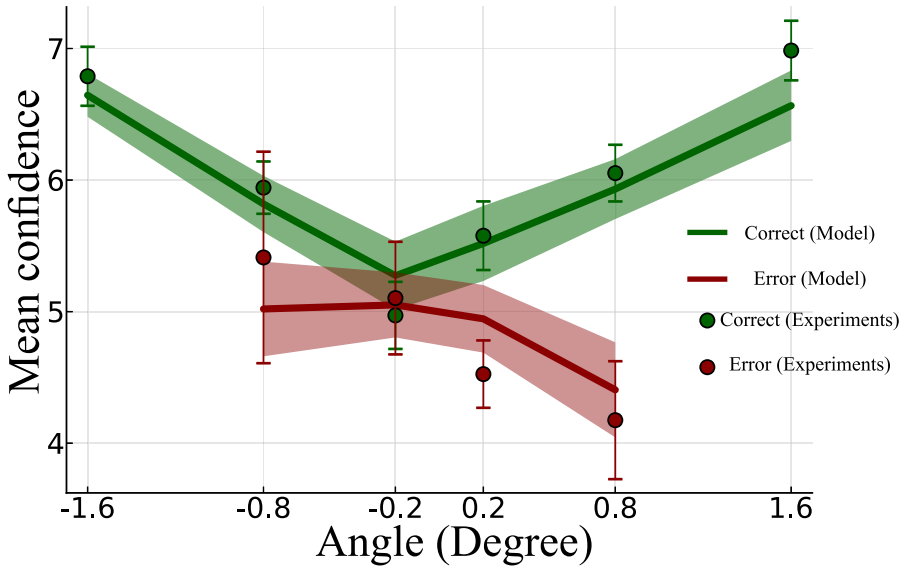


Figure 5.10: **Confidence as a function of stimulus strength.** I represent the mean confidence as a function of stimulus orientation in correct trials (green), and in error trials (red), for the experimental data (points) and the model simulations (lines). With parameters resulting from the fit on the confidence block, the numerical protocol mimics the experimental one (same number of trials, and same angle values). Due to the discrete levels of confidence, and the high performance in the task, I combined the data of all subjects to get enough statistics. The shaded areas (resp. error bars) denote the 95% bootstrapped confidence interval on the mean for the simulation (resp. data)

Another question one could ask is how the attractor network model performs with respect to other dynamical models. To address this question I will consider another non-linear model that has been used to model decision-making, the Usher-McClelland model (Usher and McClelland, 2001). The equations of this model are the following:

$$\tau dx_1 = -kx_1 dt - \beta f(x_2)dt + I_1 + \sigma\mu_1(t) \tag{5.18}$$
$$\tau dx_2 = -kx_2 dt - \beta f(x_1)dt + I_2 + \sigma\mu_2(t) \tag{5.19}$$

with $\mu_i(t)$ a white noise process and I_i the input current to the system. The external input is defined as $I_i = 0.5 \pm c_\theta$, with c_θ the strength per angle as in the attractor neural network.

$\sigma = 0.4$ denotes the strength of the noise, k the relaxation strength, $\tau = 0.1$ the relaxation time and β the inhibitory term. Finally, the function f is a sigmoidal function of gain $G = 0.4$ and half-activity offset $d = 0.5$, $f(x_i) = 1/[1 + exp(-G(x_i - d)]$. The dynamics occurs until a threshold z is reached for one of the two units. It should be noted that, despite the non-linearity, the Usher-McClelland model is closer to drift-diffusion models than to biophysical attractor models. This is principally due to the fact that the only non-linearity is in the interaction between both units. Thus, reductions to one-dimensional drift diffusion models can be made in various ranges of parameters (Bogacz et al., 2006).

I fit this model to the experimental data using the same procedure as for the attractor neural network (see Appendix A for more details). The resulting parameters are the following:

The next step is to define confidence in a similar way as with the attractor model using the balance of evidence and the histogram matching procedure.

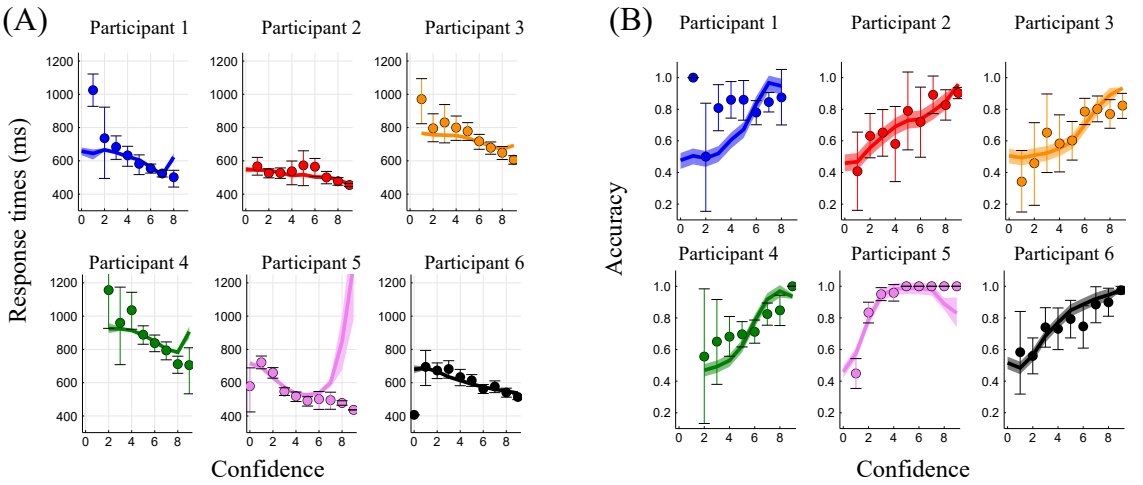


Figure 5.11: **Mean response times (A,C) and accuracies (B,D) as a function of the absolute value of stimulus orientation for the Usher-McClelland model**, in the *pure* (A and B) and *confidence* (C and D) blocks. For each subject I represent the behavioral data (red dots) and the associated fitted Usher-McClelland model (blue line). Error bars are 95% confidence interval using the bootstrap method.

Figure 5.11 presents the relation between reaction times, accuracies and confidence for this model. At first sight the fit between numerical simulations and experimental data looks accurate. However, it is only true for intermediate value of confidence. For some of the participants (Participant 1, 4 and 5), there is a strong divergence of the model at high confidence. This can be understood by the fact that, in this model, *firing rate* variables can take negative values (the steady state corresponds to a symmetric state with negative values). This leads to extreme values of confidence for long trials. Moreover, for some participants (such as 1 and 4), the trend in accuracy, despite being always increasing, is not correct. This highlights the fact that, even with the same model for confidence, discrepancies between models of decision-making exist. It is then possible to distinguish different models by comparing them on different aspects of decision-making.

Finally, the last model I will compare to the attractor model is the independent race model (IRM). I chose this model because it is possible to define confidence using the balance of evidence. Such models have been successfully used to model decision-making experiments. I will investigate the notion of sequential effects with the IRM. As already mentioned in a previous chapter, when studying sequential effects with models such as the

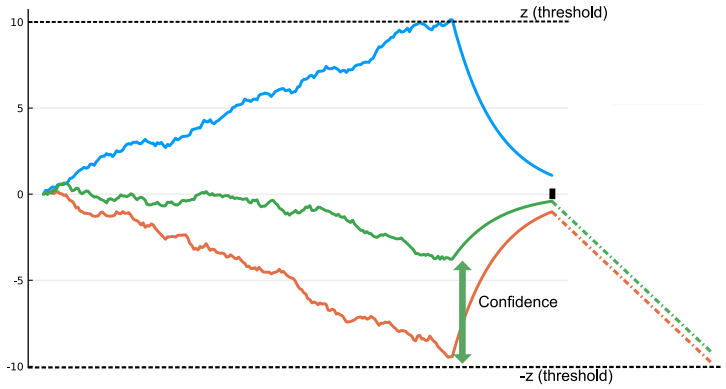


Figure 5.12: **Schematic dynamics of a race model with a relaxation mechanism.** The upper and bottom dash lines correspond to the two opposite decision thresholds. The blue trajectory is a typical winning race. The black rectangle on the x-axis denotes the onset of the next stimulus, hence the end of the relaxation period. The green and orange trajectories are the loosing races in two trials with different confidence outcomes. The green and orange dashed lines represent the mean dynamics of these two races during the presentation of the next stimulus.

IRM, the parameters are allowed to change between each condition. This characteristic does not seem to be biophysically plausible and, to compare with the attractor network, it is necessary to extend the IRM with a relaxation dynamics. After a decision is made, both units receive a non-specific inhibitory input leading to a relaxation until the next stimulus is presented (Figure 5.12). Within this extended IRM framework, one can study how the sequential effects would be correlated with confidence in an IRM model with a fixed set of parameters.

Since in the IRM there is no interaction between the two races, the relaxation of the winning race is the same in both low and high confidence trials. However, the ending point of the relaxation following a decision is closer to the base-line (0 line) for a high confidence trial than when it comes to a trial with low confidence trial (Figureé5.12). For the next trial, if the winning race is the same as previously, then the mean response times are identical in low and high confidence cases. However, if the opposite decision is made, the response time in the post-low confidence case is faster than the one in the post-high confidence case (Figure 5.12).This behavior is in contradiction with the experimental data for which the opposite effect is observed. This conclusion applies more generally to any race-type model without interactions between units. It highlights the fact that the non-linearity of the attractor neural networks is a key property in order to reproduce the sequential effects due to confidence in perceptual decision-making.

II Neural coding

In this part I describe the work I have performed during the second half of my PhD. It concerns the initial project of my PhD that was the study of a dynamical network that would perform both part of a categorization task: the coding of the stimulus and the decoding in order to perform the decision. Categorization tasks have many perceptual effects that are observed at different levels: from psychophysics to neurophysics.

My work has focused on understanding the effect of the neural coding on the learning process of the categorization task, when the decision layer is composed of an attractor neural network. I show that a classic Hebbian learning does not allow the network to learn efficiently the categorization task. I make use of the previous part to study the effect that confidence can have on the learning process. I show that a modulation of the Hebbian learning by the confidence allows the network to learn efficiently the categorization task.

The different results that I present in this part will be published in an article *currently in preparation:*

- Berlemont, K. and Nadal, J. P. (2020). Confidence-modulated Hebbian learning efficiently extracts category membership from stimuli encoded in view of a categorization task.

6

An overview of categorical perception

6.1. Categorical perception

Our perception of the world is divided in many different categories. One typical example is the perception of colors. Colors differ only in their wavelengths, which are continuously distributed among the spectrum of visible colors. However, we perceive the changes in colors gradually, from red to yellow to green for example. This is an example of how perceptual boundaries have arisen on a continuum to divide it into discrete categories. In this chapter I review the different perceptual findings that have been observed in categorical perception, and the different models that have been proposed to mimic them.

6.1.1 Categorical perception

The effect of categorization on perception can be studied thanks to psychophysics experiment. The phenomenon is called *categorical perception* (CP) and denotes the fact that differences between objects that belong to different categories are exaggerated, and objects within the same category look more similar. These differences can be quantitatively tested experimentally by comparing the performances in the discrimination and the categorization task. The discrimination task requires to tell apart stimuli presented in pairs (if they are different or not), and the categorization task to tell if the two stimuli belong to the same category.

Categorical perception has been observed in many different frameworks such as in color perception (Roberson and Davidoff, 2000), phonemic perception in speech (Liberman et al., 1957), perception of facial expressions (Etcoff and Magee, 1992). Figure 6.1 shows an exemple of phoneme discrimination (Liberman et al., 1957). Panel A represents the identification between different phonemes boundaries such as /be/ and /de/. The person's perception varies rapidly as the stimulus crosses the boundary between the categories. The perceiver's ability to discriminate between two sounds peaks near the boundary separating phonemic categories (Panel B).

6.1.2 Neuronal evidence

Despite the various psychophysics studies on categorical perception, very few studies focus on the underlying neuronal mechanisms (Sigala et al., 2002). During the training of a visual categorization task, improvements of performances are accompanied by a small change of tuning in the early visual cortex (Yang and Maunsell, 2004). However, the changes in inferior temporal and posterior parietal cortex are more significant. Neurons in the LIP learn to represent the category membership of the stimuli, almost in a binary manner (Freedman et al., 2001; Freedman and Assad, 2006). Moreover, this encoding can shift when the monkeys learn to associate new categories to the stimuli. In constrast, neurons in the MT

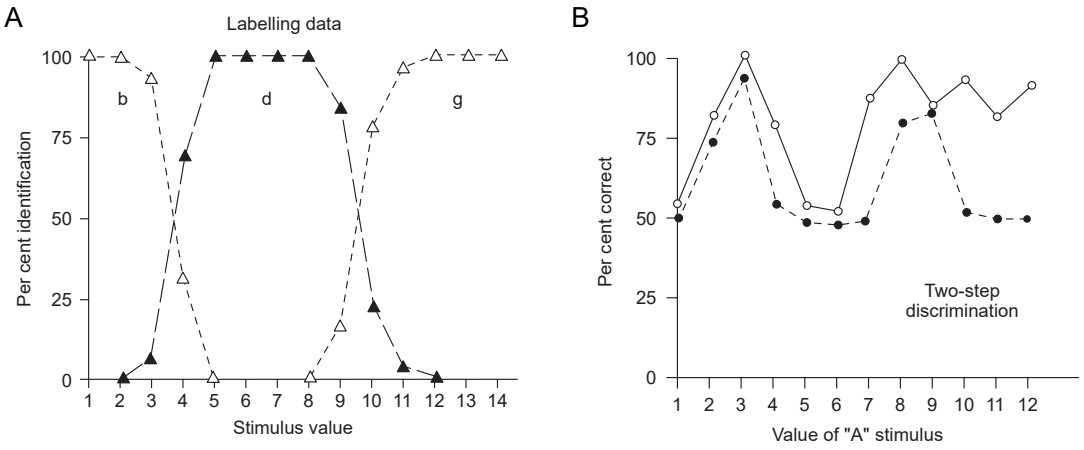


Figure 6.1: Discrimination between phonemes in a discrimination task with human participants. Figure adapted from Liberman et al. (1957).

area do not present this selectivity with respect to the categories (Freedman and Assad, 2006).

Very recently, researchers have investigated the stimulus categorization task in rodents (Xin et al., 2019). They study the reshaping of the tuning maps of neurons in the auditory cortex during the categorization task and the passive condition. They found that neurons in the auditory cortex are modulated by the task. They exhibit dynamic reorganization of population response, with neurons exhibiting emergent selectivity to stimuli near the category boundary. These neurophysiological experiments highlight the fact that learning a categorization task not only modifies the perceptual response but the neuronal response too.

6.2. Population coding

Population coding has been deeply studied, principaly in the context where the task is to infer a stimulus from the neural code. As described in the previous section, categorization has many effects on neural coding and on the behavior. However, very few studies focused on population coding in the context of a categorization task. Here, I will present the different results and hypotheses that have been proposed.

6.2.1 Fisher information

Neural coding of categories can be studied in the context of the information theory framework (Cover and Thomas, 2012). The goal is to code a discrete set of categories, $\mu = 1, \cdots M$. To do so, a layer of N neurons receives a stimulus $x \in \mathcal{R}^K$ and produces a response $r = (r_1, \cdots, r_N)$. Each neuron is a Poisson neuron characterized by a tuning curve under the form of a bell-shaped function (Tolhurst et al., 1983):

$$f_i(x) = \exp\left(-(x-\mu_i)^2 / \left(2\sigma_i^2\right)\right) \tag{6.1}$$

with μ_i the center of the tuning curve and σ_i its width. The number of spikes r_i that neuron i emits during an interval τ is:

$$P_i(r_i|x) = \frac{f_i(x)^{r_i}}{r_i!}\exp(-f_i(x)) \tag{6.2}$$

Under the assumption that $N >> 1$ and that the input dimension $K << N$, one has (Bonnasse-Gahot and Nadal, 2008):

$$I(\mu, x) - I(\mu, r) = \frac{1}{2} \int dx\, p(x) \frac{F_{\text{cat}}(x)}{F_{\text{code}}} \tag{6.3}$$

with μ the category membership, x the stimulus,r the neural response of the coding layer and I the mutual information. The Fisher information that compose the formula are the following:

$$F_{code}(x) = -\int d^N r P(r|x) \frac{\partial^2 \log P(r|x)}{\partial x^2} \tag{6.4}$$

$$F_{cat}(x) = -\sum_{\mu} P(\mu|x) \frac{\partial^2 \log P(\mu|x)}{\partial x^2} \tag{6.5}$$

F_{code} denotes the Fisher information characterizing the sensitivity of r with respect to x, and F_{cat} the Fisher information characterizing the sensitivity of μ with respect to x.

The optimal code will depend on F_{cat}. In the case of smoothly overlapping categories, F_{cat} is zero in domains where $P(\mu|x)$ is flat. Therefore, the optimization will lead to neurons that have sharp tuning curves at the boundary.

Now I present the analytical computations to obtain the distributions of tuning curves that I have performed during my PhD. I minimize the quantity of Equation 6.3 using the Lagrange multiplier method. I make the assumption that all the tuning curves correctly tile the space which leads to a dependency of the Fisher information of the neural code Ganguli and Simoncelli (2014) in d^2, with d the density of tuning curves. The loss function is the following:

$$\begin{aligned} \mathcal{L} &= \frac{1}{2} \int p(s) \frac{F_{cat}(s)}{F_{code}(s)} ds + \lambda \left(\int d(s) ds - N \right) \\ &= \frac{1}{2} \int p(s) \frac{F_{cat}(s)}{d^2(s)} ds + \lambda \left(\int d(s) ds - N \right) \end{aligned} \tag{6.6}$$

with λ the parameter corresponding to the constraint of a fixed number of neurons. Taking the derivative with respect to $d(s)$, I obtain that:

$$d(s) \propto (p(s) F_{\text{cat}}(s))^{1/3} \tag{6.7}$$

The width of the tuning curves is obtained with $1/d(s)$. From this formula one can first observe that the density of tunig curves follows F_{cat}. As already mentionned, more tuning curves with sharper slopes are present near the categories boundary. Moreover, the distribution of stimulus $p(s)$ has the same effect as F_{cat}. For highly non-uniform stimuli distribution, the density of tuning curves will be shaped in order to represent this non-uniformity and not just the behavior of F_{cat}. Finally, it is worth noting that the equation on $d(s)$ can be obtained for different normalization constraints. For example, if the maximal values of the tuning curves is modulated by a function $g(s)$ with a constraint on $\int p(s) g^k(s) ds$, one can obtain the following result:

$$g(s) \propto (p(s) F_{cat}(s))^{1/(3k+1)} \tag{6.8}$$

$$d(s) \propto (p(s) F_{\text{cat}}(s))^{\frac{k}{3k+1}} \tag{6.9}$$

One can note that, as $k \to \infty$, the result of the optimization is similar to the one ithout the parameter $g(s)$.

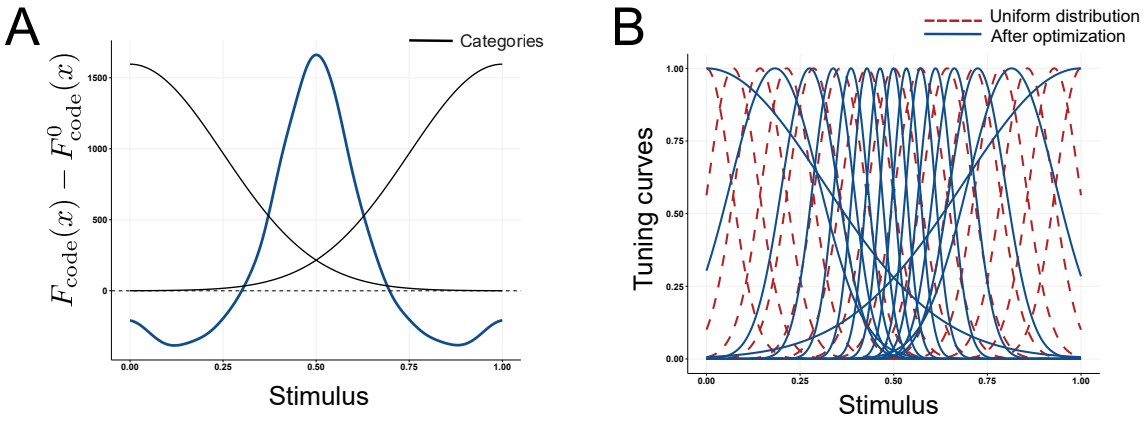


Figure 6.2: Tuning curves in a coding layer. (A): Results of the optimization of the coding layer with respect to Fisher information. In black I show the categories used in this example, and in blue $\Delta F_{code} = F_{code} - F_{code}(\text{uniform})$ the difference of Fisher information between the optimized layer and the uniform distribution of tuning curves. (B): Example of the tuning curves of a coding layer of 15 neurons in the case of two gaussian categories of width 0.25 and centered in 0 and 1 respectively. The curves in red correspond to the tuning curves in the case of an uniform distribution, and in blue after optimization of the distribution of tuning curves.

6.2.2 Top-down feedback

The other class of models that analyze the neural coding of categories makes use of a top-down feedback (see Figure 6.3) (Tajima et al., 2016, 2017). The input is sent to the hue selective neurons that project to the category selective neurons. These neurons send a feedback projection to the hue selective neurons. This leads to a modulation of the activity of the hue selective neurons towards the center of the different categories. The model has shown to replicate clustering of population toward categorical centers, an effect that has been observed in experimental studies (Brouwer and Heeger, 2013) (Figure 6.3.B).

However, in the model, there are two main problematics. First, the connections between the hue selective neurons and the category neurons are imposed. It is not clear how the feedforward and the feedback connections could be learned due to the strong interaction between the layers. Second, the different frameworks that have been proposed with this architecture have all one thing in common: there is no noise in the system. This can not be adressed by just adding noise within the system. For instance, in the model of Tajima et al. (2017) the category neurons are modelled by an attractor neural network. However, if there was noise in the dynamics, the top-down feedback would bias the decision towards the wrong decision if the noise was in this direction. Therefore, it is not clear how to characterize the effect of top-down feedback in a dynamical model that would characterize the neural coding of categories.

In the next chapter, I adress the question of how a network can learn a categorization task and the different effects that are observed.

