A biologically inspired model of decision making controlled by random neural activity

Benoit Gaillard

Supervisors:

Prof. Jianfeng Feng, Prof. Hilary Buxton, Dr. Si Wu

Examinors:

Prof. Roman Borisyuk, Dr. David Young

Submitted for the degree of Doctor of Philosophy

University of Sussex

April, 2007

Declaration

I hereby declare that this thesis has not been submitted, either in the same or in different form to this or any other university for a degree.

Signature:

Acknowledgements

Throughout this doctorate, I have had the chance to meet many inspiring and supportive people, who showed interest in my work and helped me to accomplish it, one way or another. It would not be possible to mention all of them here, and the list that follows is therefore necessarily incomplete.

First of all, I would like to thank my supervisors. Jianfeng Feng supported me during the whole PhD, Hilary Buxton during the first years, and Si Wu this last year. Jianfeng Feng was always available and eager to help me and pushed me to a deeper understanding of the subject. His comments were illuminating and he has always been very good at directing me away from dead ends. Hilary Buxton was very supportive and did not count the time she spent helping me structuring my work and putting it in the perspective of the global approach of cognitive vision. I am also endebted to my examinors, David Young and Roman Borisyuk, who suggested invaluable improvements that enabled me to greatly enhance the quality of this thesis.

Then, I would like to express my gratitude to my fellow students from the bioinformatics groups of the Universities of Sussex and Warwick, and to the students from the Centre for Computational Neuroscience and Robotics of University of Sussex. They were always ready to discuss issues and provide support, whether psychological, scientific or technical. They kept giving me invaluable comments on the various unrefined ideas and manuscripts that one produces during a PhD. Enrico Rossoni, in particular has provided great and continuous help.

Writing a PhD is challenging work. It affects one's personality in many ways. I am therefore deeply indebted to my family, friends and house mates not only for their unfailing support, especially in times of inevitable crisis, but also for tolerating my occasionally 'challenging' behaviour. It has been hard work for them too. Tadzio Mueller, friend, house mate and fellow PhD student in International Relations was very inspirational at all stages, although our research areas seem unrelated.

Finally, for the longest and most unconditional support, I would like to thank my parents, Odile and Olivier Gaillard. They have always been there for me and have never ceased to encourage me as strongly as they could in all my studies and endeavours. Without them I would never have gone so far.

This PhD was partially funded by the EPSRC.

A biologically inspired model of decision making controlled by random neural activity

Benoit Gaillard

Abstract

This work describes research in detailed neural modelling of decision making. Although the study of formal models and of the psychology of decision making has a long history, only relatively recently has it been possible to explore its neural basis. Our models and simulations are based on a simple benchmark visual decision task. This is a typical task that has been used for decades in neurological studies in which a primate has to express with an eye movement a decision based on simple visual stimuli called kinematograms, whilst its neural activity is recorded. This experimental paradigm has enabled the discovery of neural pathways and patterns of activity involved in decision making. However, the neural mechanisms generating these features are still a subject of intensive investigation.

In this thesis, after reviewing the experimental background, we give an overview of current neural modelling of decision making. We then propose a more detailed and biologically inspired neural network model that generates such behaviour, from stimulus perception to decision making and the eye movement that expresses it.

First, we model early visual processing and discrimination of the general direction of movement in the stimulus. This model explains how correlation and inhibitory inputs account for the input variance to Integrate and Fire neurons and shows that increasing this variance is in fact beneficial to the discrimination accuracy of the neuron. Then we present a model for decision making based on accumulation of evidence and neural competition generated by recurrent excitation and inhibition. Our simulations of this model show that the low level stochastic background activity of the brain can control the speed and accuracy of decision making. The underlying conviction of this work is that random neural activity is fundamental to understanding the dynamics of most stages of any mental activity.