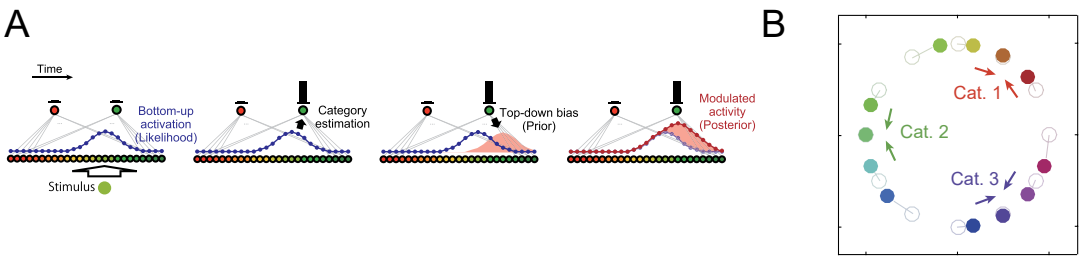


Figure 6.3: **Top-down feedback model (adapted from Tajima et al. (2016))**. (A): Cycle of the modulated activity of the network. The bottom line of neurons represents the hue selective neurons and the upper neurons the category selective neurons. (B): The colored dots indicate the stimulus properties represented by simulated neural population activities, by "decoding" the population activity. Light-colored dots indicate the input stimuli while the dark-colored dots represent the neural population representation. The colors of markers correspond to those of presented stimuli. There were three categories whose centers were in direction of $-\pi$, $-\pi/3$, π radians.

7

Learning a categorization task: the impact of confidence

7.1. Coding and Decoding during a categorization task

Task-specific neural representations develop across different cortical areas (Cromer et al., 2010; Fitzgerald et al., 2011) during the learning of a perceptual task. These representations are accompanied by a modification in the tuning properties of the neurons. However, most models of decision-making consider a uniform coding of the stimulus before the decision part (Drugowitsch et al., 2019; Beck et al., 2008b). In this section, I review two different frameworks that have studied a model that encodes the stimulus and performs the categorization task.

7.1.1 Probabilistic framework

Bonnasse-Gahot and Nadal (2012) consider a probabilistic framework that is based on the population coding that I described in Chapter 6. The model architecture is shown in Figure 7.1.A and consists in a feedforward neural architecture. A population coding layer encodes the stimulus and a read-out provides an estimate of the posterior probabilities. In a second step, the decision process occurs as a diffusion model is applied to the output of the network. This approach is different from the standard diffusion model that I described in the Introduction. Indeed, this framework allows to study the effect of neural coding on behavioral measures. Especially, it links an optimal population coding to quantities that are measurable in a categorization task.

The two main results are the following:

- One can derive a relationship between the optimal decoding (Bayesian point of view) and the encoding efficiency under the form of the mutual information between the neural activity and the categories.
- Due to the optimized population coding, the variance of the input of the diffusion model is stimulus dependent. The model reproduces the mean reaction time of the psycholinguistic experiment of Ylinen et al. (2005) and explains the difference between a group of participants for which one can assume an optimized coding and one for which this is not the case (Figure 7.1.B).

7.1.2 Recurrent neural network framework

Recently, Engel et al. (2015) have investigated the learning of categories within a recurrent neural network architecture in the context of a direction categorization task. The network architecture is described in Figure 7.2.A. Initially, both neural populations (association and sensory neurons) are described by neurons whose preferred stimuli are equally spaced within the possible directions. Therefore, at the beginning, the coding that is performed by the network consists in a uniform one. The synaptic connections that go from the

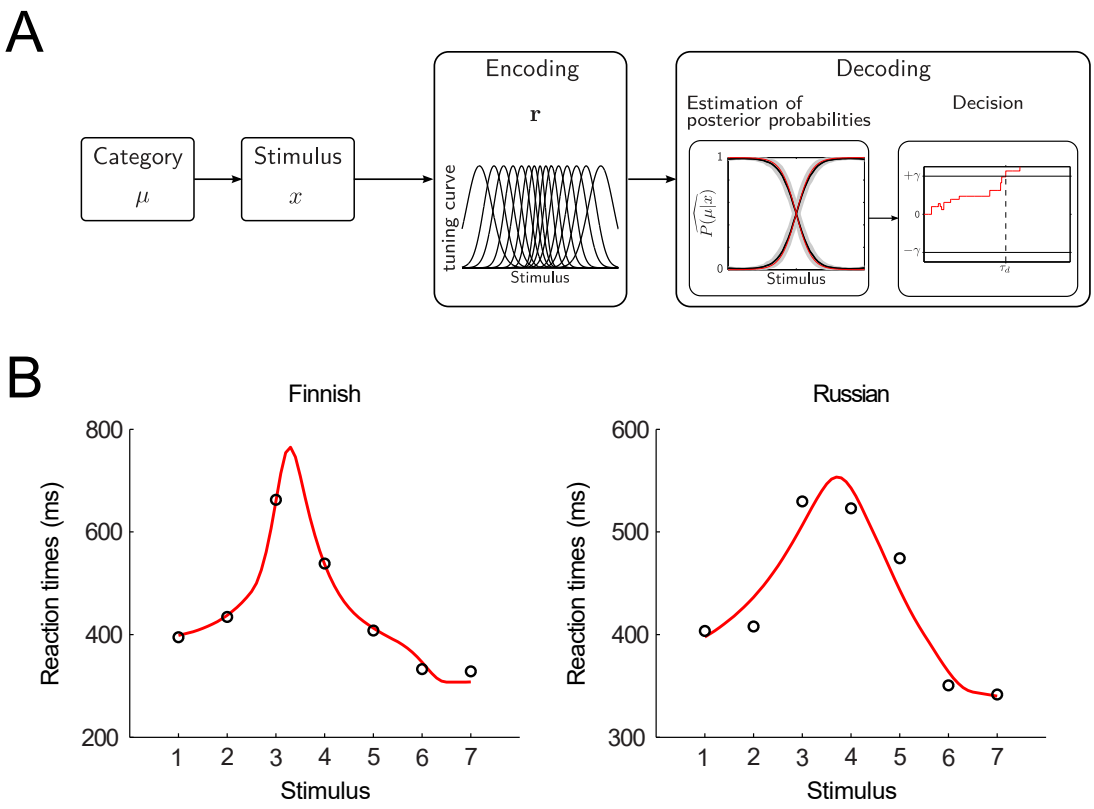


Figure 7.1: **Figure adapted from Bonnasse-Gahot and Nadal (2012).** (A): Model architecture. (B): Mean reaction times as a function of the stimulus presented, for the two groups of subjects. The open circles indicate the mean reaction times obtained in the experiment for each stimulus (Ylinen et al., 2005), whereas the red line corresponds to the model prediction. On the left, the data corresponding to the native speakers of Finnish, on the right, those corresponding to the Russian group.

sensory neurons to the association neurons and from the association neurons to the decision neurons are plastic. The plasticity consists in a reward-modulated Hebbian learning through top-down modulation (Schultz et al., 1997; Frémaux et al., 2010; Loewenstein and Seung, 2006).

During learning, a transformation of tuning in association neurons occurs and category selectivity emerges. This change is driven by the plasticity between sensory and association neurons. For the neurons that initially preferred directions near category centers, tuning curves broaden. For neurons that initially preferred directions near categories boundaries, tunig curves shift towards centers of categories (see Figure 7.2.B). Both behaviors lead to slope of tuning curves that are sharper at the boundary between categories, which is characteristic of an optimization of Fischer information (see Chapter 6).

These two frameworks have shown that specific effects arise when the coding of a stimulus is taken into account during the learning of a categorization task. In the next section I will study the effect of having an optimized coding layer within an attractor neural network framework.

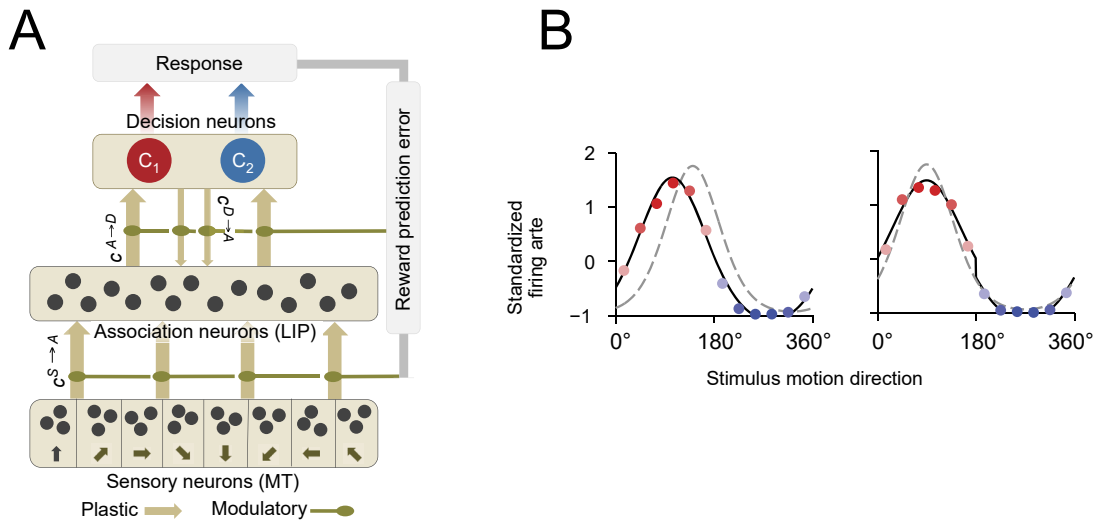


Figure 7.2: **Figure adapted from Engel et al. (2015).** (A): Model architecture. The network comprises a sensory (MT), an association (LIP) and a decision neural circuits. (B): Tuning profiles of two example association neurons before (grey dashed line) and after learning (coloured dots—firing rates, black solid line—best-fitted tuning function).

7.2. Confidence-modulated Hebbian learning

When the environment is changing and uncertain, learning is accompanied by a sense of confidence about the different predictions (Nassar et al., 2010). This sense of confidence plays a functional role in learning (Nassar et al., 2010; Meyniel and Dehaene, 2017) as it sets the balance between predictions and new information. Many studies reported the existence of surprise signals in the brain, i.e a strong signal in presence of unexpected stimulus (Hillyard et al., 1971; Summerfield and De Lange, 2014). More recently, it has been shown, using fMRI (Meyniel and Dehaene, 2017), EEG (Jepma et al., 2016; Nassar et al., 2019) or MEG (Meyniel, 2019) that this surprise signal is modulated by confidence. This suggesst that a confidence-weighting in the brain occurs in the brain in order to calibrate the response in a complex environment. This mechanism could be crucial in order to implement adjustable learning rate in the brain (Meyniel and Dehaene, 2017), as it has been shown that human learns with adjustable learning rate (Behrens et al., 2007).

From a theoretical point of view, the influence of having hard bounds on the learning rate has been studied in the framework of storing memories within recurrent neural networks (Alemi et al., 2015). The authors showed that having a modulation of the learning process by the distance to a specific decision bound allows to obtain a storage capacity close to the one predicted by analytical computation. Recently, it has been shown (Drugowitsch et al., 2019), in the context of Bayesian learning, that the optimal learning rate for categorization tasks should depend on the confidence in one's decision. Here, I adress the effect of confidence on the learning process of a categorization task in an attractor neural network framework.

7.2.1 Two-layers model

Architecture

The neural circuit model is composed of two layers: the coding layer and the decision layer (Figure 7.3.A). The coding layer is simulated by N Poissons neurons whose firing rates are

described by a bell-shaped tuning curve (Tolhurst et al., 1983):

$$f_i(x) = \exp\left(-(x-\mu_i)^2 / \left(2\sigma_i^2\right)\right) \tag{7.1}$$

with μ_i the centers of the tuning curves, x the stimulus presented and σ_i^2 the variance. The coding layer can either be uniform or optimized with respect to the categorization task (Chapter 6). The decision layer is composed of two populations representing the categorical choice. It consists in the same attractor network as I have presented through this manuscript. The decision layer pools the activity from the coding layer through synaptic connections that undergo plasticity across trials.

Each unit of the decision network is described by the equations in Chapter 3. Unless otherwise stated, the parameters and equations are the ones of Chapter 3.

$$i \in \{1,2\}, \ \frac{\mathrm{d}S_i}{\mathrm{d}t} = -\frac{S_i}{\tau_S} + (1-S_i)\gamma f\left(I_{i,tot}\right) \tag{7.2}$$

with S_i the synaptic activities of the populations. The current $I_{i,tot}$ that arrives on each population is now defined by:

$$I_{i,tot} = J_{A,ext}\mu_0 \sum_j W_j^i r_j^C + J_{i,k}S_k + J_{i,i}S_i + I_{noise,i} \tag{7.3}$$

with r_j^C the number of spikes emitted during the interval dt by neuron j of the coding layer and $J_{i,k}S_k + J_{i,i}S_i$ the coupling term between the units in the decision network (see Chapter 3). The term W_j^i denotes the synaptic strenght of the connection between neuron j of the coding layer and population i of the decision network. The noisy input of population i is defined by:

$$\frac{dI_{i,tot}}{dt} = -(I_{noise,i}(t) - I_0) + \mu_i(t)\sqrt{\tau_{noise}}\sigma_{noise} \tag{7.4}$$

with $I_0 = 0.3411$ nA and $\mu_i(t)$ a white-noise process.

Synaptic plasticity

At the end of each trial, the strength of the synaptic connection between a neuron i of the coding layer and the winning population (C_j) of the trial (W_i^j) is updated as:

$$W_i^j \leftarrow W_i^j + q\lambda R r_i \tag{7.5}$$

with $\lambda = 0.005$ the learning rate, R the reward of the trial (1 if the decision was correct, -1 otherwise), r_i the firing rate of the presynaptic neuron at the time of the decision and q a parameter modulated by the confidence of the trial.

Confidence is modelled as a function of the difference between the neural activity of the two populations of the decision layer (Δr), as in Chapter 4. A trial corresponds to a high confidence (resp. low) trial when the difference in activity at the time of the decision is greater (resp. lower) than a threshold θ_c.

To prevent the divergence of the learning algorithm, I include a synaptic normalization mechanism after the update of the weights (Oja's rule) (Oja, 1982):

$$W_i^j \leftarrow \frac{W_i^j}{\sqrt{\sum_j (W_i^j)^2}} \tag{7.6}$$

Otherwise the reward modulation would lead to a divergence of the weights. With this normalization, I ensure that the length of the weight vector remains constant. I have studied

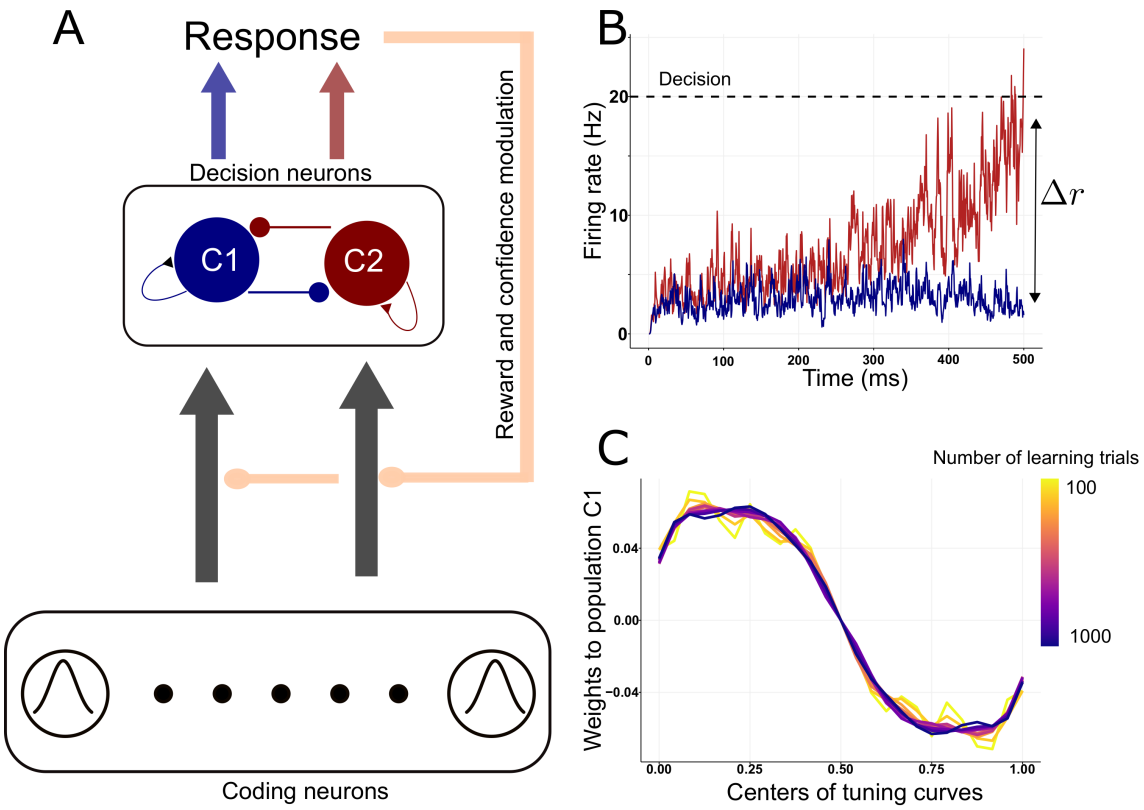


Figure 7.3: **Neural circuit model.** (A): Schematic of the circuit model. The network is composed of two layers. Neurons in the coding layer are tuned to specific values of the 1-d stimulus. They receive direct input corresponding to the stimulus. The decision layer pools activity of the coding neurons through feedforward synapses. The category decision is obtained through competitive attractor dynamics. The synaptic connections between the two layers undergo Hebbian plasticity modulated by a reward signal. (B): Example of the dynamics of the decision layer when the network is presented with a stimulus. Each color represents the activity of the corresponding population. The decision is reached when one of the population reaches the decision line (threshold z). Δr is computed as the difference of firing rates between the two populations at the time of the decision. (C): Synaptic weights going from the coding layer to the neural population C_1. Each color represents a different epoch in the learning process, starting at 100 trials. The coding layer is composed of 20 neurons, uniformly distributed and the width of the gaussian categories is 0.25. After 1000 trials, the network is already at steady-state.

other type of normalization such as a dynamical relaxation of the weights across trials. The results were similar, thus I decided to focus on the above normalization mechanism (Miller and MacKay, 1994). The synapses are initialized from an uniform distribution between $[-1,1]$.

Simulation protocol

Each simulation trial consists in the presentation of a randomly chosen stimulus among two categories and a 1 s intertrial interval. The categories are Gausian categories centered in 0 and 1, of width α. The stimulus belongs to $[0,1]$. During the presentation of the stimulus, the neurons in the coding layer fire a Poisson train (Equation 6.2). The trial lasts until one of the population of the decision circuit reaches a threshold z of 20 Hz. The choice made by the network on this trial corresponds to the one associated to the winning population. Once a decision has been reached, there is no more input from the coding neurons to the decision neurons and a brief current (corollary discharge) (Engel et al., 2015; Berlemont et al., 2020; Berlemont and Nadal, 2019) is applied to the decision neurons. This current resets the activity of the decision layer at the spontaneous level. The corollary discharge is of the form:

$$I_{CD} = I_{max} * \exp(-t/\tau_{CD}) \tag{7.7}$$

with $\tau_{CD} = 200$ ms and $I_{max} = 0.05$ nA. The performance of the network for a specific input is computed as the average over 2000 trials.

7.2.2 Learning the categorization task

Without modulation by confidence

First, I will study the case where confidence does not modulate learning rate ($\theta_c > z$). Figure 7.3.C represents the weights of the coding neurons towards the neural population C_1 during learning. The dynamics of the weights stabilize quickly, as 1000 trials seem to be enough. Figure 7.4.A shows the performances achieved by the network for different coding layers after 2000 trials. First, the network achieves learning as the synaptic connections were initialized with random values leading to performance at chance level. Two behaviors can be observed on this panel. When the distribution of the tuning curves is uniform, the performances do not seem to vary with N. However, when the coding layer is optimized with respect to rhe mutual information between the stimulus and the category membership, accuracy decreass with the number of coding neurons. Secondly, as the number of neurons increases, a network with an uniform coding layer performs better that a network whose coding layer has been optimized.

These two effects are both due to the same reasons. In the case of an optimized coding layer, the tuning curves are sharp at the boundary (see Figure 6.2). A neuron close to the boundary will only emit spikes for stimuli close to the center of its tuning curve. This has an effect on the learning dynamics. A neuron close to the boundary will update its synaptic connection towards the decision layer less than for a neuron far away from the boundary, whose firing rate is higher in average. Moreover, it is difficult to categorize a stimulus close to the boundary, thus the mean reward rate for such stimulus will be lower than for a stimulus far from the boundary. Therefore, the weights decrease near the boundary (see Figure 7.3). But this goes against the fact that the Fisher information is maximal at this location. During the learning, the network has lost information as fine tuned coding neurons are associated with a synaptic connection of strength close to 0. This leads to worse performances for the optimized case than for the uniform one.