Supervisors:

Prof. Jianfeng Feng, Prof. Hilary Buxton, Dr. Si Wu

Examinors:

Prof. Roman Borisyuk, Dr. David Young

Submitted for the degree of Doctor of Philosophy

University of Sussex

April, 2007

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Decision making: re-embodying abstraction

"Cogito, Ergo Sum". I think, therefore I am. This statement, made by one that many consider the father of occidental philosophy and science, symbolises an approach to the world that informs our lives, and especially science. Initiating the enlightenment period, in the seventeenth century, the philosopher René Descartes explicitly proposed a theory of human intelligence and its relation to the body, in Descartes (1664). According to him, two categories of processes constitute human behaviour. The first are deterministic behaviours such as reflexes or mechanistic links between a sensory and a motor system, whilst the second are mediated by an immaterial soul. This view on human behaviour and thought shaped subsequent neuroscientific research for a long time, because it encouraged scientists to experimentally study the sensory apparatus and motor control system, hoping to elucidate the neural pathways involved in reflexes. For example, Sherringdon (1906) spoke about a direct "arc" between sensory and motor functions. This paradigm has been the basis of scientific research until today and has yielded significant progress in physics, technology, biology and even in neuroscience.

On the other hand, Descartes himself started a long tradition of philosophy and science that considered mind and higher thought processes to be detached from the body. This paradigm, whereby a rational and immaterial mind controls a body-machine is an essential foundation of modernity, this world-view that has been generally accepted in the Occident over the past few centuries. In this modern cosmology, just as the thinking mind controls the body-machine, humans (and even in certain historical understandings, 'certain/some' humans), have the gift of rationality which entitles them to use the rest of the world as an

object submitted to their will. In Descartes' own words; "I perceived it to be possible to arrive at knowledge highly useful in life... and thus render ourselves the lords and possessors of nature". More recently, some (Jensen (2000)) have started to realise that this view of the human as a rational controller of the world misses essential dimensions of life and leads to catastrophic outcomes. This refutation of Cartesian dualism is rooted in the critique of the philosophy of the subject that we can find in Foucault (1970) and that leads to a radically different view on society as in for example Deleuze and Guattari (1980).

Similarly, in the study of intelligence and the brain, this dualist approach has been criticized, especially during the course of the 20th century, in light of the study of perception. Two antithetical views of perception developed. The intellectualists claimed that perception was separated from the sensory motor arc and generated by higher level thought processes of the mind, whilst the materialists would see perceptions as direct consequences of physical stimuli on the sensory system. As we see in the Phenomenology of Perception (Merleau-Ponty (1945)) these opposite views carry parallel misconceptions that prevent any research that follows their paradigms from reaching a satisfying understanding of perception. According to Merleau-Ponty, on the one hand, we have to re-embody the soul, bearing in mind that thoughts and decisions are never abstracted from bodily functions, sensory inputs and motor control. On the other hand, there is no such thing as a pure sensation, because our bodily interactions with the world are mediated by our active and motivated explorations of sensory stimuli. We actually choose to perceive this or that stimulus. Without purposeful exploration of the sensory field sensations quickly fade, and only from this exploratory perception can we again choose what course of action to take. The choices, are, obviously, dependent on our motives, our context and our history, which impregnate our mental life. One can see, here, the importance of choice or decision making at the very heart of our presence to the world. Thus, in order to re-unite cognitive processes and the body, abstractions and stimuli, psychology and biology, it seems appropriate to elucidate the interplay between stimuli, decision making and the rest of our mental life.

This thesis focuses on simple decision tasks based on visual stimuli, because the visual system is the most developed and well known in primates. In this work, we have implemented a complete model of decision making, from the discrimination between sensory inputs to the commitment to a decision and the performance of the corresponding action. The general context, previously referred to as 'mental life' is simplistically modelled as low level random background neural activity. By comparing various versions of this global model, we have been trying to evaluate the role of this 'noise' in the dynamics of decision making. In order to do so, we have chosen to focus on the example of a benchmark visually based decision task, because there is established experimental knowledge of the structures involved in this task. So, the next sections will outline the neural basis of the brain, the location of the neural systems involved in this visual decision making, and the classic experimental set up used for finding and studying them.