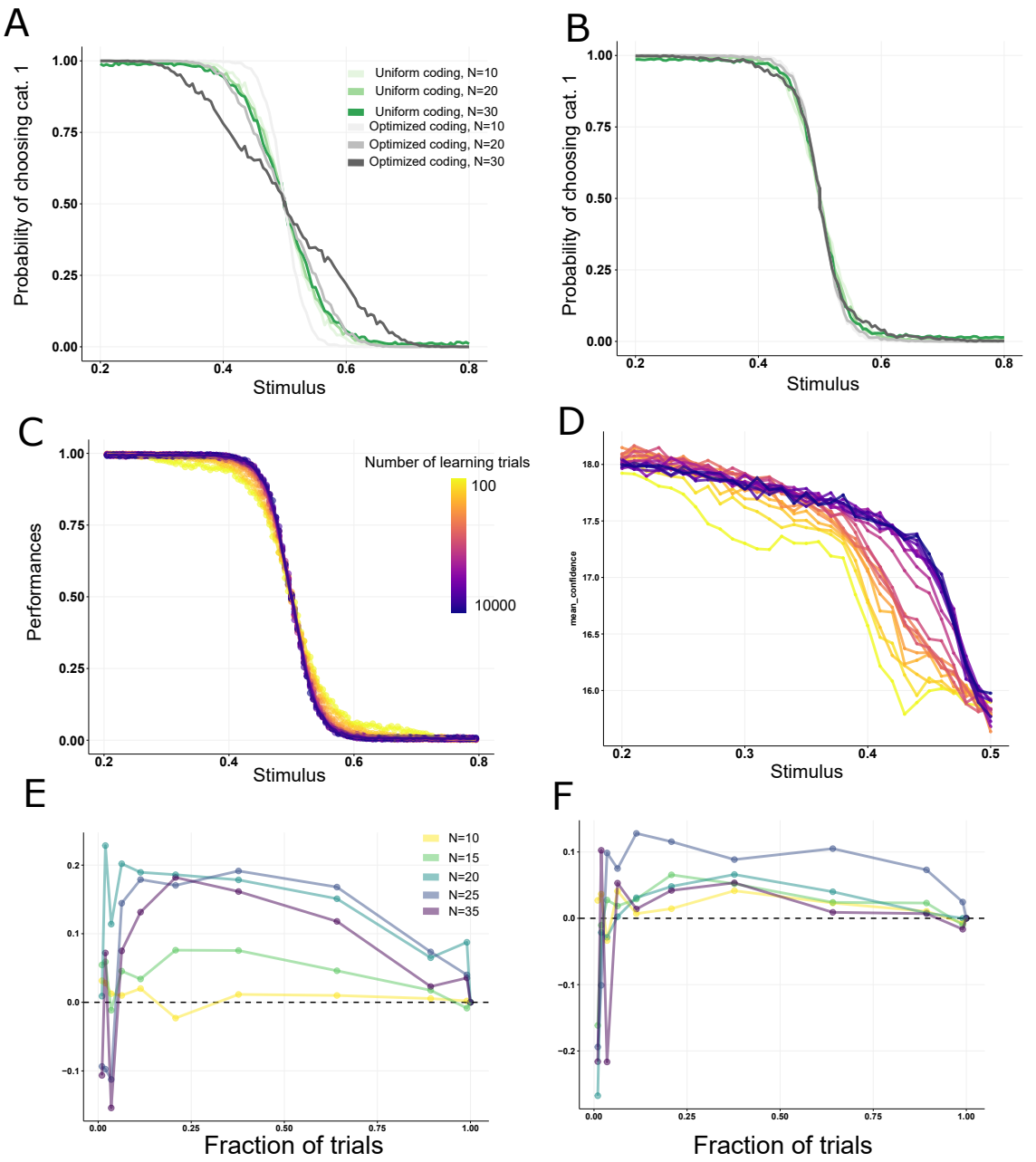


Figure 7.4: **Performance of the network after learning.** (A): Probability of choosing category C_1 for the network with respect to the stimulus presented x. The learning is performed without reward modulation by the confidence and after 2000 trials. The green curves correspond to a network with uniform coding layer, and the greys ones to an optimized coding layer. For both colors, the number of neurons in the coding layer is denoted by going from light to dark colors. The width of the gaussian categories is 0.25. (B): Same as (A) but this time for a learning with modulation by the confidence (with a threshold $z = 15$ Hz) (C): Performance of the network after different numbers of learning trials. The coding layer (20 neurons) has been optimized and the threshold of the confidence modulation is $z = 15$ Hz. (D): Average Δr on 1000 trials with respect to the stimulus ambiguity. The gaussian categories have a width of 0.25, and there is 20 neurons in the coding layer (optimized). Each represents a number of learning trials. (E): Effect of confidence modulation on the performances for a network with an optimized coding layer after 2000 trials. The y-axis represents the difference between the performances, at an ambiguity $x = 0.45$, for a learning with modulation by confidence and for a learning without. The x-axis represents the percentage of trials where Δr was lower than the threshold θ_c. (F): Same as (E) but for a neural circuit model with a uniformly distributed coding neurons.

Impact of confidence on performances.

How to counterbalance the loss of information due to reinforcement learning? The idea is to only update the synaptic connections when it is useful for the network, when it would gain information by doing so. In other words, once the categories have been more or less learned, one would like to fine tune the learning by only considering the hard trials (the ones near the categories boundary). This can be achieved through a modulation of the learning rate by the confidence. Figure 7.4.D shows the average Δr with respect to the stimulus ambiguity at different stage of learning. One can see that it decreases as it approaches the category boundary. Moreover, this behavior is even observed after a small number of trials. The representation of confidence emerges early during the learning and evolves with the number of trials. Figure 7.4.B represents the performances of the network, after 2000 learning trials with the confidence-modulated Hebbian learning ($\theta_c = 15$ Hz). One can see that the performances with the optimized coding layer are improved, especially at large N. One important point is that confidence modulation does not prevent the network from learning the task after a few number of trials (Figure 7.4.C).

Figure 7.4.E and F represent with more details the influence of confidence modulation on the learning process. Each point represents the difference in performances between the confidence modulated learning at a specific threshold and a learning without the modulation (for $x = 0.45$). Panel E shows the results for a network with an optimized coding layer and Panel F a network with a uniform coding layer. One can note two behaviors:

- Confidence modulation has more impact on the performances of a network with an optimized coding layer than for an uniform one.
- For low N, confidence modulation has only a small impact.

Therefore, the impact of confidence on the learning process strongly depends on the type of coding that is considered. This effect does not disappear with more learning trials (Figure 7.5.A and B)). The tendency of all the curves is the same, with an increase of performance when the confidence threshold decreases until it reaches a maximum at a confidence threshold ($\theta_c \simeq 15$ Hz) corresponding to $\sim 25\%$ of trials used. If the confidence threshold continues to decrease, the performances drop as too few trials are used to do the learning. This plateau in the performances exists after 10000 trials too (Figure 7.5.A and B). Confidence modulation is successful for improving the performances because the weights near the boundary are increased. Therefore, neurons with sharp tuning curves are given more weights and the gain of information due to the optimization of the coding layer is not lost (Figure 7.5.C).

Confidence-based RL accounts for almost optimal performances

Confidence modulation improves the accuracy of the network for the categorization task, but differently depending on the parameters of the task. Figure 7.6 presents the difference in accuracy between a network with an optimized coding layer, and one with an uniform coding layer, for different N and width of the categories α. Red (resp. blue) regions correspond to higher (resp. lower) performances for the optimized coding layer with respect to the uniform one. For intermediate values of categories widths (for example, $\alpha = 0.2$), there are two regimes (Figure 7.6.A). For small N, there is an increase of performance with an optimized coding layer in the case of a non-modulated learning. However as N increases, the performances become worse at the boundary and a uniform code performs better. Figure 7.6.C shows that the region where uniform coding performs better increases with α. However, the difference in accuracies between the two networks decreases if one keeps increasing α. This is explained by the fact that, when the categories are very wide, the overlap between the gaussian is big and the optimal code is more and more similar to n uniform code.

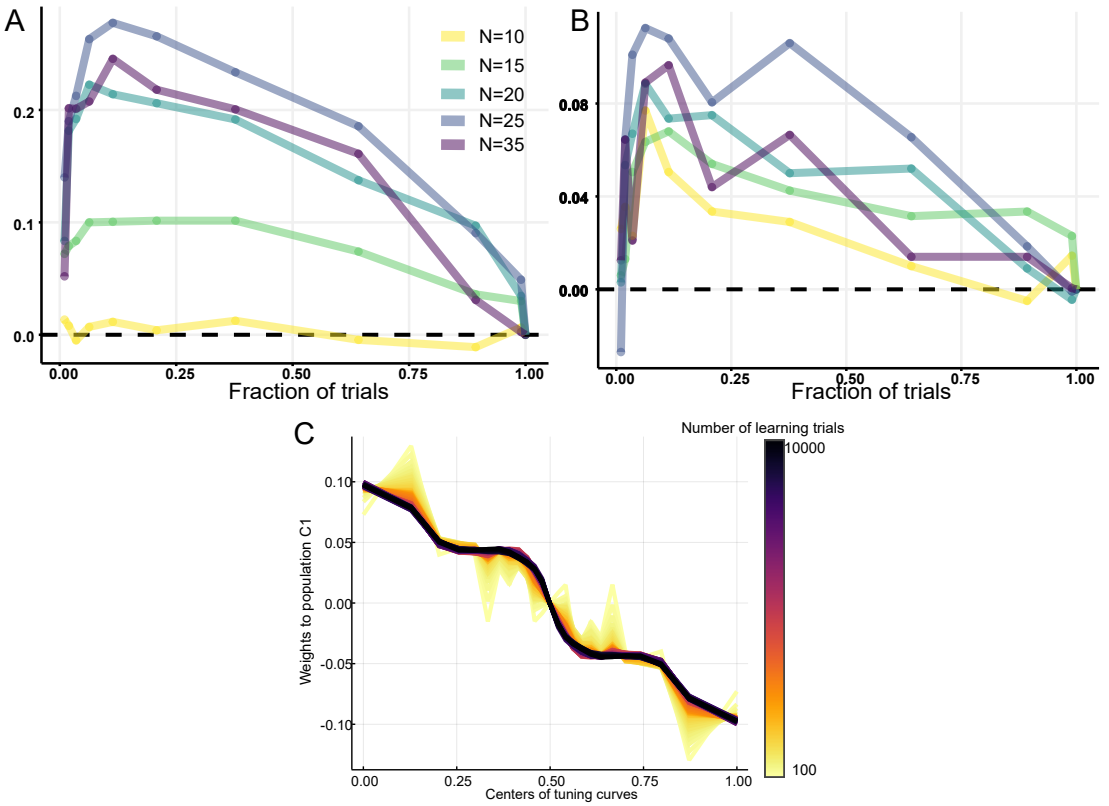


Figure 7.5: (A): Effect of confidence modulation on the performances for a network with an optimized coding layer after 10000 trials. The y-axis represents the difference between the performances, at an ambiguity $x = 0.45$, for a learning with modulation by confidence and for a learning without. The x-axis represents the percentage of trials where Δr was lower than the threshold θ_c. (B): Same as (A) but for a neural circuit model with a uniformly distributed coding neurons. (C): Synaptic weights going from the coding layer to the neural population C_1. Each color represents a different epoch in the learning process, starting at 100 trials. The coding layer is composed of 20 neurons, optimized for the categorization task and the width of the gaussian categories is 0.25.

Looking at Figure 7.6, one sees the impact of confidence modulation on the learning process. Figure 3.3.B and D show the same phase diagram as panels A and C but when the learning is modulated by confidence. In this case, a network with an optimized coding layer performs better than a uniform one, independently of the number of neurons N. Indeed, by focusing on trials with low confidence, the network tends to decrease the synaptic weights far from the boundary and increase the ones close to the boundary. This is in contrast with what happens in the Hebbian case with weights going to zero when approaching the boundary. With higher weights close to the boundary, the sharp tuning curves of the optimal code are used by the network to obtain a better accuracy (Figure 7.5.C). To conclude, confidence modulation of Hebbian learning allows the network to make use of an optimized coding layer in order to obtain better performances.

Confidence modulation improves the performances but one can ask whether such networks can achieve performances that are close to optimal ones. To find the optimal performances of the network, I use a non-linear convex optimisation procedure. The first step is to obtain the relation between the probability of choosing the most probable category

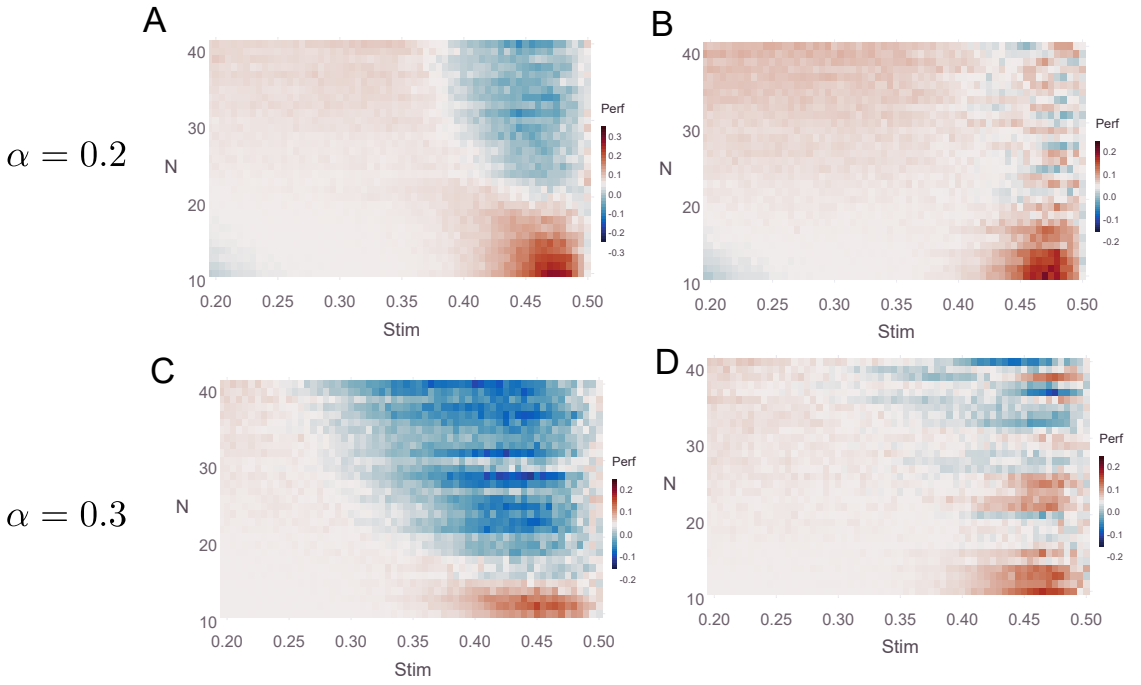


Figure 7.6: Difference of performances after learning (2000 trials) between networks with a uniform (or optimized) coding layer. Red (resp. Blue) corresponds to a positive (resp. negative) difference meaning that the network with the optimized coding layer has better (resp. worse) accuracy than the one with uniform layer. The x-axis corresponds to a variation of stimulus x, and the y-axis to the number of neurons N. (A) and (C) correspond to a learning without the modulation due to confidence. The width of the categories is 0.2 for (A), 0.3 for (C). (B) and (D) correspond to a learning with reward modulation by the confidence, with a threshold $z = 15$ Hz. The width of the categories is 0.2 for (B), 0.3 for (D).

and the input stimulus x for the attractor network. I note p this function. To perform the non-linear optimization, I will replace the attractor network by this function in the model. The second step is to define a cost function:

$$\mathcal{L} = \int_0^1 \left[1 - p(\sum_j W_j^1 f_j(s)) \right] ds \tag{7.8}$$

with $f_j(s)$ the tuning function of neuron j. The set of weights that minimize this cost function is found by solving a non-linear convex optimization problem, with the set of constraints defined as in the neural circuit model (synaptic normalization and symmetry between the units). The numerical routine was done in Matlab.

Figure 7.7.A and B present the difference of performances between the non-linear optimization and the confidence-modulated Hebbian learning, for $\alpha = 0.2$. Near the boundary, the network with optimized coding layer has better performances than the optimal model. However, the global performances are lower because it performs worse when the stimuli are a bit further from the boundary. Surprisingly, increasing the number of trials (Figure 7.7.C and D) does not lead to an increase of performances for a network with a uniform coding layer. For the optimized coding layer, the accuracy increases and

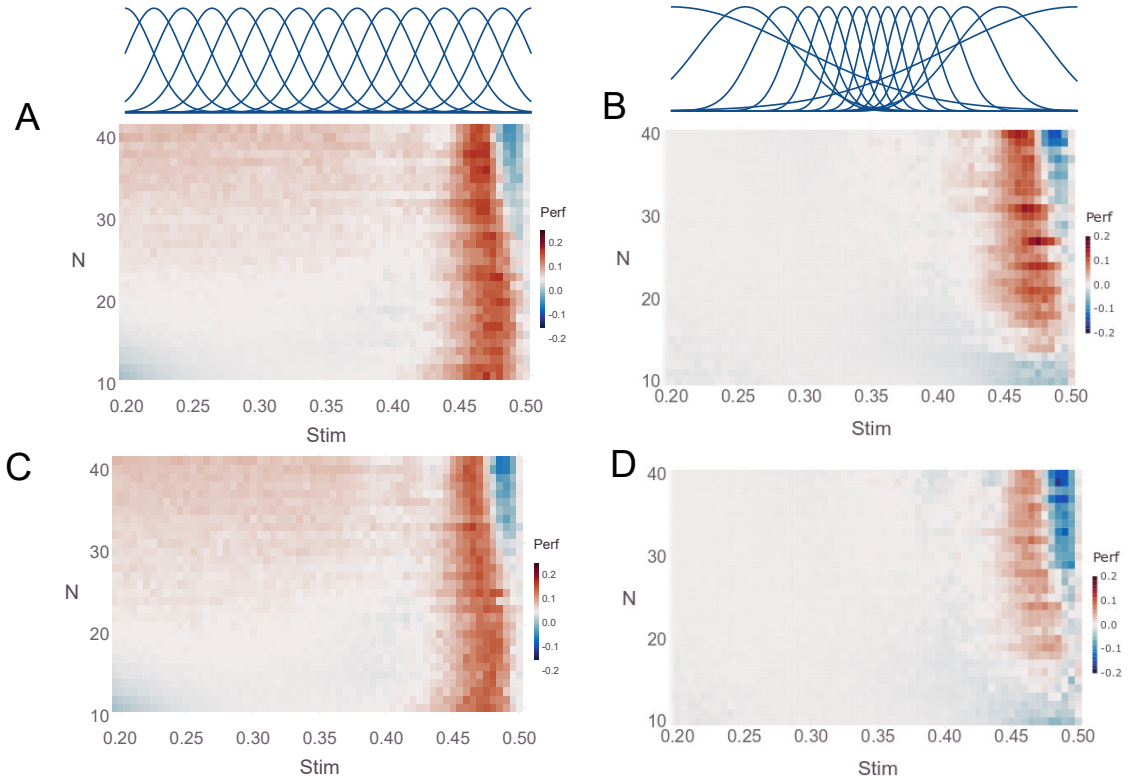


Figure 7.7: **Difference of performances between the result of the non-linear optimization and the model, with $\alpha = 0.2$.** (A) and (C): Difference with a uniform coding layer for 2000 trials (A) and 10000 trials (C). (B) and (D): Difference with an optimized coding layer for 2000 trials (B) and 10000 trials (D).

the network becomes very close to the optimal performances it could achieve (especially in the low N region). This behavior does not depend on specific values of α, as only the amplitude of this effect is modified.

Conclusion

To conclude this part, I have proposed a learning mechanism to learn categories that does not rely on a prediction error signal. I studied the efficiency of Hebbian learning in perceptual categorization tasks in the context of reward-based protocols. I assumed a coding layer providing a distributed representation of the stimulus, feeding a decision-making attractor neural network. Authors have modeled the learning of decision-making with a Reward Modulated Hebbian learning scheme. Considering a coding layer optimized in view of the categorization task (from an information-theoretic viewpoint), I found that RMHL fails to take advantage of this goal-directed stimulus encoding. Here I showed that a confidence-modulated, reward-based, Hebbian learning efficiently extracts category membership from the optimized coding layer. In the model, confidence is computed from the attractor network activity, in line with recent studies modeling the neural basis of confidence in decision-making. I have related this reward-modulation to different experimental findings that have shown that, in uncertain environments, the learning rate of human participants is correlated with their confidence. From a computational point of view, there is another advantage to use the confidence-modulated learning algorithm for a biological system. The reward-modulated Hebbian learning algorithm makes the assumption that the mean

rewards corresponding to all the possible stimuli are stored somewhere, and can be recalled to modulate the learning rate. In contrast, the confidence is only a local quantity and does not need an external memory storage. Yet, it still leads to better performance than the RMHL. Different extension of this work will be discussed in the Conclusion chapter.

Mean-field equations of neural networks

This part presents the results of a project that started during my M2 internship under the supervision of Gianluigi Mongillo, and that I have continued to work on during my PhD. Despite being a side project during my PhD thesis, I wanted to include in the manuscript as it is related to the general topic of attractor networks in neuroscience that constitue the first part of this manuscript. The project consisted in the study of dynamically balanced and partially connected neural networks. Very few studies have investigate partially connected neural networks, and more specially dynamically balanced neural networks. Most of the studies focused on a chaotic regime of activity for these networks. Here, I propose to investigate the effect of partial connectivity on this regime. I make use of a method entitled *cavity method* that has been successfully used in statistical physics.

As a result, I obtain the statistics of the neurons in the steady states and study their stabilities. There are two regions within the bifurcation diagram. In the first one, there is a unique stable fixed point for the network. In the second region, a breaking of ergodicity appends. I study with more details this breaking of ergodicity and show that, partially symmetric balanced neural networks exhibit glassy-like dynamics. Moreover, this breaking of ergodicity appends at levels of symmetry comparable to the ones observed in synaptic connections in neuroscience.

The different results that I present in this part will be published in an article *currently in preparation:*

- Berlemont, K. and Mongillo, G. (2020). Glassy-like dynamics of partially symmetric, dynamically balanced networks.

8

Mean-field equations of neural networks

8.1. Temporal dynamics of neural networks

Two main theoretical frameworks have been proposed to study the temporal dynamics in neuronal networks: rate-based models with Gaussian connections and spiking dynamics of sparsely connected excitatory-inhibitory networks.

Spiking networks

Networks of spiking neurons have been studied numerically or analytically in the case of fully symmetric networks (Amit et al., 1990; Abbott and van Vreeswijk, 1993; Hansel et al., 1995). During the last twenty years, sparsely connected networks have been studied, in the case of binary neurons or integrate-and-fire neurons (Van Vreeswijk and Sompolinsky, 1996). In a sparsely connected network, the mean number of connections K is much smaller than the number of neurons N. Due to the fact that the connections are random, it is possible, in the thermodynamic limit, to use a mean-field approach in order to self-consistently obtain the mean and the variance of the inputs of the neurons (Brunel, 2000). Doing so, it is possible to obtain the phase diagrams of such networks.

Depending on different parameters, such as the balance between excitation and inhibition, networks of sparsely connected excitatory-inhibitory neurons show different states (Figure 8.1):

- Synchronous regular state: neurons are almost fully synchronized.
- Asynchronous regular state: stationnary global activity and quasi-regular individual firing rates.
- Asynchronous irregular state: stationnary global activity but irregular individual firing rates.
- Synchronous irregular state: oscillatory global activity but strongly irregular individual firing rates.

Rate models

The second type of networks that has been studied consists in a network in which each unit is described by a firing-rate dynamics. Each neuron is characterized by a transfer function that maps the synaptic input into the output firing rate. Various transfer functions have been used such as $\tanh(x)$, rectified linear or sigmoid function (Sompolinsky et al., 1988; Kadmon and Sompolinsky, 2015). The structure of the connectivity matrix is a Gaussian distribution with zero mean and a variance g/N. It was shown that the system exhibits a transition from a stable fixed-point for low g to chaotic state (Sompolinsky et al., 1988). To illustrate this behavior, I will use the model of Stern et al. (2014). The network is

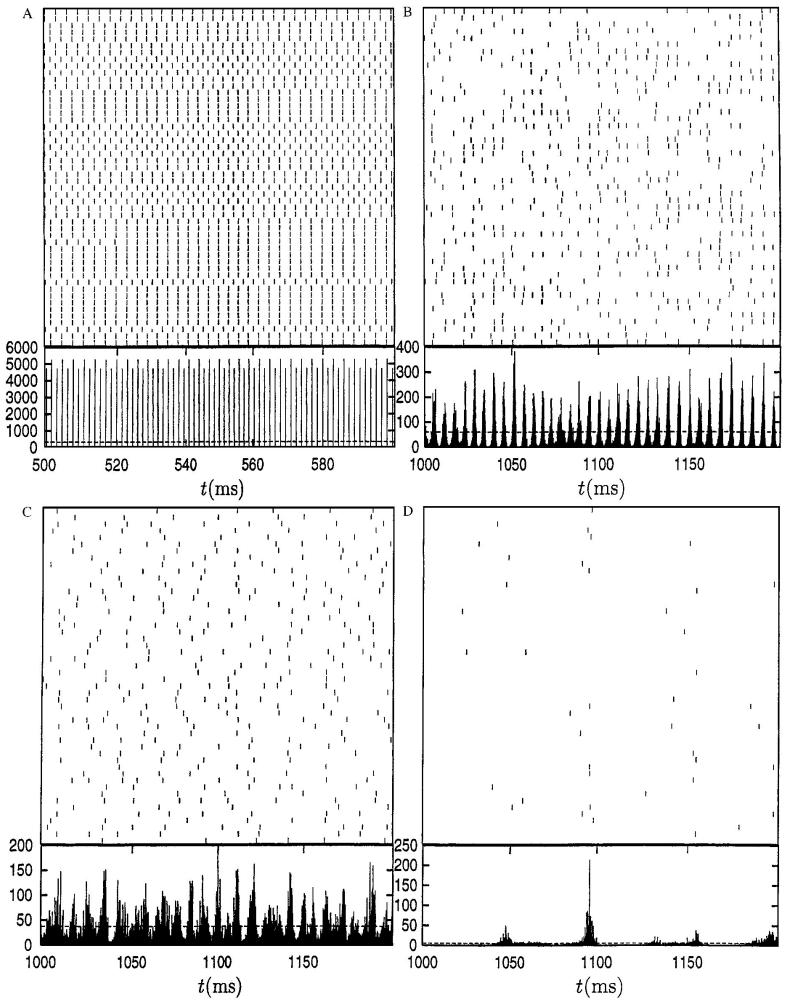


Figure 8.1: **Simulation of a spiking network (adapted from Brunel (2000)).** (A) Almost fully synchronized network. (B) Fast oscillation of the global activity. (C) Stationnary global activity. (D) Slow oscillation of the global activity.

composed of N units obeying the equations:

$$\frac{dh_i}{dt} - h_i + s\tanh(h_i) + g\sum_{j\neq i} J_{ij}\tanh(h_j) \tag{8.1}$$

with J_{ij} the connection from neuron j to neuron i (drawn from a Gaussian distribution with zero mean and $1/N$ variance). In this model, the state in which all the neurons have $h_i = 0$ is always a fixed point. The stability will depend on the synaptic gain g and the self-coupling s. The zero-fixed point is the matrix defined by:

$$M_{ij} = (-1+s)\delta_{ij} + gJ_{ij} \tag{8.2}$$

has all of its eigenvalues with negative real part. Thus, the bifurcation line between stability and unstability is defined by $s = 1-g$. The bifurcation diagram of the network is shown in Figure 8.2. There are three regions: a chaotic region, region with a decay to zero, and a region with non-zero fixed points (Figure 8.2). Networks with random connections can exhibit various dynamical behaviors depending on the parameters. There can be a decay to a fixed point, a chaotic activity or even a transient chaotic activity until convergence to a fixed point.

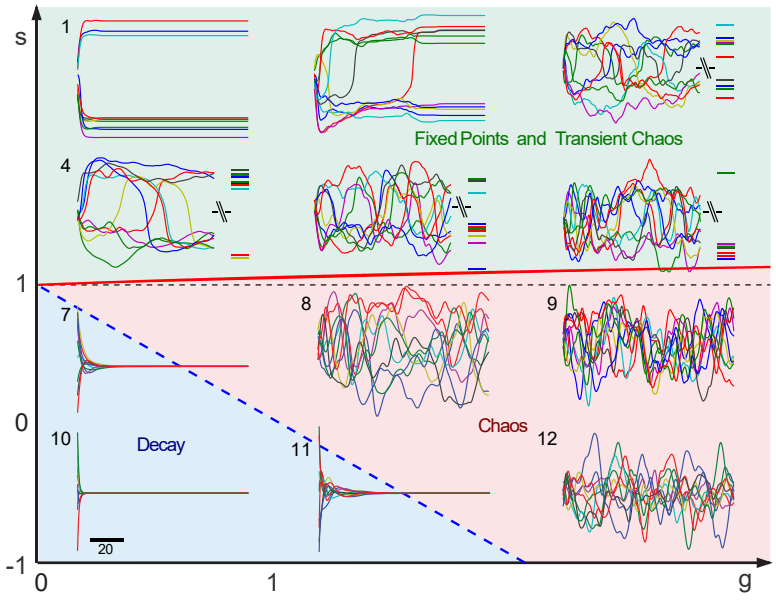


Figure 8.2: **Example of activity of a rate based model (adapted from Stern et al. (2014)).**

Partially symmetric networks

In the mammalian brain, cortical networks exhibit a complex connectivit that, to a first approximation, can be regarded as random. As mentionned previously, random neural networks can exhibit a regime where they satisfy the local chaos hypothesis (Amari, 1972; Geman and Hwang, 1982). This dynamical regime has been shown as potentially useful for computations (Bertschinger and Natschläger, 2004).

But, what happens when such networks are neither totally symmetric or randomly connected? It is known that local connectivity can have a strong effect on network activity (Litwin-Kumar and Doiron, 2012). Experimental studies on cortical connectivity have identified many patterns of connectivity that strongly deviate from the independancy between neurons assumption. Reciprocal connections are overpresent in cortical networks and are consistent with a partially symmetric structure (Markram et al., 1997; Song et al., 2005; Harris and Mrsic-Flogel, 2013). Very few studies have tackled the question of the impact of partial symmetry on the networks dynamics, principally due to the fact that it renders the mathematical analyses more challenging (Crisanti and Sompolinsky, 1987; Renart et al., 2010). In the case of a network with zero mean connections, it has been shown that symmetry in the network slows the dynamics down, which is important to perform computations with neural networks (Martí et al., 2018). However, in this model, neurons can have negative activity and can have excitatory and inhibitory projections which does not satisfy Dale's law.

Cavity method

A large part of the study of statistical physics of disordered systems has focused on the physics of spin glasses. Spin glasses are similar to networks of binary neurons as the models are described by Hamiltonian of the form:

$$H = -\sum_{ij} J_{ij}\sigma_i\sigma_j \tag{8.3}$$

with $\sigma_i = \pm$, and J_{ij} the coupling between two spins. The model 8.3 has been studied for many configurations of the matrix of couplings such as fully connected or partially connected. Physicists have developped two main methods to study these problems: the replica method and the cavity method (Mézard et al., 1987; Mézard and Parisi, 2003). Both methods are equivalent, but the latter is more intuitive in the case of partially connected networks. The cavity method has been applied to many types of problems: spin glasses (Mézard and Parisi, 2003), optimization problem (Zdeborová and Krzakala, 2010), ecological problems (Barbier and Arnoldi, 2017). The cavity method has been used to study random neural networks and more specifically the effect of the assymetry on the local chaos hypothesis (Cessac, 1995). The author found that non-assymetric couplings lead to the failure of the local chaos hypothesis.

More recently, this method has been used to study the distributions of cortical connectivity in order to maximize the storing of a large number of attractor within a neural network (Brunel, 2016). One interesting result is that the resulting connectivity matrix of this excitatory network is sparse (large number of zero weights connections) and has an over-representation of bidirectionally couplings. In the following section, I will use this method in order to study the steady-states of dynamically balanced partially symmetric networks of rate inhibitory neurons. I will show that this method gives the bifurcation diagram of the system. There are several differences I would like to point out with the previous study I mentionned, that used cavity method to study random neural networks. First, the network I consider is dynamically balanced. This means that the state *all neurons are silent* does not correspond to a steady-state in my model. In addition, I propose to perform the computations without the introduction of additional hypothesis, such as the local chaos hypothesis. This allows for a general framework in order to study the steady-state of such networks. Moreover, Cessac (1995) does not study the behavior of the random networks when the local chaos hypothesis can not be applied anymore. Here, I investigate how the network behaves in the regime where it exhibits a breaking of ergodicity.

8.2. Dense strongly coupled partially symmetric networks

I will consider partially symmetric neural networks that are dense and strongly coupled. Considering strongly coupled networks means that the synaptic weights are of order $1/\sqrt{N}$, with N the total number of neurons. Each neuron is characterized by its firing rate r_i and receives inputs from its neighbors neurons. The total current that is received is:

$$h_i = \sqrt{N} h_{ext} - \frac{1}{\sqrt{N}} \sum_j w_{ij} \nu_j \tag{8.4}$$

with h_{ext} the external current, and w_{ij} the synaptic weight from neuron j to neuron i. The firing rate is defined as:

$$\nu_i = \phi(h_i) \tag{8.5}$$

with ϕ the transfer function.

The dynamics of the network is the following:

$$\frac{dh_i}{dt} = -h_i + \sqrt{N} h_{ext} - \frac{1}{\sqrt{N}} \sum_j w_{ij} \nu_j \tag{8.6}$$

The synaptic efficacies are log-normalyy distributed:

$$w_{ij} = \exp\left(\mu_t + t_{ij} \sigma_t\right) \tag{8.7}$$

where t_{ij} is Gaussian with mean zero and unitary variance, μ_t and σ_t are free parameters used to set the mean $\langle w \rangle$ and the variance σ_w^2 of the synaptic efficacies. These are given by:

$$\langle w \rangle = \exp\left(\mu_t + \frac{1}{2}\sigma_t^2\right) \tag{8.8}$$

$$\sigma_w^2 = \exp\left(2\mu_t + \sigma_t^2\right)\left(e^{\sigma_t^2} - 1\right) \tag{8.9}$$

The level of symmetry of the network, ρ_w, can be controlled by adjusting the correlation $\rho_t = \langle t_{ij} t_{ji} \rangle$ according to:

$$\rho_w = \frac{\langle w_{ij} w_{ji} \rangle - \langle w \rangle^2}{\sigma_w^2} \tag{8.10}$$

The w_{ij} are non-negative, which leads to:

$$\rho_w \geq \frac{e^{-\sigma_t^2} - 1}{e^{\sigma_t^2} - 1} \tag{8.11}$$

It is worth noting that, in this particular case, a full anti-symmetric network ($\rho_w = -1$) is only achievable in a limiting sense for $\sigma_t^2 \to 0$.

8.2.1 Mean-field solution

In this section, I derive a mean-field theory for the model netork previously described, which is exact (for infinitely large networks) in the region of parameters where the network exhibits a unique stable fixed point. However, the unicity of the fixed point is not necessary for the theory to be valid. What is necessary is that the fixed point is asymptotically stable.

In a stable fixed point, $\frac{dh_i}{dt} = 0$, thus the input to neuron i can be written as:

$$h_i = \sqrt{N} h_{ext} - \frac{1}{\sqrt{N}} \sum_j w_{ij} \nu_j \tag{8.12}$$

The idea of the cavity-like approach is to decompose the input as:

$$h_i = \sqrt{N} h_{ext} - \frac{1}{\sqrt{N}} \sum_j w_{ij} \nu_j^- - \frac{1}{\sqrt{N}} \sum_j w_{ij} \delta\nu_j^- \tag{8.13}$$

The term ν_j^- corresponds to the activity of neuron j in a network where all the connections going from neuron i to the other neurons have been removed (Figure 8.3). At the contrary, $\delta\nu_j^-$ corresponds to the difference in the activity of neuron j that is due to the fact that neuron i has connections to the other neurons. I will define the cavity input as:

$$u_i = \sqrt{N} h_{ext} - \frac{1}{\sqrt{N}} \sum_j w_{ij} \nu_j^- \tag{8.14}$$

Cavity input

The above decomposition is useful as each term can be studied separately. For the cavity input, as the network is at a fixed point, there are no long-range correlations in the pattern of activity. Therefore, the cavity input becomes Gaussian in the large N limit because the ν_j^- are independent of the w_{ij}'s (as it corresponds to the activity in a network where w_{ji} has been removed):

$$\nu_j^- = \sqrt{N} h_{ext} - \frac{1}{\sqrt{N}} \sum_{k \neq i} w_{jk} r_k \tag{8.15}$$

It is worth noting that this notion of independency between the ν_j and the w_{ij} has been first mentionned in (Amari, 1972; Crisanti et al., 1990) under the *local chaos hypothesis*, in the general case. More recently, this hypothesis has been studied within a mathematical framework (Faugeras and Maclaurin, 2015; Salhi et al., 2018), using large deviations theory. However, in the present framework, I do not need to make use of the local chaos hypothesis in order to obtain the independance, as neuron i has been *removed* from the network at this step of the computations. Such the specific use of this cavity-inspired method allows me to directly obtain the independency property between the ν_j^- and the w_{ij}'s.

Thus, u_i can be rewritten as (in the large N limit):

$$u_i = \mu + z_i \sigma \tag{8.16}$$

with z_i a Gaussian variable with zero mean and unitary variance. The variables μ and σ can be computed using standard mean-field computation:

$$\langle u \rangle = \sqrt{N} \left(h^{(ext)} - \langle w \rangle \langle \nu \rangle \right) \tag{8.17}$$

$$s^2 = \langle w^2 \rangle \langle \nu^2 \rangle - \langle w \rangle^2 \langle \nu \rangle^2 \tag{8.18}$$

To obtain this formula, I have used the fact that, in the limit $N \to \infty$, $\langle (\nu)^n \rangle = \langle (\nu^-)^n \rangle$.

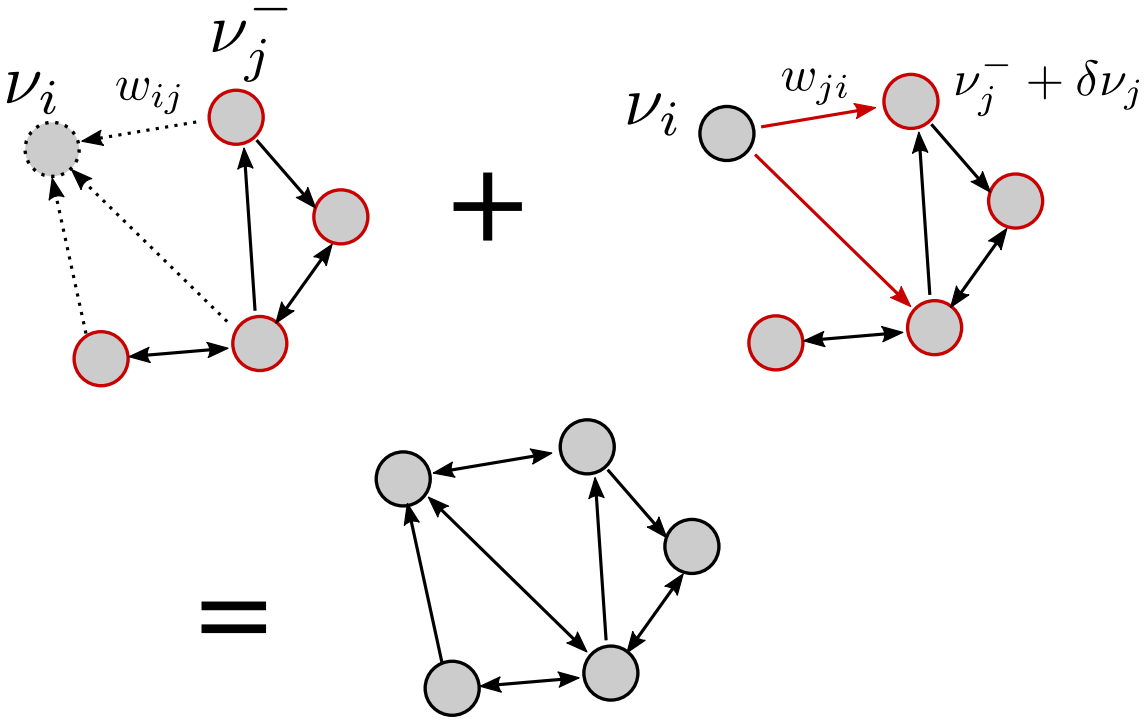


Figure 8.3: **Schema of the cavity approach.** Each circle represents a neuron in the network. The red neurons are the ones that are in the cavity network. The upper left network corresponds to the cavity network: without the connections from neuron i to the other neurons. The upper right network corresponds to the network with the reaction due to neuron i. By imposing the self-consistency between these two networks, one can obtain the full network (bottom).

Reaction term

The presence of neuron i induces a perturbation of order $1/\sqrt{N}$ in the input to neuron j. To compute the impact of neuron i on the network I will make use of linear response theory. In the previous section I mentionned some results about the local chaos hypothesis and the propagation of chaos. In order for the linear response theory to be valid, it is necessary to take into account the second-order (and more) effect of neuron i on the network. This has been done using large deviations theory (Moynot and Samuelides, 2002; Faugeras and Maclaurin, 2015; Salhi et al., 2018), and, in this case, linear response theory can be used to compute the reaction term.

One quantity that arises naturally in linear response theory is the static susceptibility. In this network, the static susceptibility is defined as:

$$\chi_{jk} = \frac{\delta \nu_j}{\delta h_k} \tag{8.19}$$

It represents the static susceptibility of neuron j to a small change in the input to neuron k. With this quantity, the total input h_i can be written as:

$$h_i = u_i - \frac{1}{\sqrt{N}} \sum_j w_{ij} \delta \nu_j \tag{8.20}$$

$$= u_i - \frac{1}{\sqrt{N}} \sum_j \sum_k w_{ij} \frac{\delta \nu_j}{\delta h_k} \delta h_k \tag{8.21}$$

$$= u_i - \frac{1}{\sqrt{N}} \sum_j \sum_k w_{ij} \frac{\delta \nu_j}{\delta h_k} \left(-\frac{-1}{\sqrt{N}} w_{ki} \nu_i \right) \tag{8.22}$$

This double sum can be decomposed into diagonal terms and off-diagonal terms:

$$\begin{aligned} -\frac{1}{\sqrt{N}} \sum_j \sum_k w_{ij} \frac{\delta \nu_j}{\delta h_k} \left(-\frac{-1}{\sqrt{N}} w_{ki} \nu_i \right) =& \nu_i \left\{ \frac{1}{N} \sum_j w_{ij} w_{ji} \chi_{jj} \right\} \\ &+ \nu_i \left\{ \frac{1}{N} \sum_k w_{ki} \left(\sum_{j \neq k} w_{ij} \chi_{jk} \right) \right\} \end{aligned} \tag{8.23}$$

In order to compute both sums one should note that $\sum_j \chi_{jk} = 0$. This equation can be obtained by the following argument. A balanced solution of the dynamical equations is so that only a negligible number of ν_i is saturated in the limit $N \to \infty$. In order to understand the emergence of such a solution, and the effects of small perturbation on it, it is possible to write:

$$\nu_j = \langle \nu^{(0)} \rangle + \delta \nu_j^{(0)} + \frac{1}{\sqrt{N}} \left(\langle \nu^{(1)} \rangle + \delta \nu_j^{(1)} \right) + \frac{1}{N} \left(\langle \nu^{(2)} \rangle + \delta \nu_j^{(2)} \right) + \cdots \tag{8.24}$$

$$w_{ij} = \langle w \rangle + \delta w_{ij} \tag{8.25}$$

This expression can be inserted into the expression for the input h_i:

$$\begin{aligned} h_i &= \sqrt{N} \left(h^{(ext)} - \langle w \rangle \langle \nu^{(0)} \rangle \right) + \\ &= -\frac{1}{\sqrt{N}} \sum_j \delta w_{ij} \delta \nu_j^{(0)} - \langle w \rangle \langle \nu^{(1)} \rangle + \\ &= -\frac{1}{N} \sum_j \delta w_{ij} \delta \nu_j^{(1)} - \frac{1}{\sqrt{N}} \langle w \rangle \langle \nu^{(2)} \rangle + O\left(\frac{1}{N} \right) \end{aligned} \tag{8.26}$$

In a balanced solution, $h_i = O(1)$ and thus, if such solution exists, it must be:

$$h_{ext} - \langle w \rangle \langle \nu^{(0)} \rangle = 0 \tag{8.27}$$

Similarly, it must be:

$$\frac{1}{\sqrt{N}} \sum_j \delta w_{ij} \delta \nu_j^{(0)} = O(1) \tag{8.28}$$

$$\frac{1}{N} \sum_j \delta w_{ij} \delta \nu_j^{(1)} = O\left(\frac{1}{\sqrt{N}}\right) \tag{8.29}$$

A perturbation of $O(1)$ will change the $\delta\nu_j^{(0)}$'s but not $\langle \nu^{(0)} \rangle$. Similarly, a perturbation of $O(1/\sqrt{N})$ will change the $\delta\nu_j^{(1)}$'s but not $\langle \nu^{(1)} \rangle$. This is true for any possible perturbation (of a given order), and thus it must be

$$\sum_{k=1}^{N} \chi_{jk} = 0; \quad k = 1 \cdots N \tag{8.30}$$

In the following I will note χ the mean susceptibility of the network:

$$\chi = \frac{1}{N} \sum_k \chi_{kk} \tag{8.31}$$

Let's get back to the Equation 8.23. The diagonal term can be derived as follows:

$$\begin{aligned} \nu_i \frac{1}{N} \sum_j w_{ij} w_{ji} \chi_{jj} &= -\nu_i \frac{1}{N} \sum_j w_{ij} w_{ji} \sum_{k \neq j} \chi_{jk} \\ &= -\nu_i \frac{1}{N} \sum_j \sum_{k \neq j} w_{ij} w_{ji} \chi_{jk} \\ &= -\nu_i \sum_{k \neq j} \langle w_{ij} w_{ji} \rangle \chi_{jk} \end{aligned} \tag{8.32}$$

The last line is deduced from the fact that χ_{jk} is independent of w_{ij} and w_{ji}. Finally, it leads to:

$$\nu_i \frac{1}{N} \sum_j w_{ij} w_{ji} \chi_{jj} = \nu_i \left(\langle w \rangle^2 + \rho_w \sigma_w^2 \right) \chi \tag{8.33}$$

The off-diagonal term is obtained with the following computations:

$$\nu_i \left\{ \frac{1}{N} \sum_k w_{ki} \left(\sum_{j \neq k} w_{ij} \chi_{jk} \right) \right\} = \frac{1}{N} \sum_k w_{ki} \left(\langle w \rangle \sum_{j \neq k} \chi_{jk} + \sum_{j \neq k} \delta w_{ij} \chi_{jk} \right) \tag{8.34}$$

$$= -\frac{\langle w \rangle}{N} \sum_k w_{ki} \chi_{kk} + O\left(\frac{1}{\sqrt{N}}\right) \tag{8.35}$$

$$= -\langle w \rangle^2 \chi + O\left(\frac{1}{\sqrt{N}}\right) \tag{8.36}$$

To obtain the last equation, one needs to use the fact that the χ_{jk} are independent of both w_{ij} and w_{ki}.

Self-consistency

The total input h_i can now be written:

$$h_i = u_i + R\phi(h_i) = \psi(u_i) \tag{8.37}$$

with $R = \rho_w \sigma_w^2 \chi$. This equation defines a mapping between the cavity input u_i and the input in the full network h_i through the function ψ.

$$\nu_i = \phi(u_i + R\nu_i) = \hat{\phi}(u_i) \tag{8.38}$$

In the limit $N \to \infty$, the average activity in the network is determined by the balance condition:

$$\langle \nu \rangle = \frac{h_{ext}}{\langle w \rangle} \tag{8.39}$$

Moreover, one must have:

$$\langle \nu \rangle = \int Dz \phi(\psi(\mu + z\sigma)) = \langle \hat{\phi} \rangle = \frac{h_{ext}}{\langle w \rangle} \tag{8.40}$$

with Dz the standard Gaussian measure. The variance can be written as:

$$\sigma^2 = \sigma_w^2 \langle \nu^2 \rangle = \int Dz \phi^2(\psi(\mu + z\sigma)) = \sigma_w^2 \langle \hat{\phi}^2 \rangle \tag{8.41}$$

The last equation one needs to obtain self-consistency of the system is one on the local susceptibility:

$$\chi = \int Dz \frac{d\hat{\phi}}{du}(\mu + z\sigma) = \langle \hat{\phi}' \rangle \tag{8.42}$$

For each of these equations, the effective mapping $\hat{\phi}$ is defined through Equation 8.38.

8.2.2 Stability

The mean-field theory developped previously is based on the assumption that the fixed point is stable. To check the validity of this assumption, one needs to study the stability of the mean-field solution. To do so I will compute the variance of the $\delta\nu_j$'s. Indeed, in a stable solution $N \cdot \langle \delta\nu^2 \rangle = O(1)$.

Using the effective f-I curve of the network:

$$\begin{aligned} \delta\nu_j &= \hat{\phi}(u_j) - \hat{\phi}(u_j^-) \\ &= -\frac{\hat{\phi}'(h_j^-)}{\sqrt{N}} \left(w_{ji}\nu_i + \sum_k w_{jk}\delta\nu_k \right) \end{aligned} \tag{8.43}$$

with u_j^- the cavity input to neuron j when neuron i is not connected to the rest of the network. This equation can be rewritten in power of $\sqrt{N}$:

$$\delta\nu_j = \frac{\delta\nu_j^{(1)}}{\sqrt{N}} + \frac{\delta\nu_j^{(2)}}{N} \tag{8.44}$$

Due to the fact that $\langle \delta\nu \rangle = O(1/N)$, $\langle \Delta\nu^{(1)} \rangle = 0$. An identification of the different terms in Equation 8.43 leads to:

$$\delta\nu_j^{(1)} = -\hat{\phi}'(u_j^-) \times \left(w_{ji}\nu_i + \langle w \rangle \langle \delta\nu^{(2)} \rangle + \frac{1}{\sqrt{N}} \sum_k \delta w_{jk} \delta\nu_k^{(1)} \right) \tag{8.45}$$

Averaging over j leads to $\langle \delta\nu^{(2)} \rangle = -\nu_i$. Injecting this relation into Equation 8.45:

$$\delta\nu_j = -\hat{\phi}'(u_j^-)\left(\delta w_{ji}\nu_i + \frac{1}{\sqrt{N}}\sum_k \delta w_{jk}\delta\nu_k^{(1)}\right) \tag{8.46}$$

Finally, by taking the square and averaging over j one can obtain:

$$\langle \delta\nu^2 \rangle = \frac{\sigma_w^2\langle\hat{\phi}'^2\rangle}{1-\sigma_w^2\langle\hat{\phi}'^2\rangle}\nu_i^2 = R_{nl}\nu_i^2 \tag{8.47}$$

The term R_{nl} denotes the non-linear reaction term due to neuron i. When this term diverges, the variances of the $\delta\nu_j$'s diverges and the solution is unstable.

8.3. Rectified linear transfer function

The previous theory does not depend on the specific form of the transfer function. For numerical purpose, I will be considering rectified linear units in the following section as it will allow me to derive exact analytical formula and to compare them with numerical simulations.

8.3.1 Steady-state of the network

In the case of a rectified linear transfer function, one has:

$$h = u + R\beta[h]_+ = \psi(u) \tag{8.48}$$

This leads to:

$$\psi(u) = \begin{cases} u & \text{if } u < 0 \\ \frac{u}{1-\beta R} & \text{otherwise} \end{cases} \tag{8.49}$$

The mapping betwen ν and u, denoted by $\hat{\phi}$ is obtained as:

$$\nu = \hat{\phi}(u) = \frac{\beta}{1-\beta R}[u]_+ \tag{8.50}$$

The derivative is exprimed as:

$$\frac{d\hat{\phi}}{du}(u) = \begin{cases} 0 & \text{if } u < 0 \\ \frac{\beta}{1-\beta R} & \text{otherwise} \end{cases} \tag{8.51}$$

I define the following functions:

$$G(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-x^2/2\right) \tag{8.52}$$

$$H(x) = \int_x^\infty dy G(y) \tag{8.53}$$

To solve the cavity equations, one can rewrite them as functions of the parameter $x = \frac{-\mu}{\sigma}$.

$$\begin{aligned} \langle\nu\rangle &= \int Dz\hat{\phi}(\mu + z\sigma) \\ &= \int Dz\frac{\beta}{1-\beta R}[\mu + z\sigma]_+ \\ &= \frac{\beta\sigma}{1-\beta R}\int Dz[\frac{\mu}{\sigma} + z]_+ \\ &= \frac{\beta\sigma}{1-\beta R}\left[G(x) - xH(x)\right] \end{aligned}$$

$$\begin{aligned}\langle \nu^2 \rangle &= \int Dz \hat{\phi}^2(\mu + z\sigma) \\ &= \int Dz (\frac{\beta}{1-\beta R})^2 [\mu + z\sigma]_+^2 \\ &= \int_x^\infty Dz (\frac{\beta}{1-\beta R})^2 (\mu + z\sigma)^2 \\ &= \int_x^\infty Dz (\frac{\beta\sigma}{1-\beta R})^2 (-x + z)^2 \\ &= (\frac{\beta\sigma}{1-\beta R})^2 \left[-xG(x) + (1+x^2)H(x) \right]\end{aligned}$$

$$\chi = \frac{1}{1-\beta R} H(x) \tag{8.54}$$

In order to solve these equations it is necessary to find the value of x in the fixed point. One can write:

$$\sigma^2 = \sigma_w^2 \langle \nu^2 \rangle = \sigma_w^2 (\frac{\beta\sigma}{1-\beta R})^2 \left[-xG(x) + (1+x^2)H(x) \right] \tag{8.55}$$

Dividing both sides by σ^2, I obtain:

$$1 - \beta R = \sigma_w \beta \sqrt{[-xG(x) + (1+x^2)H(x)]} = B(x) \tag{8.56}$$

But, using Equation 8.54:

$$R = \rho_w \sigma_w^2 \chi = \frac{\rho_w \sigma_w^2 H(x)}{B(x)} \tag{8.57}$$

with $\rho_w = \langle w_{ij} w_{ji} \rangle - \langle w \rangle^2$. In the case of the numerical simulations, $\beta = 1$, which leads to the following equation that is used to determine x in the fixed point:

$$1 - B(x) - \rho_w \sigma_w^2 \frac{H(x)}{B(x)} = 0 \tag{8.58}$$

Numerical simulations

Figure 8.4 shows the total and the cavity input in the network. First, one can note that the cavity input follows a Gaussian distribution as expected. For $\rho_w = 0$, the cavity input and the total input h are similar as the reaction term R is equal to zero. For $\rho_w \neq 0$, the theory matches perfectly the numerical simulations and captures the non-gaussian character of h_i.

Figure 8.5 shows the different variables of the steady-state of the network with respect to the degree of symmetry ρ_w. The first thing one can note is that theory and numerical simulations are in accordance. In the previous figure I have shown that the form of the distribution of the input h_i was modified due to ρ_w. Figure 8.5 highlights the fact that the symmetry in the network has a strong impact on the steady-state. For instance, if ρ_w is increased or decreased the value of μ can be reduced by half or doubled. The mean-field solution would have led to a solution far from the reality.

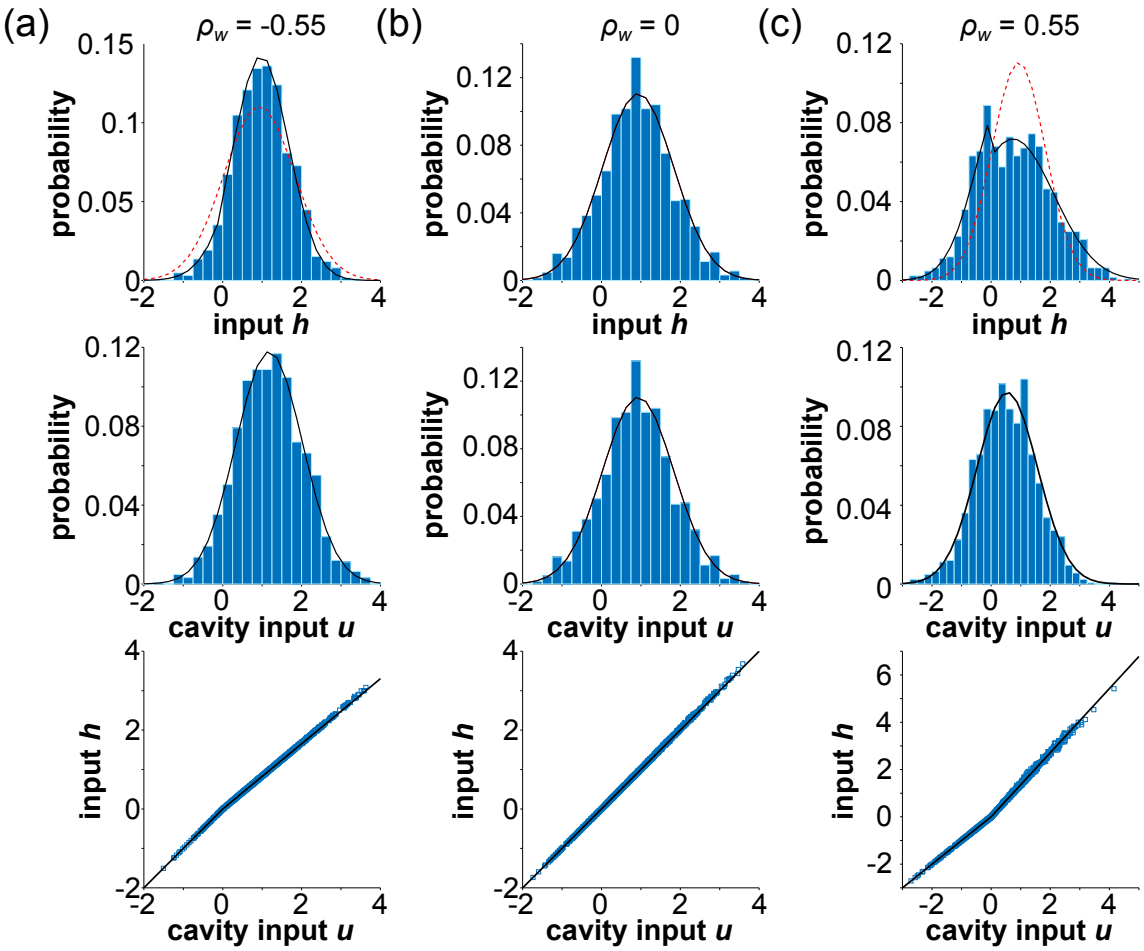


Figure 8.4: **Total and cavity input in the network.** The first line shows the histogram of the input h in the network. The black curve represents the theoretical result using the cavity method. The red curve shows the mean-field solution. The second line of panels represents the histogram of the cavity input in the network. Finally, the last line shows the mapping between cavity input and input h. Each column stands for a different value of ρ_w, $\rho_w = -0.55$ (a), $\rho_w = 0$ (B), $\rho_w = 0.55$ (c).

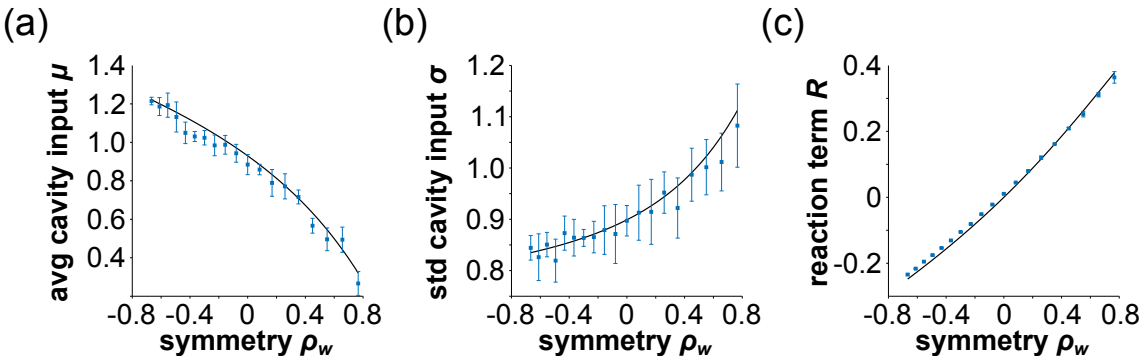


Figure 8.5: **Steady state of the network.** (A): Mean cavity input with respect to ρ_w. The black curve represents the theoretical solution and the blue points the numerical simulation. The error bars are the boostrapped confidence interval at 95%. (B): Standard deviation of the cavity input σ. (C): Reaction term R with respect to ρ_w.

8.3.2 Stability of the fixed point

At the fixed point the variance of the $\delta\nu_j$'s is of order $1/N$. Figure 8.6 represents the numerical simulations corresponding to this analysis. As described, the mean effect of neuron i in the network is of order $1/N$ (Figure 8.6.A), and the variance of the $\delta\nu_j$'s too (Figure 8.6.B). Figure 8.6.C shows the dependence of R_{nl} with ρ_w. It increases until the divergence that corresponds to the lose of stability. With these parameters it corresponds, here, to $\rho_w = 1$.

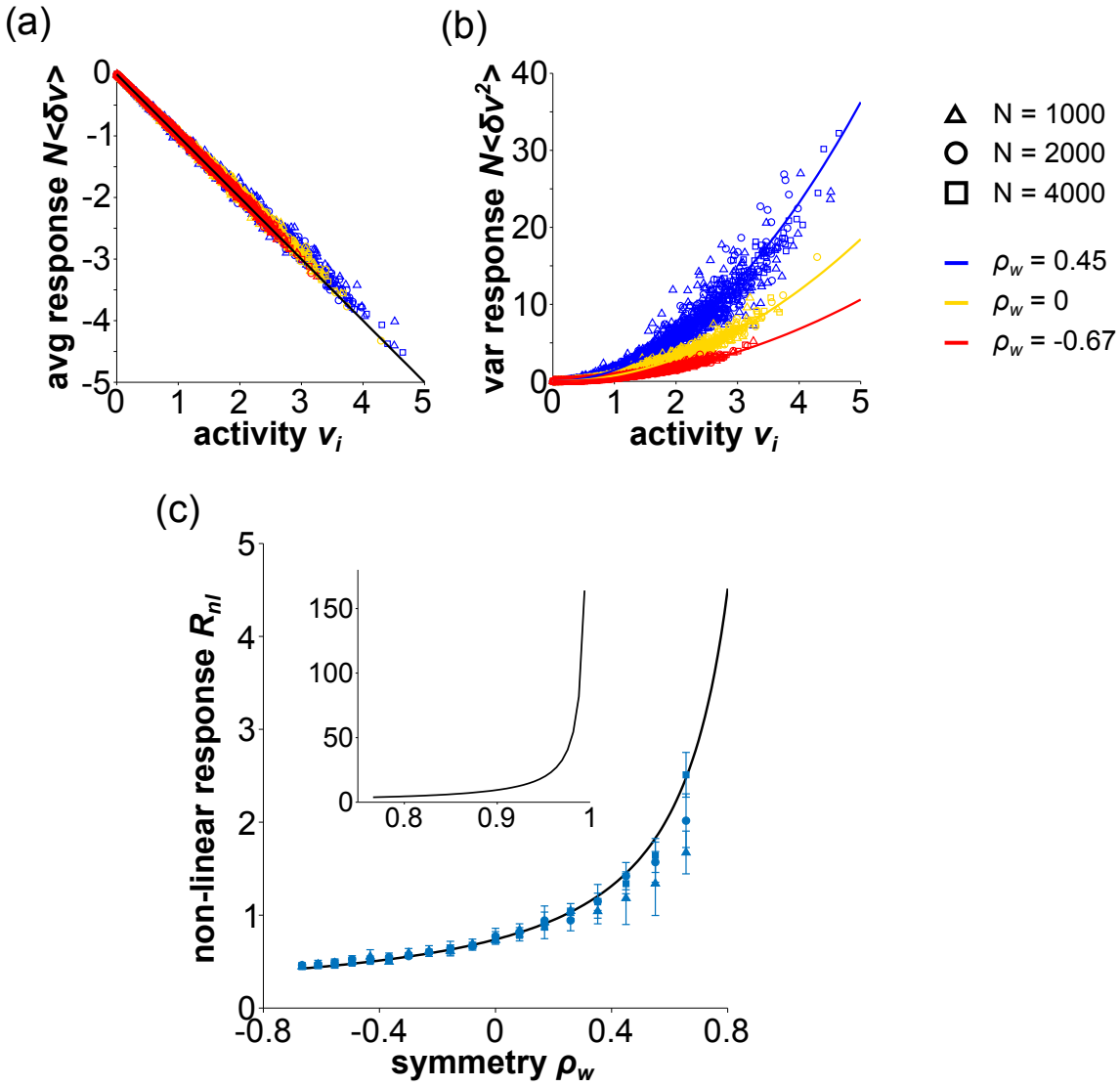


Figure 8.6: **Stability of the fixed point.** (A) Mean average response due to neuron i in the network. Each color represents a different ρ_w. (B) Variance of $\delta\nu_j$ with respect to the activity of neuron i. (C) Non-linear reaction term R_{nl} with respect to degree of symmetry ρ_w.

In a more systematical way, one can obtain the bifurcation diagram of the network with respect to the two parameters σ and ρ_w (Figure 8.7). It is composed of two regions: the grey one denotes a region where only one fixed point exists. In this region, the dynamics is going to converge to this fixed point. The second region represents a region where there is not one unique fixed point and the cavity-like approach is not valid anymore. In the

following I will give insight about what is happening beyond the bifurcation.

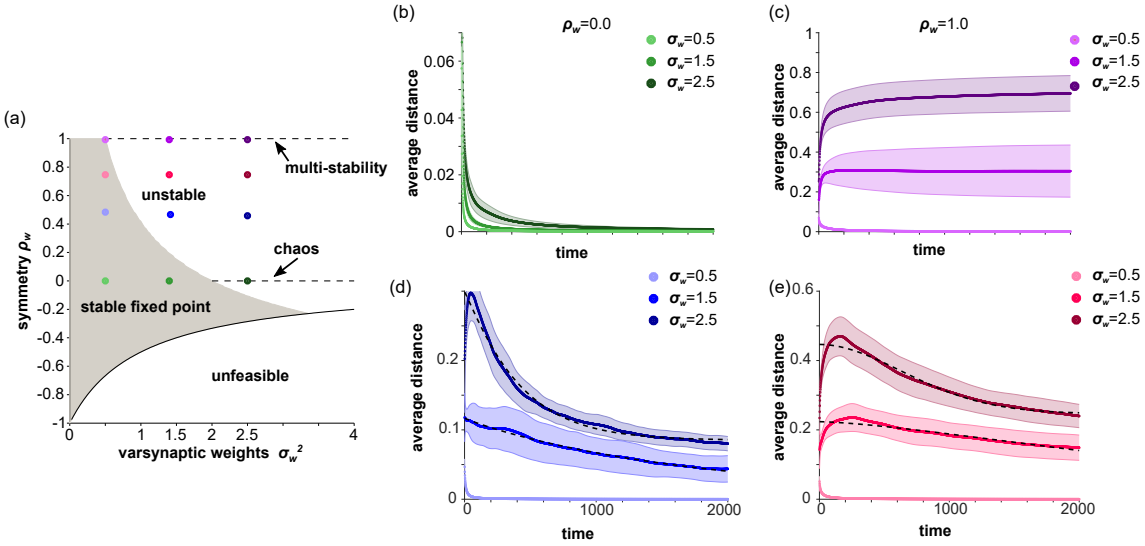


Figure 8.7: **Breaking of ergodicity.** (A) Bifurcation diagram of the network. (B,C,D,E) Measure of ergodicity with respect to time. Each colour is associated to a point in the bifurcation diagram. The dashed black curve corresponds to a fit by a streched exponential function.

8.3.3 Beyond the bifurcation

Small size networks

To understand what happens beyond the bifurcation I will first study a small size network. To do so, I consider a network with $N = 400$ neurons and I will look at the fixed point that is reached for 100 initial conditions (for parameters outside the fixed point region). The results are shown in Figure 8.8. Each color represents a different ρ_w and the circle the mean on five different realizations of the synaptic connectivity. $\sigma^2_{w,c}$ represents the critical σ_w at which a bifurcation appends for the corresponding value of ρ_w.

As σ_w^2 crosses the critical value, the number of different fixed points that are obtained for 100 different initial conditions starts to increase (Figure 8.8.B). Moreover, the maximal eigenvalue increases to 1 at the critical value. This behavior occurs even when ρ_w is decreased. This means that for these different behaviors the steady states that are reached by the dynamics are marginally stable. Beyond the bifurcation, it seems that the network has many different fixed points that are marginally stables.

Breaking of ergodicity

To characterize the behavior of the network, I will introduce a measure of ergodicity. Let's consider two different initial conditions a and b. It is possible to define the distance between these two initial conditions for each neuron:

$$\Delta^{a,b}h_i(t) = \frac{1}{t}\left[\int_0^t ds\, h_i^a(s) - \int_0^t ds\, h_i^b(s)\right] \tag{8.59}$$

The measure for the full network will be:

$$D^2(t) = \frac{1}{N}\sum_i (\Delta^{a,b}h_i(t))^2 \tag{8.60}$$

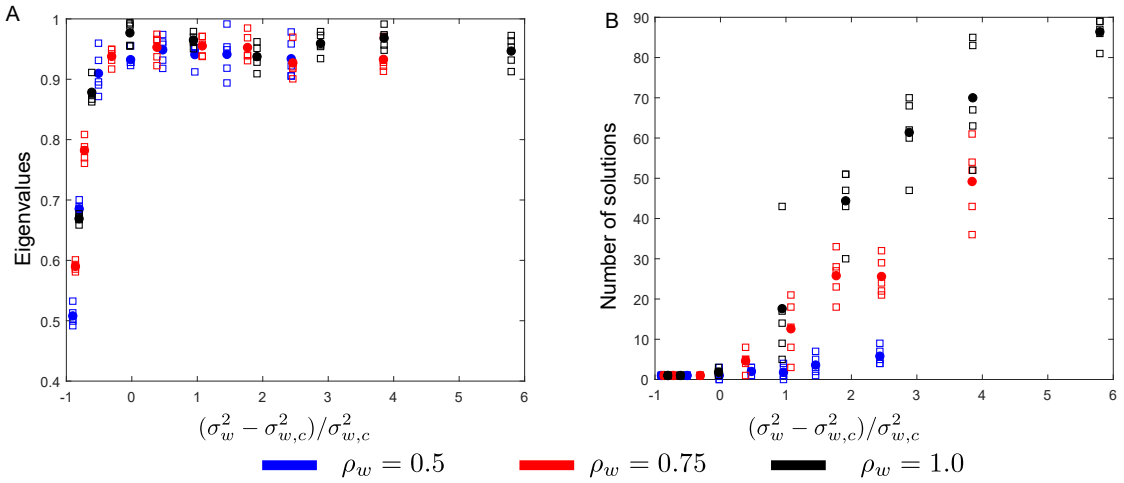


Figure 8.8: **Criticality at the bifurcation.** The squares denotes the results of the simulation for 100 initial conditions. The circle is the mean of 5 different synaptic matrices. (A) Maximal eigenvalues of the fixed point that has been reached by the network. (B) Number of different fixed points that are reached by a network for 100 initial conditions.

Why is this measure a good measure of ergodicity ? If the network is ergodic (such as chaotic or converging to a fixed point), one will have:

$$\lim_{t\to\infty} \Delta^{a,b} h_i(t) = 0 \tag{8.61}$$

Indeed, in an ergodic system the mean activity of each neuron is going to be the same (at infinite time) independently of the initial conditions as the structure of the synaptic connections is fixed. When this distance does not converge to zero, it will mean that the system is not ergodic.

In Figure 8.7, I show the time variations of this measure for different values of parameters for a network of 4000 neurons. First, for $\rho_w = 0$ (Figure 8.7.B) it is known that there are two possible behaviors: either the system converges to a stable fixed point (grey region), either it is chaotic (Doyon et al., 1993). As expected, in Figure 8.7.B the measure $D(t)$ always goes to zero as t increases. This is characteristic of an ergodic system. On the contrary, it is known that at $\rho_w = 1$ the system should exhibit multi-stability. In Figure 8.7.C, $D(t)$ does not go to zero, except for parameters in the region of the fixed point. For lower values of ρ_w, an intermediate regime is observed. For values of parameters that are not within the stable fixed point region $D(t)$ decreases with time. However, this decrease is very long (the time unit is the time constant of the network). To obtain more insight about this behavior, I fitted $D(t)$ by a streched exponential function:

$$f(t) = a \cdot \exp\left(-(t/\tau)^{\beta}\right) + b \tag{8.62}$$

The parameter b is not equal to zero when I fit this function on D for values of parameters for which the system is non-ergodic. This means that, at least for networks of this size and within this time-range, partially symmetric networks can show a breaking of ergodicity depending on the level of correlation between the weights.

9

Conclusion and Perspectives

This thesis presents results on different aspects of decision-making using attractor neural networks. This work has focused on two main parts: understanding the implication of having a non-linear dynamics for the decision-making process and understanding the effects of the neural coding on the learning process of a categorization task.

Researchers have shown that attractor neural networks reproduce, qualitatively, the neurophysiological signatures of decision-making. Howeer, they did not fully compare to behavioral specificities of decision-making, especially with humans. In the first part of this thesis, I have focused on the sequential effects that are observed during perceptual decision-making experiments. The effects are various and go from repetition biases to post-error effects. I have shown that an attractor neural network can be used to study the sequential effects in perceptual decision-making. Moreover, first-order effects result from the intrinsic non-linear dynamics of the network. This result adresses a central question about the sequential effects as they do not reflect an optimal behavior for the participant. I proposed that they are not voluntary and are in fact a constraint due to the non-linear dynamics of the decision network.

History biases and post-error effects have been studied within the framework of drift-diffusion model. With these models the trials are divided into groups, such as repetition trials or alternation trials for example, and a DDM is fitted separately on each group. It has been found that the drift parameter is modified between the different conditions (Urai et al., 2019). This has been proposed as a success of the DDM as the build-up rate of neuronal activity has been found to differ between post-error and post-correct trials for example (Purcell and Kiani, 2016). I have shown that the behavioral observations of sequential effects can be explained by a relaxation mechanism that induces a modification of the starting point in attractor networks. An interesting feature of non-linear systems is that the starting point of the dynamics has an impact on the speed too. It could be very interesting to study this feature and to relate this to the neuronal findings on the build-up rate.

Without any additional memory module, an attractor neural network cannot reproduce the transition between automatic facilitation and strategic expectancy (Gao et al., 2009; Laming, 1968). In the network that I have presented in this manuscript, for too short RSIs (such as a few dozens of milliseconds) the sequential effects are too strong to be plausible. Decision conflict mechanisms (Jones et al., 2002) could be implemented to correct this effect and to investigate other effects of repetitions and alternations. The use of additional memory module accounts for higher-order sequential effects too. Due to the nature of the dynamics in my network, I do not expect to reproduce higher-order sequential effects. In fact, for parameters for which the model exhibits first-order sequential effects, I do not find neither second-order sequential effects, nor post-error adjustments at second-order. One

may ask whether a more complex architecture, taking into account other brain areas, could account for higher-order repetition biases and post-error adjustments effects as resulting from some intrinsic properties of the dynamics, in the absence of specific memory units.

Attractor neural networks are usually compared qualitatively with behavioral data. In this work, I have shown that such models can be fitted on partially behavioral data such as mean reaction times and accuracies. Doing so, I was able to quantitatively compare the attractor network to various behavioral measures. For instance, the sense of confidence that participants have in their decisions, and its various effects, can be modelled by an attractor neural network. Most of the research on confidence during perceptual decision-making experiments has focused on models of confidence that model it as the probability of having made the correct decision. This framework has its successes but does not give insights into the neurophysiological signatures of confidence. In this work, I have shown that a sense of confidence based on an attractor neural network reproduces many effects that were thought to be characteristic of a statistical confidence. Moreover, it has been recently found, using more complex psychophysics tasks, that confidence could not be modelled according to the Bayesian framework. The next step of my work would be to study the specific effects that have been found by Adler and Ma (2018) within the framework of attractor neural networks. This would give insights into specific effects that could be modelled by attractor networks and not within the Bayesian framework.

Another aspect of my thesis was to study the relations between a neural coding and a decision-making network. The motivation was that, in most of the research on this thematic, researchers study the optimality of different systems in a categorization task by considering a population code which has been constructed to discreminate between stimuli and not to categorize them. This is kind of contradictory as the network is tested on a discrimination task but has learned a categorization task. I have presented a study of the effects of the neural coding on the learning process.

The association between the continuum of sensory stimuli and the discrete stimuli can be learned with a trial-and-error mechanism (Law and Gold, 2009; Engel et al., 2015). Different learning algorithms have been proposed to perform this learning, such as reward-modulated Hebbian learning (Legenstein et al., 2008, 2010; Brea and Gerstner, 2016). This type of learning has some experimental grounds as dopamine could act as a neuromodulator and its acticity is correlated with reward signals (Schultz et al., 1997; Schultz, 2002). However, for such learning algorithms, the mean reward needs to be remembered in order to adapt the learning. In this work, I have proposed a learning process that does not need to store the mean reward as the modulation is done through the confidence in each trial.

I have shown that using the sense of confidence during the learning process improves the performances of the network. Moreover it allows for an online learning as it does not require to store the previous rewards. These results have to be compared with recent work (Engel et al., 2015) on categorical learning in an attractor network using top-down feedback. The authors found that the distribution of effective tuning curves was shifted after learning, leading to categorical effects. By mixing their framework and mine it could be possible to theoretically compute the effect of the learning on the coding neurons. Therefore, one could compare the effects and optimality of different learning processes. This will constitute my research direction during the next months.

The last part of my thesis did not focus on decision-making but rather on statistical physics methods to study neural networks. Most of the research has been focusing on random neural networks as they exhibit behaviors that are useful for neural computations and can be analyzed using standard statistical physics methods. However, parameters and structural connectivity are strongly constrained in such networks. Moreover, they do not

correspond to experimental evidences. Here, I propose a new approach to compute the steady-states of non-random inhibitory neural networks using the cavity method. I have shown that a partial degree of symmetry does not necessary lead to a chaotic activity and that a breaking of ergodicity occurs. This work has focused on a unique neural population in order to establish the theory. However, many populations of excitatory and inhibitory neurons interact with each other in the brain. One direct extension to the presented framework would be to consider different neural pools, mixing excitatory and inhibitoy, connected with each other and to adapt the theory to obtain the bifurcation diagram. Such theory has been developped for multi populations random neural networks (Faugeras et al., 2009). This should be feasible with partially connected networks as it will principally increase the number of equations but the hypotheses would still be valid.

IV Appendix

A

Fitting an attractor neural network

For each participant the model is calibrated by fitting both the mean response times and the accuracies for each orientation, this separately for each block. The procedure is shown in Figure A.1.

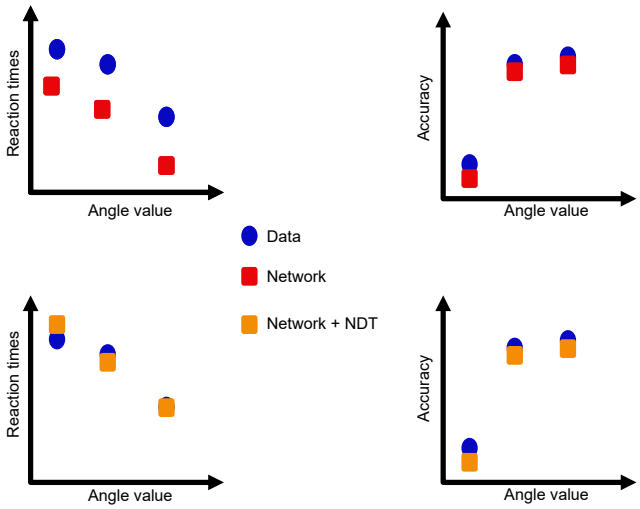


Figure A.1: Schematic cost function of the model. The blue circles correspond to the experimental data. The red squares are the results of the numerical simulations and in orange it corresponds to the model in which the NDT as been added. This addition doesn't change the accuracy part of the cost function.

The idea is to estimate the non-decision time using the numerical simulations for each participant and each block. Using this NDT, the cost function is be the difference with the performances and the response times.

$$\begin{aligned}\text{Cost function} = \lambda \frac{1}{m} \sum_{\theta} \left([\langle RT \rangle_{network}(\theta) - \langle RT \rangle_{network}] - [\langle RT \rangle_{data}(\theta) - \langle RT \rangle_{data}] \right)^2 \\ + \frac{1}{n} \sum_{\theta} \left(\langle accuracy \rangle_{network}(\theta) - \langle accuracy \rangle_{data}(\theta) \right)^2 \end{aligned} \tag{A.1}$$

where the sums are over the orientation values, $\theta = \{0.2°, 0.5°, 0.8°, 1.6°\}$, the brackets $\langle ... \rangle$ design averages, and the normalization factors n (for response times) and m (for the

accuracy) are given by

$$m = \max_{\theta} \left([\langle RT \rangle_{network}(\theta) - \langle RT \rangle_{network}] - [\langle RT \rangle_{data}(\theta) - \langle RT \rangle_{data}] \right)^2$$
$$n = \max_{\theta} \left[\langle accuracy \rangle_{network}(\theta) - \langle accuracy \rangle_{data}(\theta) \right]^2$$

In this expression, $\langle RT \rangle_{data}(\theta)$ denotes the mean experimental response time obtained by averaging over all trials at the orientations $\pm\theta$, $\langle RT \rangle_{data}$ is the average over all orientations; $\langle RT \rangle_{network}(\theta)$ and $\langle RT \rangle_{network}$ are the corresponding averages obtained from the model simulations. The coefficient λ denotes the relative weight given to the response time and accuracy cost terms. The results are obtained when taking $\lambda = 2$, but it should be noted that the choice of this parameter does not impact drastically the fitted parameters.

The fitting procedure consists in a Monte Carlo Markov Chain procedure. At each step, I choose M sets of parameters randomly within a certain distance of an initial set of parameters P_0. The cost function is evaluated for each of these parameters. The next step is performed with a new P_0 that corresponds to the set of parameters that had the lowest cost function.

1.1. Fitting the parameters

For most model parameters I take the value used in a previous study (Berlemont and Nadal, 2019), as reproduced below. For the models calibration I consider $I_{CD,max}, \tau_{CD}, c_\theta$ and z as free parameters. I impose the two parameters $I_{CD,max}$ and τ_{CD} to be common to all participants (joint optimization). I optimize the parameters c_θ (one for each orientation value) and the decision threshold z across subjects and blocks.

The rationale for this choice of free parameters is as follows. To avoid overfit, one has to restrict as much as possible the number of free parameters. I rely on the model calibrations done in previous works (Wang, 2002; Wong and Wang, 2006; Berlemont and Nadal, 2019) which suggest to keep as much as possible the parameters values resulting from the initial work of Wong and Wang. In particular, the original parameters values were chosen such as to reproduce empirical data with the mean field model. Now since the empirical data are only behavioral data, it is difficult to make a calibration of the synaptic weights. A significant change of these parameters would be required to change the behavioral outcomes. Importantly, I tried to restrict the calibration to a small set of reasonably independent parameters. For instance, a change in the weights values may be compensated by a change in the decision threshold (so that the cost function may be flat on a large domain of the parameters space). With the weights fixed, I can optimize the fit with respect to the decision threshold in a safer way. An important quantity is the signal-to-noise ratio. By keeping the internal noise constant during the fitting procedure, I explore the whole range of this ratio. The choice of free parameters can be paralleled with the one made in the DDM framework (drift, threshold and level of noise). This facilitates the comparison with the DDM approach. Finally, imposing some of the free parameters to be common to all participants allows to further reduce the number of free parameters, at a price of a more complex optimization (a partially joint calibration of all the participant-specific networks).

Parameter	**Value**	**Parameter**	**Value**	**Fitted parameter**	**Status**
a	270 Hz/nA	σ_{noise}	0.02 nA	$I_{CD,max}$	common to all participants
b	108 Hz	τ_{noise}	2 mS	τ_{CD}	common to all participants
d	0.154 s	I_0	0.3255 nA	decision threshold z	specific to each participant
γ	0.641	J_{ext}	0.182 nA	$c_\theta = \alpha_1\theta + \alpha_2\theta^2$	specific to each participant
$J_{C,C} = J_{AC,AC}$	0.2609 nA	$J_{C,AC} = J_{AC,C}$	0.0497 nA		
τ_S	100 ms				

For each subject, I minimize the cost function with respect to the choice of c_θ and z, making use of a Monte Carlo Markov Chain fitting procedure, coupled to a subplex procedure (Rowan, 1990). This method is particularly adapted to handle simulation based models with stochastic dynamics. Finally, $I_{CD,max}$ and τ_{CD} are fitted using a grid search algorithm as they have less influence on the cost function. In the model, the parameter c represents the stimulus ambiguity, which I expect here to be a monotonous function of the amplitude of the angle, θ. When allowed to be independent parameter values for each value of the orientation, $\theta = \{0.2°, 0.5°, 0.8°, 1.6°\}$, I find that the c_θ values can be approximated by a linear or quadratic function of θ depending on the participant. I performed an AIC test between the linear and quadratic fit in order to choose which function to use for each participant. These approximations reduce the number of free parameters.

In order to obtain a confidence interval for the different parameters, I used the likelihood estimation of confidence interval for Monte-Carlo Markov Chains method. The confidence interval on the parameters is thus the 70% confidence interval, assuming a Gaussian distribution of the cost function. This provides an approximation of the reliability of the parameters values found. In order to assess the reliability of this method I checked that the threshold z and stimulus strength c_θ parameters have an almost non-correlated influence onto the cost function.
The results of the calibrating procedure are summarized in the following section, with $I_{CD,max} = 0.033$ nA and $\tau_{CD} = 150$ ms.

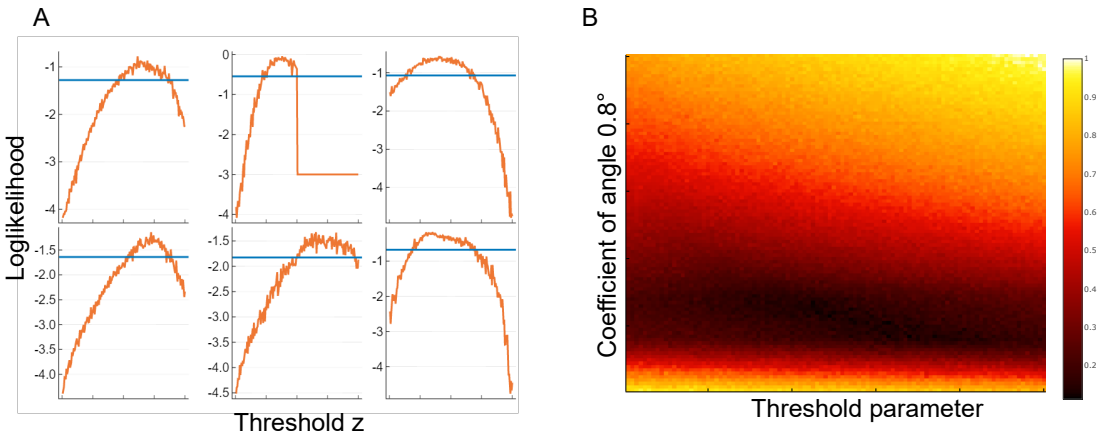


Figure A.2: (A) Each panel represents the log-likelihood of the cost function for each participant when the parameter z is varied. The blue line shows the confidence interval. (B) The color code represents the normalized value of the cost function for participant 5 when two parameters are varied. The minimal value of the cost function corresponds to a line. This means that these two parameters are almost non-correlated.

1.2. Parameters of the model

Block	Parameter	Participant 1	Participant 2	Participant 3	Participant 4	Participant 5	Participant 6
Pure Block	$z(Hz)$	10.78	12.69	14.07	12.80	10.05	12.89
	$\Delta z(Hz)$	(-0.7,+1.75)	(-2.1,+0.175)	(-1.92,+1.75)	(-0.1,+2.275)	(-1.8,2.1)	(-1.05,+2.45)
Confidence	$z(Hz)$	13.08	13.70	14.95	12.96	12.55	14.65
Block	Δz (Hz)	(-0.18,+2.28)	(-1.93,+0.18)	(-2.45,+1.40)	(-0.17,+2.1)	(-0.35,+2.8)	(-2.1,+0.53)

	Type of fit	**Pure Block**	**Confidence Block**
Participant 1	Quadratic	$c_\theta = 554.8 * \theta - 1444 * \theta^2$	$c_\theta = 650 * \theta 1.3 \times 10^4 * \theta^2$
Participant 2	Quadratic	$c_\theta = 729.6 * \theta - 9819 * \theta^2$	$c_\theta = 1056 * \theta - 1.4 \times 10^4 * \theta^2$
Participant 3	Linear	$c_\theta = 524.4 * \theta$	$c_\theta = 634.6 * \theta$
Participant 4	Quadratic	$c_\theta = 269.8 * \theta - 577.6 * \theta^2$	$c_\theta = 190 * \theta + 2.17 \times 10^4 * \theta^2$
Participant 5	Quadratic	$c_\theta = 387.6 * \theta + 4188 * \theta^2$	$c_\theta = 182.4 * \theta + 4.9 \times 10^4 * \theta^2$
Participant 6	Linear	$c_\theta = 551 * \theta$	$c_\theta = 1030 * \theta$

Usher-McClelland model

Parameter	**Subject 1**	**Subject 2**	**Subject 3**	**Subject 4**	**Subject 5**	**Subject 6**
z	1.0	1.0	1.0	1.4	1.3	1.3
β	0.25	0.10	0.18	0.10	0.15	0.12
k	0.15	0.18	0.18	0.11	0.11	0.14
$c_{0.2}$	0.02	0.04	0.02	0.02	0.02	0.04
$c_{0.8}$	0.15	0.12	0.07	0.08	0.14	0.132
$c_{1.6}$	0.23	0.20	0.17	0.225	0.295	0.235

B

Collapsing bounds models

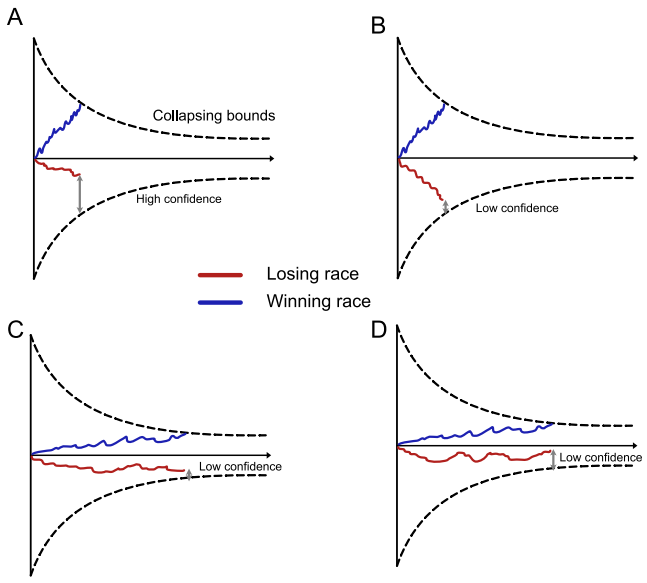


Figure B.1: **IRM collapsing bounds.** Each panel represents a different case in the IRM with collapsing bounds. The blue race corresponds to the winning race and the blue one to the losing race. (A) This panel corresponds to the case where the decision was fast and lead to a high confidence trial. (B) Fast decision but low confidence trial. (C) Slow decision and losing race very close to the decision threshold. (D) Slow decision and losing race far to the decision threshold.

Figure B.1 shows a schematic representation of the IRM with collapsing bounds. The same analysis as the one in the main text for the IRM without collapsing bounds leads to the following conclusions. One can observe two behaviors. Either the decision is fast and it is possible for the model to be in a high-confidence trial, either the decision is slow and the confidence will be low.

If the decision was fast, the model will lead to the opposite sequential effect due to confidence after the relaxation period (similar case as in the main text). However, if the decision was slow the losing race is going to be close to the decision threshold and the confidence will be low for all of these trials. After the relaxation, the losing race will have the same state as in the high-confidence fast trial case. Therefore there will be no sequential effects. This analysis shows that even an IRM with collapsing bounds and a relaxation mechanism will not account for the sequential effects due to confidence that I

can model with an attractor neural network.

Bibliography

Abbott, L. and F. S. Chance
2005. Drivers and modulators from push-pull and balanced synaptic input. *Progress in brain research*, 149:147–155.

Abbott, L. and C. van Vreeswijk
1993. Asynchronous states in networks of pulse-coupled oscillators. *Physical Review E*, 48(2):1483.

Abraham, N. M., H. Spors, A. Carleton, T. W. Margrie, T. Kuner, and A. T. Schaefer
2004. Maintaining accuracy at the expense of speed: stimulus similarity defines odor discrimination time in mice. *Neuron*, 44(5):865–876.

Adler, W. T. and W. J. Ma
2018. Comparing bayesian and non-bayesian accounts of human confidence reports. *PLoS computational biology*, 14(11):e1006572.

Alemi, A., C. Baldassi, N. Brunel, and R. Zecchina
2015. A three-threshold learning rule approaches the maximal capacity of recurrent neural networks. *PLoS computational biology*, 11(8).

Amari, S.-I.
1972. Characteristics of random nets of analog neuron-like elements. *IEEE Transactions on Systems, Man, and Cybernetics*, (5):643–657.

Amit, D. J., M. Evans, and M. Abeles
1990. Attractor neural networks with biological probe records. *Network: Computation in Neural Systems*, 1(4):381–405.

Baranski, J. V. and W. M. Petrusic
1994. The calibration and resolution of confidence in perceptual judgments. *Perception & psychophysics*, 55(4):412–428.

Barbier, M. and J.-F. Arnoldi
2017. The cavity method for community ecology. *bioRxiv*, P. 147728.

Bates, D., M. Mächler, B. Bolker, and S. Walker
2014. Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

Bates, D., M. Mächler, B. Bolker, and S. Walker
2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48.

Beck, J. M., W. Ma, R. Kiani, T. Hanks, A. K. Churchland, J. Roitman, M. N. Shadlen, P. E. Latham, and A. Pouget
2008a. Probabilistic population codes for bayesian decision making. 60(6):1142–1152.

Beck, J. M., W. J. Ma, R. Kiani, T. Hanks, A. K. Churchland, J. Roitman, M. N. Shadlen, P. E. Latham, and A. Pouget
2008b. Probabilistic population codes for bayesian decision making. *Neuron*, 60(6):1142–1152.

Behrens, T. E., M. W. Woolrich, M. E. Walton, and M. F. Rushworth
2007. Learning the value of information in an uncertain world. *Nature neuroscience*, 10(9):1214–1221.

Bennur, S. and J. I. Gold
2011. Distinct representations of a perceptual decision and the associated oculomotor plan in the monkey lateral intraparietal area. *Journal of Neuroscience*, 31(3):913–921.

Berlemont, K., J.-R. Martin, J. Sackur, and J.-P. Nadal
2020. Nonlinear neural network dynamics accounts for human confidence in a sequence of perceptual decisions. *Scientific reports*, 10(1):1–16.

Berlemont, K. and J.-P. Nadal
2019. Perceptual decision-making: Biases in post-error reaction times explained by attractor network dynamics. *Journal of Neuroscience*, 39(5):833–853.

Bertelson, P.
1961. Sequential redundancy and speed in a serial two-choice responding task. *Quarterly Journal of Experimental Psychology*, 13(2):90–102.

Bertschinger, N. and T. Natschläger
2004. Real-time computation at the edge of chaos in recurrent neural networks. *Neural computation*, 16(7):1413–1436.

Bliss, D. P., J. J. Sun, and M. D'Esposito
2017. Serial dependence is absent at the time of perception but increases in visual working memory. *Scientific reports*, 7(1):1–13.

Bogacz, R.
2007. Optimal decision-making theories: linking neurobiology with behaviour. *Trends in cognitive sciences*, 11(3):118–125.

Bogacz, R., E. Brown, J. Moehlis, P. Holmes, and J. D. Cohen
2006. The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological review*, 113(4):700.

Bonaiuto, J. J., A. de Berker, and S. Bestmann
2016. Response repetition biases in human perceptual decisions are explained by activity decay in competitive attractor models. *Elife*, 5:e20047.

Bonnasse-Gahot, L. and J.-P. Nadal
2008. Neural coding of categories: information efficiency and optimal population codes. *Journal of computational neuroscience*, 25(1):169–187.

Bonnasse-Gahot, L. and J.-P. Nadal
2012. Perception of categories: from coding efficiency to reaction times. *Brain research*, 1434:47–61.

Brainard, D. H.
1997. The psychophysics toolbox. *Spatial vision*, 10(4):433–436.

Braun, A., A. E. Urai, and T. H. Donner
2018. Adaptive history biases result from confidence-weighted accumulation of past choices. *Journal of Neuroscience*, Pp. 2189–17.

Brea, J. and W. Gerstner
2016. Does computational neuroscience need new synaptic learning paradigms? *Current opinion in behavioral sciences*, 11:61–66.

Britten, K. H., M. N. Shadlen, W. T. Newsome, and J. A. Movshon
1992. The analysis of visual motion: a comparison of neuronal and psychophysical performance. *Journal of Neuroscience*, 12(12):4745–4765.

Britten, K. H., M. N. Shadlen, W. T. Newsome, and J. A. Movshon
1993. Responses of neurons in macaque mt to stochastic motion signals. *Visual neuroscience*, 10(6):1157–1169.

Brody, C. D. and T. D. Hanks
2016. Neural underpinnings of the evidence accumulator. *Current opinion in neurobiology*, 37:149–157.

Brouwer, G. J. and D. J. Heeger
2013. Categorical clustering of the neural representation of color. *Journal of Neuroscience*, 33(39):15454–15465.

Brunel, N.
2000. Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *Journal of computational neuroscience*, 8(3):183–208.

Brunel, N.
2016. Is cortical connectivity optimized for storing information? *Nature neuroscience*, 19(5):749.

Brunel, N. and X.-J. Wang
2001. Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. *Journal of computational neuroscience*, 11(1):63–85.

Brunton, B. W., M. M. Botvinick, and C. D. Brody
2013. Rats and humans can optimally accumulate evidence for decision-making. *Science*, 340(6128):95–98.

Cessac, B.
1995. Increase in complexity in random neural networks. *Journal de Physique I*, 5(3):409–432.

Cho, R. Y., L. E. Nystrom, E. T. Brown, A. D. Jones, T. S. Braver, P. J. Holmes, and J. D. Cohen
2002. Mechanisms underlying dependencies of performance on stimulus history in a

two-alternative forced-choice task. *Cognitive, Affective, & Behavioral Neuroscience*, 2(4):283–299.

Clarke, F. R., T. G. Birdsall, and W. P. Tanner Jr
1959. Two types of roc curves and definitions of parameters. *The Journal of the Acoustical Society of America*, 31(5):629–630.

Cohen, M. X., S. Van Gaal, K. R. Ridderinkhof, and V. Lamme
2009. Unconscious errors enhance prefrontal-occipital oscillatory synchrony. *Frontiers in human neuroscience*, 3:54.

Compte, A., N. Brunel, P. S. Goldman-Rakic, and X.-J. Wang
2000. Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cerebral cortex*, 10(9):910–923.

Cover, T. M. and J. A. Thomas
2012. *Elements of information theory*. John Wiley & Sons.

Crapse, T. B. and M. A. Sommer
2008. Corollary discharge across the animal kingdom. *Nature Reviews Neuroscience*, 9(8):587–600.

Crapse, T. B. and M. A. Sommer
2009. Frontal eye field neurons with spatial representations predicted by their subcortical input. *Journal of Neuroscience*, 29(16):5308–5318.

Crisanti, A., H. Sommers, and H. Sompolinsky
1990. Chaos in neural networks: chaotic solutions. *preprint*.

Crisanti, A. and H. Sompolinsky
1987. Dynamics of spin systems with randomly asymmetric bonds: Langevin dynamics and a spherical model. *Physical Review A*, 36(10):4922.

Cromer, J. A., J. E. Roy, and E. K. Miller
2010. Representation of multiple, independent categories in the primate prefrontal cortex. *Neuron*, 66(5):796–807.

Danielmeier, C., T. Eichele, B. U. Forstmann, M. Tittgemeyer, and M. Ullsperger
2011. Posterior medial frontal cortex activity predicts post-error adaptations in task-related visual and motor areas. *Journal of Neuroscience*, 31(5):1780–1789.

Danielmeier, C. and M. Ullsperger
2011. Post-error adjustments. *Frontiers in psychology*, 2:233.

Danielmeier, C., J. R. Wessel, M. Steinhauser, and M. Ullsperger
2009. Modulation of the error-related negativity by response conflict. *Psychophysiology*, 46(6):1288–1298.

de Gardelle, V. and P. Mamassian
2014. Does confidence use a common currency across two visual tasks? *Psychological science*, 25(6):1286–1288.

Debener, S., M. Ullsperger, M. Siegel, K. Fiehler, D. Y. Von Cramon, and A. K. Engel
2005. Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. *Journal of Neuroscience*, 25(50):11730–11737.

Deco, G., A. Ponce-Alvarez, D. Mantini, G. L. Romani, P. Hagmann, and M. Corbetta
2013. Resting-state functional connectivity emerges from structurally and dynamically shaped slow linear fluctuations. *Journal of Neuroscience*, 33(27):11239–11252.

Desender, K., A. Boldt, T. Verguts, and T. H. Donner
2018a. Post-decisional sense of confidence shapes speed-accuracy tradeoff for subsequent choices. *bioRxiv*, P. 466730.

Desender, K., P. R. Murphy, A. Boldt, T. Verguts, and N. Yeung
2018b. A post-decisional neural marker of confidence predicts information-seeking. *bioRxiv*, P. 433276.

Ding, L. and O. Hikosaka
2006. Comparison of reward modulation in the frontal eye field and caudate of the macaque. *Journal of Neuroscience*, 26(25):6695–6703.

Ditterich, J.
2006. Evidence for time-variant decision making. *European Journal of Neuroscience*, 24(12):3628–3641.

Ditterich, J., M. E. Mazurek, and M. N. Shadlen
2003. Microstimulation of visual cortex affects the speed of perceptual decisions. *Nature neuroscience*, 6(8):891–898.

Doyon, B., B. Cessac, M. Quoy, and M. Samuelides
1993. Control of the transition to chaos in neural networks with random connectivity. *International Journal of Bifurcation and Chaos*, 3(02):279–291.

Drugowitsch, J., A. G. Mendonça, Z. F. Mainen, and A. Pouget
2019. Learning optimal decisions with confidence. *Proceedings of the National Academy of Sciences*, 116(49):24872–24880.

Dutilh, G., D. van Ravenzwaaij, S. Nieuwenhuis, H. L. van der Maas, B. U. Forstmann, and E.-J. Wagenmakers
2012. How to measure post-error slowing: a confound and a simple solution. *Journal of Mathematical Psychology*, 56(3):208–216.

Efron, B. and R. J. Tibshirani
1994. *An introduction to the bootstrap*. CRC press.

Engel, T. A., W. Chaisangmongkon, D. J. Freedman, and X.-J. Wang
2015. Choice-correlated activity fluctuations underlie learning of neuronal category representation. *Nature communications*, 6(1):1–12.

Engel, T. A. and X.-J. Wang
2011. Same or different? a neural circuit mechanism of similarity-based pattern match decision making. *Journal of Neuroscience*, 31(19):6982–6996.

Etcoff, N. L. and J. J. Magee
1992. Categorical perception of facial expressions. *Cognition*, 44(3):227–240.

Farrell, S. and C. J. Ludwig
2008. Bayesian and maximum likelihood estimation of hierarchical response time models. *Psychonomic bulletin & review*, 15(6):1209–1217.

Faugeras, O. and J. Maclaurin
2015. Asymptotic description of neural networks with correlated synaptic weights. *Entropy*, 17(7):4701–4743.

Faugeras, O. D., J. D. Touboul, and B. Cessac
2009. A constructive mean-field analysis of multi population neural networks with random synaptic weights and stochastic inputs. *Frontiers in computational neuroscience*, 3:1.

Fay, M. P. and M. A. Proschan
2010. Wilcoxon-mann-whitney or t-test? on assumptions for hypothesis tests and multiple interpretations of decision rules. *Statistics surveys*, 4:1.

Fetsch, C. R., R. Kiani, and M. N. Shadlen
2014. Predicting the accuracy of a decision: a neural mechanism of confidence. In *Cold Spring Harbor symposia on quantitative biology*, volume 79, Pp. 185–197. Cold Spring Harbor Laboratory Press.

Finkel, A. S. and S. J. Redman
1983. The synaptic current evoked in cat spinal motoneurones by impulses in single group 1a axons. *The Journal of Physiology*, 342(1):615–632.

Fitzgerald, J. K., D. J. Freedman, and J. A. Assad
2011. Generalized associative representations in parietal cortex. *Nature neuroscience*, 14(8):1075.

Fleming, S. M. and H. C. Lau
2014. How to measure metacognition. *Frontiers in human neuroscience*, 8:443.

Fleming, S. M., R. S. Weil, Z. Nagy, R. J. Dolan, and G. Rees
2010. Relating introspective accuracy to individual differences in brain structure. *Science*, 329(5998):1541–1543.

Freedman, D. J. and J. A. Assad
2006. Experience-dependent representation of visual categories in parietal cortex. *Nature*, 443(7107):85–88.

Freedman, D. J., M. Riesenhuber, T. Poggio, and E. K. Miller
2001. Categorical representation of visual stimuli in the primate prefrontal cortex. *Science*, 291(5502):312–316.

Freeman, W. J.
1995. The hebbian paradigm reintegrated: local reverberations as internal representations. *Behavioral and brain sciences*, 18(4):631–631.

Frémaux, N., H. Sprekeler, and W. Gerstner
2010. Functional requirements for reward-modulated spike-timing-dependent plasticity. *Journal of Neuroscience*, 30(40):13326–13337.

Fründ, I., F. A. Wichmann, and J. H. Macke
2014. Quantifying the effect of intertrial dependence on perceptual decisions. *Journal of vision*, 14(7):9–9.

Funahashi, S., C. J. Bruce, and P. S. Goldman-Rakic
1989. Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *Journal of neurophysiology*, 61(2):331–349.

Galvin, S. J., J. V. Podd, V. Drga, and J. Whitmore
2003. Type 2 tasks in the theory of signal detectability: Discrimination between correct and incorrect decisions. *Psychonomic bulletin & review*, 10(4):843–876.

Ganguli, D. and E. Simoncelli
2014. Efficient sensory coding and bayesian decoding with neural populations. neural computation.

Ganguli, S., J. W. Bisley, J. D. Roitman, M. N. Shadlen, M. E. Goldberg, and K. D. Miller
2008. One-dimensional dynamics of attention and decision making in lip. *Neuron*, 58(1):15–25.

Gao, J., K. Wong-Lin, P. Holmes, P. Simen, and J. D. Cohen
2009. Sequential effects in two-choice reaction time tasks: decomposition and synthesis of mechanisms. *Neural Computation*, 21(9):2407–2436.

Gehring, W. J. and D. E. Fencsik
2001. Functions of the medial frontal cortex in the processing of conflict and errors. *Journal of Neuroscience*, 21(23):9430–9437.

Gelman, A. and J. Hill
2007. Data analysis using regression and hierarchical/multilevel models. *New York, NY: Cambridge*.

Geman, S. and C.-R. Hwang
1982. A chaos hypothesis for some large systems of random equations. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 60(3):291–314.

Gherman, S. and M. G. Philiastides
2015. Neural representations of confidence emerge from the process of decision formation during perceptual choices. *Neuroimage*, 106:134–143.

Gnadt, J. W. and R. A. Andersen
1988. Memory related motor planning activity in posterior parietal cortex of macaque. *Experimental brain research*, 70(1):216–220.

Gold, J. I., C.-T. Law, P. Connolly, and S. Bennur
2008. The relative influences of priors and sensory evidence on an oculomotor decision variable during perceptual learning. *Journal of neurophysiology*, 100(5):2653–2668.

Gold, J. I. and M. N. Shadlen
2007. The neural basis of decision making. *Annual review of neuroscience*, 30.

Goldfarb, S., K. Wong-Lin, M. Schwemmer, N. E. Leonard, and P. Holmes
2012. Can post-error dynamics explain sequential reaction time patterns? *Frontiers in Psychology*, 3:213.

Goldman-Rakic, P. S.
1995. Cellular basis of working memory. *Neuron*, 14(3):477–485.

Gonzalez, R. C., R. E. Woods, et al.
2002. Digital image processing.

Green, D. M., J. A. Swets, et al.
1966. *Signal detection theory and psychophysics*, volume 1. Wiley New York.

Griffin, D. and A. Tversky
1992. The weighing of evidence and the determinants of confidence. *Cognitive psychology*, 24(3):411–435.

Guo, Z. V., N. Li, D. Huber, E. Ophir, D. Gutnisky, J. T. Ting, G. Feng, and K. Svoboda
2014. Flow of cortical activity underlying a tactile decision in mice. *Neuron*, 81(1):179–194.

Hall, W. and A. Moschovakis
2004. The superior colliculus: new approaches for studying sensorimotor integration (crc, boca raton, fl).

Hampton, R. R.
2001. Rhesus monkeys know when they remember. *Proceedings of the National Academy of Sciences*, 98(9):5359–5362.

Hanes, D. P. and J. D. Schall
1996. Neural control of voluntary movement initiation. *Science*, 274(5286):427–430.

Hanks, T. D., J. Ditterich, and M. N. Shadlen
2006. Microstimulation of macaque area lip affects decision-making in a motion discrimination task. *Nature neuroscience*, 9(5):682–689.

Hansel, D., G. Mato, and C. Meunier
1995. Synchrony in excitatory neural networks. *Neural computation*, 7(2):307–337.

Harris, K. D. and T. D. Mrsic-Flogel
2013. Cortical connectivity and sensory coding. *Nature*, 503(7474):51–58.

Hernández, A., A. Zainos, and R. Romo
2002. Temporal evolution of a decision-making process in medial premotor cortex. *Neuron*, 33(6):959–972.

Hester, R., J. J. Foxe, S. Molholm, M. Shpaner, and H. Garavan
2005. Neural mechanisms involved in error processing: a comparison of errors made with and without awareness. *Neuroimage*, 27(3):602–608.

Hick, W. E.
1952. On the rate of gain of information. *Quarterly Journal of experimental psychology*, 4(1):11–26.

Hillyard, S. A., K. C. Squires, J. W. Bauer, and P. H. Lindsay
1971. Evoked potential correlates of auditory signal detection. *Science*, 172(3990):1357–1360.

Huk, A. and M. L. Meister
2012. Neural correlates and neural computations in posterior parietal cortex during perceptual decision-making. *Frontiers in integrative neuroscience*, 6:86.

Huk, A. C. and M. N. Shadlen
2005. Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. *Journal of Neuroscience*, 25(45):10420–10436.

Jaramillo, J., J. F. Mejias, and X.-J. Wang
2019. Engagement of pulvino-cortical feedforward and feedback pathways in cognitive computations. *Neuron*, 101(2):321–336.

Jentzsch, I. and C. Dudschig
2009. Short article: Why do we slow down after an error? mechanisms underlying the effects of posterror slowing. *Quarterly Journal of Experimental Psychology*, 62(2):209–218.

Jentzsch, I. and W. Sommer
2002. Functional localization and mechanisms of sequential effects in serial reaction time tasks. *Perception & psychophysics*, 64(7):1169–1188.

Jepma, M., P. R. Murphy, M. R. Nassar, M. Rangel-Gomez, M. Meeter, and S. Nieuwenhuis
2016. Catecholaminergic regulation of learning rate in a dynamic environment. *PLoS computational biology*, 12(10).

Jones, A. D., R. Y. Cho, L. E. Nystrom, J. D. Cohen, and T. S. Braver
2002. A computational model of anterior cingulate function in speeded response tasks: Effects of frequency, sequence, and conflict. *Cognitive, Affective, & Behavioral Neuroscience*, 2(4):300–317.

Kadmon, J. and H. Sompolinsky
2015. Transition to chaos in random neuronal networks. *Physical Review X*, 5(4):041030.

Katz, L. N., J. L. Yates, J. W. Pillow, and A. C. Huk
2016. Dissociated functional significance of decision-related activity in the primate dorsal stream. *Nature*, 535(7611):285–288.

Kepecs, A. and Z. F. Mainen
2012. A computational framework for the study of confidence in humans and animals. *Philosophical Transactions of the Royal Society B Biological Sciences*, 367(1594):1322–1337.

Kepecs, A., N. Uchida, H. A. Zariwala, and Z. F. Mainen
2008. Neural correlates, computation and behavioural impact of decision confidence. *Nature*, 455(7210):227–231.

Kiani, R., T. D. Hanks, and M. N. Shadlen
2008. Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. *Journal of Neuroscience*, 28(12):3017–3029.

Kiani, R. and M. N. Shadlen
2009. Representation of confidence associated with a decision by neurons in the parietal cortex. *science*, 324(5928):759–764.

King, J. A., F. M. Korb, D. Y. von Cramon, and M. Ullsperger
2010. Post-error behavioral adjustments are facilitated by activation and suppression of task-relevant and task-irrelevant information processing. *Journal of Neuroscience*, 30(38):12759–12769.

Kirby, N. H.
1972. Sequential effects of serial reaction time. *Journal of Experimental Psychology*, 96(1):32.

Kirby, N. H.
1976. Sequential effects in two-choice reaction time: Automatic facilitation or subjective expectancy? *Journal of Experimental Psychology: Human perception and performance*, 2(4):567.

Klayman, J., J. B. Soll, C. Gonzalez-Vallejo, and S. Barlas
1999. Overconfidence: It depends on how, what, and whom you ask. *Organizational behavior and human decision processes*, 79(3):216–247.

Komura, Y., A. Nikkuni, N. Hirashima, T. Uetake, and A. Miyamoto
2013. Responses of pulvinar neurons reflect a subject's confidence in visual categorization. *Nature neuroscience*, 16(6):749.

Kreyszig, E.
1979. Advanced engineering mathematics. fourth edi.

Lak, A., G. M. Costa, E. Romberg, A. A. Koulakov, Z. F. Mainen, and A. Kepecs
2014. Orbitofrontal cortex is required for optimal waiting based on decision confidence. *Neuron*, 84(1):190–201.

Laming, D.
1979. Choice reaction performance following an error. *Acta Psychologica*, 43(3):199–224.

Laming, D. R. J.
1968. Information theory of choice-reaction times.

Lauwereyns, J., K. Watanabe, B. Coe, and O. Hikosaka
2002. A neural correlate of response bias in monkey caudate nucleus. *Nature*, 418(6896):413–417.

Law, C.-T. and J. I. Gold
2009. Reinforcement learning can account for associative and perceptual learning on a visual-decision task. *Nature neuroscience*, 12(5):655.

Legenstein, R., S. M. Chase, A. B. Schwartz, and W. Maass
2010. A reward-modulated hebbian learning rule can explain experimentally observed network reorganization in a brain control task. *Journal of Neuroscience*, 30(25):8400–8410.

Legenstein, R., D. Pecevski, and W. Maass
2008. A learning theory for reward-modulated spike-timing-dependent plasticity with application to biofeedback. *PLoS computational biology*, 4(10).

Lemus, L., A. Hernández, R. Luna, A. Zainos, V. Nácher, and R. Romo
2007. Neural correlates of a postponed decision report. *Proceedings of the National Academy of Sciences*, 104(43):17174–17179.

Liberman, A. M., K. S. Harris, H. S. Hoffman, and B. C. Griffith
1957. The discrimination of speech sounds within and across phoneme boundaries. *Journal of experimental psychology*, 54(5):358.

Lichtenstein, S. and B. Fischhoff
1977. Do those who know more also know more about how much they know? *Organizational behavior and human performance*, 20(2):159–183.

Lisi, M., G. Mongillo, and A. Gorea
2018. Discrete confidence levels revealed by sequential decisions. *bioRxiv*, P. 169243.

Litwin-Kumar, A. and B. Doiron
2012. Slow dynamics and high variability in balanced cortical networks with clustered connections. *Nature neuroscience*, 15(11):1498.

Lo, C.-C. and X.-J. Wang
2006. Cortico–basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *Nature neuroscience*, 9(7):956–963.

Loewenstein, Y. and H. S. Seung
2006. Operant matching is a generic outcome of synaptic plasticity based on the covariance between reward and neural activity. *Proceedings of the National Academy of Sciences*, 103(41):15224–15229.

Luce, R. D. et al.
1986. *Response times: Their role in inferring elementary mental organization*, number 8. Oxford University Press on Demand.

Machens, C. K., R. Romo, and C. D. Brody
2005. Flexible control of mutual inhibition: a neural model of two-interval discrimination. *Science*, 307(5712):1121–1124.

Marco-Pallarés, J., E. Camara, T. F. Münte, and A. Rodríguez-Fornells
2008. Neural mechanisms underlying adaptive actions after slips. *Journal of cognitive neuroscience*, 20(9):1595–1610.

Markram, H., J. Lübke, M. Frotscher, A. Roth, and B. Sakmann
1997. Physiology and anatomy of synaptic connections between thick tufted pyramidal neurones in the developing rat neocortex. *The Journal of physiology*, 500(2):409–440.

Martí, D., N. Brunel, and S. Ostojic
2018. Correlations between synapses in pairs of neurons slow down dynamics in randomly connected neural networks. *Physical Review E*, 97(6):062314.

Mazurek, M., J. Roitman, J. Ditterich, and M. Shadlen
2003a. A role for neural integrators in perceptual decision making. *Cereb Cortex*, 13(11):1257–1269.

Mazurek, M. E., J. D. Roitman, J. Ditterich, and M. N. Shadlen
2003b. A role for neural integrators in perceptual decision making. *Cerebral cortex*, 13(11):1257–1269.

Merkle, E. C. and T. Van Zandt
2006. An application of the poisson race model to confidence calibration. *Journal of Experimental Psychology: General*, 135(3):391.

Meyniel, F.
2019. Brain dynamics for confidence-weighted learning. *bioRxiv*, P. 769315.

Meyniel, F. and S. Dehaene
2017. Brain networks for confidence weighting and hierarchical inference during probabilistic learning. *Proceedings of the National Academy of Sciences*, 114(19):E3859–E3868.

Mézard, M. and G. Parisi
2003. The cavity method at zero temperature. *Journal of Statistical Physics*, 111(1-2):1–34.

Mézard, M., G. Parisi, and M. Virasoro
1987. *Spin glass theory and beyond: An Introduction to the Replica Method and Its Applications*, volume 9. World Scientific Publishing Company.

Miller, K. D. and D. J. MacKay
1994. The role of constraints in hebbian learning. *Neural computation*, 6(1):100–126.

Miller, P. and D. B. Katz
2013. Accuracy and response-time distributions for decision-making: linear perfect integrators versus nonlinear attractor-based neural circuits. *Journal of computational neuroscience*, 35(3):261–294.

Moreno-Bote, R.
2010. Decision confidence and uncertainty in diffusion models with partially correlated neuronal integrators. *Neural computation*, 22(7):1786–1811.

Morgan, M. J., A. Mason, and J. Solomon
1997. Blindsight in normal subjects. *Nature*, 385(6615):401–402.

Moynot, O. and M. Samuelides
2002. Large deviations and mean-field theory for asymmetric random recurrent neural networks. *Probability Theory and Related Fields*, 123(1):41–75.

Munoz, D. and J. Schall
2003. The superior colliculus: New approaches for studying sensorimotor integration.

Munoz, D. P., J. H. Fecteau, et al.
2002. Vying for dominance: dynamic interactions control visual fixation and saccadic initiation in the superior colliculus. *Progress in brain research*, 140:3–20.

Nassar, M. R., R. Bruckner, and M. J. Frank
2019. Statistical context dictates the relationship between feedback-related eeg signals and learning. *eLife*, 8.

Nassar, M. R., R. C. Wilson, B. Heasly, and J. I. Gold
2010. An approximately bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *Journal of Neuroscience*, 30(37):12366–12378.

Navajas, J., B. Bahrami, and P. E. Latham
2016. Post-decisional accounts of biases in confidence. *Current Opinion in Behavioral Sciences*, 11:55–60.

Navajas, J., C. Hindocha, H. Foda, M. Keramati, P. E. Latham, and B. Bahrami
2017. The idiosyncratic nature of confidence. *Nature human behaviour*, 1(11):810–818.

Newsome, W. T., K. H. Britten, and J. A. Movshon
1989. Neuronal correlates of a perceptual decision. *Nature*, 341(6237):52–54.

Notebaert, W., F. Houtman, F. Van Opstal, W. Gevers, W. Fias, and T. Verguts
2009. Post-error slowing: an orienting account. *Cognition*, 111(2):275–279.

núñez Castellar, E., S. Kühn, W. Fias, and W. Notebaert
2010. Outcome expectancy and not accuracy determines posterror slowing: Erp support. *Cognitive, Affective, & Behavioral Neuroscience*, 10(2):270–278.

Odoemene, O., S. Pisupati, H. Nguyen, and A. K. Churchland
2018. Visual evidence accumulation guides decision-making in unrestrained mice. *Journal of Neuroscience*, 38(47):10143–10155.

Oja, E.
1982. Simplified neuron model as a principal component analyzer. *Journal of mathematical biology*, 15(3):267–273.

Padoa-Schioppa, C.
2013. Neuronal origins of choice variability in economic decisions. *Neuron*, 80(5):1322–1336.

Park, I. M., M. L. Meister, A. C. Huk, and J. W. Pillow
2014. Encoding and decoding in parietal cortex during sensorimotor decision-making. *Nature neuroscience*, 17(10):1395.

Paz, L., A. Insabato, A. Zylberberg, G. Deco, and M. Sigman
2016. Confidence through consensus: a neural mechanism for uncertainty monitoring. *Scientific reports*, 6:21830.

Peirce, C. S.
1873. On the theory of errors of observation. *Report of the Superintendent of the United States Coast Survey Showing the Progress of the Survey During the Year 1870*, Pp. 220–224.

Peirce, C. S. and J. Jastrow
1884. On small differences in sensation.

Perfect, T. J., E. L. Watson, and G. F. Wagstaff
1993. Accuracy of confidence ratings associated with general knowledge and eyewitness memory. *Journal of Applied Psychology*, 78(1):144.

Persaud, N., P. McLeod, and A. Cowey
2007. Post-decision wagering objectively measures awareness. *Nature neuroscience*, 10(2):257–261.

Pleskac, T. J. and J. R. Busemeyer
2010. Two-stage dynamic signal detection: a theory of choice, decision time, and confidence. *Psychological review*, 117(3):864.

Purcell, B. A. and R. Kiani
2016. Neural mechanisms of post-error adjustments of decision policy in parietal cortex. *Neuron*, 89(3):658–671.

Raab, D. H.
1962. Division of psychology: Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences*, 24(5 Series II):574–590.

Rabbitt, P. and B. Rodgers
1977. What does a man do after he makes an error? an analysis of response programming. *Quarterly Journal of Experimental Psychology*, 29(4):727–743.

Rahnev, D., K. Desender, A. L. Lee, W. T. Adler, D. Aguilar-Lleyda, B. Akdoğan, P. Arbuzova, L. Y. Atlas, F. Balcı, J. W. Bang, et al.
2020. The confidence database. *Nature human behaviour*, Pp. 1–8.

Rao, V., G. C. DeAngelis, and L. H. Snyder
2012. Neural correlates of prior expectations of motion in the lateral intraparietal and middle temporal areas. *Journal of Neuroscience*, 32(29):10063–10074.

Ratcliff, R.
1978. A theory of memory retrieval. *Psychological review*, 85(2):59.

Ratcliff, R., P. Gomez, and G. McKoon
2004. A diffusion model account of the lexical decision task. *Psychological review*, 111(1):159.

Ratcliff, R. and J. N. Rouder
1998. Modeling response times for two-choice decisions. *Psychological science*, 9(5):347–356.

Ratcliff, R. and J. N. Rouder
2000. A diffusion model account of masking in two-choice letter identification. *Journal of Experimental Psychology: Human perception and performance*, 26(1):127.

Ratcliff, R. and P. L. Smith
2004. A comparison of sequential sampling models for two-choice reaction time. *Psychological review*, 111(2):333.

Ratcliff, R. and J. J. Starns
2013. Modeling confidence judgments, response times, and multiple choices in decision making: recognition memory and motion discrimination. *Psychological review*, 120(3):697.

Ratcliff, R., A. Thapar, and G. Mckoon
2003. A diffusion model analysis of the effects of aging on brightness discrimination. *Perception & psychophysics*, 65(4):523–535.

Ratcliff, R. and F. Tuerlinckx
2002. Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic bulletin & review*, 9(3):438–481.

Renart, A., J. De La Rocha, P. Bartho, L. Hollender, N. Parga, A. Reyes, and K. D. Harris
2010. The asynchronous state in cortical circuits. *science*, 327(5965):587–590.

Rizzo, M. L. and G. J. Székely
2016. Energy distance. *wiley interdisciplinary reviews: Computational statistics*, 8(1):27–38.

Roberson, D. and J. Davidoff
2000. The categorical perception of colors and facial expressions: The effect of verbal interference. *Memory & cognition*, 28(6):977–986.

Rodriguez-Fornells, A., A. R. Kurzbuch, and T. F. Münte
2002. Time course of error detection and correction in humans: neurophysiological evidence. *Journal of Neuroscience*, 22(22):9990–9996.

Roitman, J. D. and M. N. Shadlen
2002. Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *Journal of neuroscience*, 22(21):9475–9489.

Rolls, E. T., F. Grabenhorst, and G. Deco
2010. Choice, difficulty, and confidence in the brain. *Neuroimage*, 53(2):694–706.

Romo, R., A. Hernández, and A. Zainos
2004. Neuronal correlates of a perceptual decision in ventral premotor cortex. *Neuron*, 41(1):165–173.

Rowan, T.
1990. *The subplex method for unconstrained optimization.* PhD thesis, PhD thesis, Ph. D. thesis, Department of Computer Sciences, Univ. of Texas.

RStudio Team
2015. *RStudio: Integrated Development Environment for R.* RStudio, Inc., Boston, MA.

Saito, Y. and T. Isa
2003. Local excitatory network and nmda receptor activation generate a synchronous and bursting command from the superior colliculus. *Journal of Neuroscience*, 23(13):5854–5864.

Salhi, J., J. MacLaurin, and S. Toumi
2018. On uniform propagation of chaos. *Stochastics*, 90(1):49–60.

Samaha, J., M. Switzky, and B. R. Postle
2018. Confidence boosts serial dependence in orientation estimation. *bioRxiv*, P. 369140.

Sanders, J. I., B. Hangya, and A. Kepecs
2016. Signatures of a statistical computation in the human sense of confidence. *Neuron*, 90(3):499–506.

Schultz, W.
2002. Getting formal with dopamine and reward. *Neuron*, 36(2):241–263.

Schultz, W., P. Dayan, and P. R. Montague
1997. A neural substrate of prediction and reward. *Science*, 275(5306):1593–1599.

Scudder, C. A., C. R. Kaneko, and A. F. Fuchs
2002. The brainstem burst generator for saccadic eye movements. *Experimental brain research*, 142(4):439–462.

Shadlen, M. N. and W. T. Newsome
1996. Motion perception: seeing and deciding. *Proceedings of the national academy of sciences*, 93(2):628–633.

Shadlen, M. N. and W. T. Newsome
2001. Neural basis of a perceptual decision in the parietal cortex (area lip) of the rhesus monkey. *Journal of neurophysiology*, 86(4):1916–1936.

Siegel, M., T. J. Buschman, and E. K. Miller
2015. Cortical information flow during flexible sensorimotor decisions. *Science*, 348(6241):1352–1355.

Sigala, N., F. Gabbiani, and N. K. Logothetis
2002. Visual categorization and object representation in monkeys and humans. *Journal of Cognitive Neuroscience*, 14(2):187–198.

Smith, J. D., J. Schull, J. Strote, K. McGee, R. Egnor, and L. Erb
1995. The uncertain response in the bottlenosed dolphin (tursiops truncatus). *Journal of Experimental Psychology: General*, 124(4):391.

Smith, J. D., W. E. Shields, J. Schull, and D. A. Washburn
1997. The uncertain response in humans and animals. *Cognition*, 62(1):75–97.

Soetens, E., L. C. Boer, and J. E. Hueting
1985. Expectancy or automatic facilitation? separating sequential effects in two-choice reaction time. *Journal of Experimental Psychology: Human Perception and Performance*, 11(5):598.

Sommer, M. A. and R. H. Wurtz
2002. A pathway in primate brain for internal monitoring of movements. *Science*, 296(5572):1480–1482.

Sommer, M. A. and R. H. Wurtz
2008. Visual perception and corollary discharge. *Perception*, 37(3):408–418.

Sompolinsky, H., A. Crisanti, and H.-J. Sommers
1988. Chaos in random neural networks. *Physical review letters*, 61(3):259.

Song, S., P. J. Sjöström, M. Reigl, S. Nelson, and D. B. Chklovskii
2005. Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS biology*, 3(3).

Stankov, L.
2000. Complexity, metacognition, and fluid intelligence. *Intelligence*, 28(2):121–143.

Stern, M., H. Sompolinsky, and L. Abbott
2014. Dynamics of random neural networks with bistable units. *Physical Review E*, 90(6):062710.

Summerfield, C. and F. P. De Lange
2014. Expectation in perceptual decision making: neural and computational mechanisms. *Nature Reviews Neuroscience*, 15(11):745–756.

Székely, G. J. and M. L. Rizzo
2013. Energy statistics: A class of statistics based on distances. *Journal of statistical planning and inference*, 143(8):1249–1272.

Tajima, C. I., S. Tajima, K. Koida, H. Komatsu, K. Aihara, and H. Suzuki
2016. Population code dynamics in categorical perception. *Scientific reports*, 6:22536.

Tajima, S., K. Koida, C. I. Tajima, H. Suzuki, K. Aihara, and H. Komatsu
2017. Task-dependent recurrent dynamics in visual cortex. *eLife*, 6:e26868.

Tolhurst, D. J., J. A. Movshon, and A. F. Dean
1983. The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision research*, 23(8):775–785.

Uchida, N., A. Kepecs, and Z. F. Mainen
2006. Seeing at a glance, smelling in a whiff: rapid forms of perceptual decision making. *Nature Reviews Neuroscience*, 7(6):485–491.

Uchida, N. and Z. F. Mainen
2003. Speed and accuracy of olfactory discrimination in the rat. *Nature neuroscience*, 6(11):1224–1229.

Urai, A. E., J. W. De Gee, K. Tsetsos, and T. H. Donner
2019. Choice history biases subsequent evidence accumulation. *ELife*, 8.

Usher, M. and J. L. McClelland
2001. The time course of perceptual choice: the leaky, competing accumulator model. *Psychological review*, 108(3):550.

Van Vreeswijk, C. and H. Sompolinsky
1996. Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science*, 274(5293):1724–1726.

Verdonck, S. and F. Tuerlinckx
2016. Factoring out nondecision time in choice reaction time data: Theory and implications. *Psychological review*, 123(2):208.

Verguts, T., W. Notebaert, W. Kunde, and P. Wühr
2011. Post-conflict slowing: cognitive adaptation after conflict processing. *Psychonomic Bulletin & Review*, 18(1):76–82.

Vickers, D.
1970. Evidence for an accumulator model of psychophysical discrimination. *Ergonomics*, 13(1):37–58.

Vickers, D.
1979 (reeditited in 2014). *Decision processes in visual perception*. Academic Press.

Wald, A. and J. Wolfowitz
1948. Optimum character of the sequential probability ratio test. *The Annals of Mathematical Statistics*, Pp. 326–339.

Wang, X.-J.
2002. Probabilistic decision making by slow reverberation in cortical circuits. *Neuron*, 36(5):955–968.

Wang, X.-J.
2008. Decision making in recurrent neuronal circuits. *Neuron*, 60(2):215–234.

Wei, Z. and X.-J. Wang
2015. Confidence estimation as a stochastic process in a neurodynamical system of decision making. *Journal of neurophysiology*, 114(1):99–113.

Wickelgren, W. A.
1977. Speed-accuracy tradeoff and information processing dynamics. *Acta psychologica*, 41(1):67–85.

Wilcoxon, F.
1945. Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6):80.

Wong, K.-F., A. C. Huk, M. N. Shadlen, and X.-J. Wang
2007. Neural circuit dynamics underlying accumulation of time-varying evidence during perceptual decision making. *Frontiers in Computational Neuroscience*, 1:6.

Wong, K.-F. and X.-J. Wang
2006. A recurrent network mechanism of time integration in perceptual decisions. *Journal of Neuroscience*, 26(4):1314–1328.

Xin, Y., L. Zhong, Y. Zhang, T. Zhou, J. Pan, and N.-l. Xu
2019. Sensory-to-category transformation via dynamic reorganization of ensemble structures in mouse auditory cortex. *Neuron*, 103(5):909–921.

Yang, T. and J. H. Maunsell
2004. The effect of perceptual learning on neuronal responses in monkey visual area v4. *Journal of Neuroscience*, 24(7):1617–1626.

Yang, Y., P. Cao, Y. Yang, and S.-R. Wang
2008. Corollary discharge circuits for saccadic modulation of the pigeon visual system. *Nature neuroscience*, 11(5):595.

Yeung, N. and C. Summerfield
2012. Metacognition in human decision-making: confidence and error monitoring. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1594):1310–1321.

Ylinen, S., A. Shestakova, P. Alku, and M. Huotilainen
2005. The perception of phonological quantity based on durational cues by native speakers, second-language users and nonspeakers of finnish. *Language and Speech*, 48(3):313–338.

Zdeborová, L. and F. Krzakala
2010. Generalization of the cavity method for adiabatic evolution of gibbs states. *Physical Review B*, 81(22):224205.

Zhou, Y. and D. J. Freedman
2019. Posterior parietal cortex plays a causal role in perceptual and categorical decisions. *Science*, 365(6449):180–185.

Zizlsperger, L., T. Sauvigny, B. Händel, and T. Haarmeier
2014. Cortical representations of confidence in a visual perceptual decision. *Nature communications*, 5(1):1–13.

Znamenskiy, P. and A. M. Zador
2013. Corticostriatal neurons in auditory cortex drive decisions during auditory discrimination. *Nature*, 497(7450):482–485.

RÉSUMÉ

De nombreuses expériences neurophysiologiques sur les singes et rongeurs ont mis en évidence certains mécanismes neuronaux de la prise de décision. Des signaux neuronaux qui sont corrélés avec des éléments spécifiques de la prise de décision réflètent un processus d'accumulation d'évidence afin de prendre la décision. Dans cette thèse, j'étudie un modèle de réseau neuronal dynamique qui utilise un phénomène d'équilibre en excitation et inhibition pour rendre compte du mécanisme d'accumulation d'évidence. Ce modèle rend compte, qualitativement, des données neurophysiologiques mais a très peu été comparé aux résultats comportementaux d'expériences de prise de décision avec des humains. Lors de la prise de décision, de nombreux effets comportementaux sont observables comme par exemple le ralentissement de la décision après une erreur. Reproduire ces effets est crucial pour comprendre le processus de prise de décisions chez les humains et les animaux. J'explore ces différents effets du point de vue du modèle à attracteur neuronaux. Bien qu'il ne s'agisse pas du cadre le plus utilisé dans l'étude de la prise de décision et ses effets, il permet l'étude de méchanismes biophysiques détaillés.

Dans cette thèse je montre que ce niveau de modélisation ne constitue pas uniquement un simple perfectionnement du cadre standard mais qu'il est essentiel pour certains effets comportementaux. Grâce à un phénomène de relaxation, le réseau attracteur reproduit de nombreux séquentiels comme les biais de répétitions, le ralentissement après-erreur ou encore l'amélioration des performances après-erreur. Dans un second temps j'ai développé une expérience psychophysique qui permet l'étude de la confiance lors de la prise de décision. J'ai montré qu'un réseau attracteur reproduit la notion de confiance des participants, ainsi que les effets séquentiels dues à la confiance. Ces différents résultats montre que la dynamique non-linéaire, caractéristique des réseau attracteurs, est essentielle pour reproduire de nombreux aspcts de la prise de décision.

La dernière partie de ce manuscrit consiste en l'étude du codage neuronal de l'information dans un réseau de prise de décisions. Je me concentre sur le processus d'apprentissage d'une tâche de catégorisation. Je montre que la modulation par la confiance du signal de récompense conduit à un apprentissage plus efficace de la part du système.

MOTS CLÉS

prise de décision perceptuelle, attracteurs neuronaux, effets séquentiels, confiance, codage neuronal

ABSTRACT

Neurophysiological experiments on monkeys and rodents have highlighted the neural mechanisms of decision-making. Neural signals, that are correlated with specific elements of the decision-making process, reflect an accumulation of evidence until the decision is reached. In this thesis, I study a dynamic neural network model that uses a balance between excitation and inhibition to account for the mechanism of evidence accumulation. This model qualitatively accounts for neurophysiological data but has not been compared quantitatively with behavioral results from decision experiments with humans. During decision-making experiments, many behavioral effects can be observed, such as the slowing down of the decision after an error. Modeling these effects is critical to reproduce the decision-making process in animals and humans. I explore these different effects from the point of view of the neural attractor model. Despite the fact that this framework does not consist in the most common to study decision-making and its effects, it allows for detailed biophysical mechanisms.

In this thesis, I show that this level of modeling does not just correspond to a refinement of the standard framework but is essential to reproduce some behavioral measures. Using a relaxation dynamics, the network accounts for many of the sequential effects such as history biases, post-error slowing and post-error improvement in accuracy. In a second step I have developped a psychophysics experiment in order to study confidence in decision-making. I have shown that an attractor network reproduces the sense of confidence of the participants, as well as the sequential effects due to confidence. These results show that the non-linear dynamics, characteristic of attractor neural networks, is essential to reproduce various aspects of decision-making.

The last part of this manuscript consists in a study of the neural coding of information in a decision network. I focus on the learning process of a categorization task by the network. I show that a modulation of the reward signal by the confidence leads to a more efficient learning of the categorization task.

KEYWORDS

perceptual decision-making, attractor neural networks, sequential effects, confidence, neural coding