1.2 The locus of visual decision making

1.2.1 Neural architecture of the brain

Since the work of Ramón y Cajal (1909), we know that the central nervous system is constituted of small units called neurons, that communicate via electrical impulses called action potentials or spikes. The central nervous system is divided in grey matter and white matter. The grey matter is where the cell bodies of neurons are gathered, whereas the white matter is where the wires through which they communicate are gathered, excluding cell bodies. The grey matter comprises the cerebral cortex, the cerebellar cortex and some nuclei such as the brainstem nuclei and the nuclei of the thalamus. The cerebral cortex will be our main focus, because this is where most of the information exchanges underlying our thoughts occur. Neurons in the cortex come in various shapes and with various characteristics, but some features can be stressed that describe the majority of them. There are 20 billion neurons in a human brain, and they communicate through synapses that connect each neuron to many others. The number of synapses is even more impressive: hundreds of trillions. This means that their density is very high: in 1 cubic millimetre of human cortex, there are one hundred thousand neurons and a billion synapses. Each neuron receives input from a few thousand to tens of thousands of synapses. These figures show us an important feature of neural computation: only in very exceptional cases will the output of one neuron depend on one or a few other neurons' signals. It depends on thousands of other neurons, which make the contribution of each irrelevant with regards to the global input.

Many cortical areas have been named and categorised, such as, for example the hippocampus, the thalamus, the substancia nigra, the basal ganglia, and two hemispheres. This division helps linking certain activities to certain parts of the brain. As stated in Churchland and Sejnowski (1992) the motor control system is situated in the frontal cortex, whilst the sensory cortex is situated at the back. Descriptions of the neural system are often structured in a hierarchy of neural groups, successive layers being ranked according to how far they are from the stimulus, in terms of number of synaptic connections. However, it has been known for a few decades that this description is far from capturing the stunning variety of neural pathways. The more we study a neural area or a task in detail, the more we realise that other areas are involved, or that the area under study performs other tasks. Moreover this hierarchical structure should not make us believe in a one-way system where the information would travel from sensory areas to motor control. In fact, as stated in Churchland and Sejnowski (1992), there are as many if not more feedback connections (from higher to lower areas) as there are feedforward connections. For example, the recurrent projections from the visual cortical area V1 back to the Lateral Geniculate Nucleus (LGN) are about ten times as numerous as those from the LGN to V1. We cannot give here a very detailed neuroscientific description of all brain areas. However, since we want to study visually based decision making, we will present in the next section a basic introduction to the neurology of the visual system.

1.2.2 The visual system

In “An introduction to the visual system” of Tovée (1996), there is a clear basic description of the cortical neurons involved in the visual system.

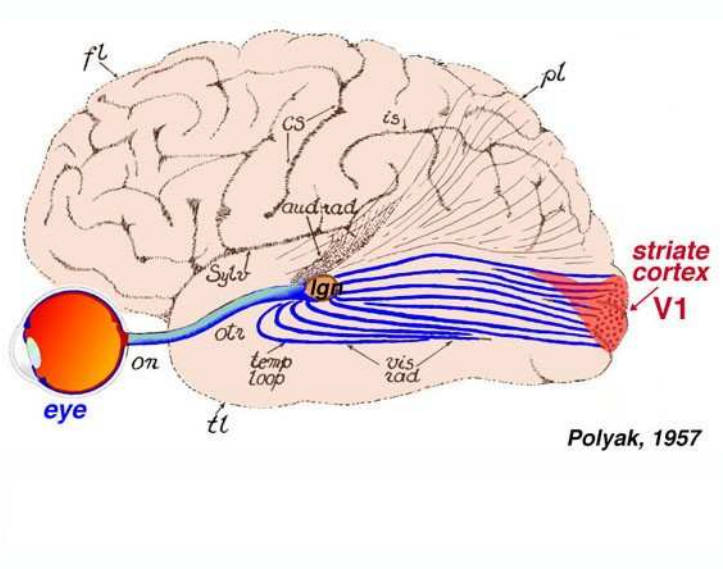


Figure 1.1: The visual stimuli go from the retina to the area V1 of the cortex, also called primary visual cortex, via the Lateral Geniculate Nucleus. The visual cortex is situated at the back of the head. Adapted from Polyak (1957)

The visual cortex is divided into five successive areas, V1, V2, V3 V4 and V5, organised in a hierarchical manner. We can see these areas in Fig. 1.2, adapted from Van Essen et al. (1992). The neurons in these areas respond to more and more complex stimuli, neurons from V1 (primary visual cortex), being sensitive to, for example, straight lines, while neurons of higher level areas can be sensitive to faces. That is why the connections between the areas are called feedforward and feedback connections.

As we can see from Fig. 1.1 adapted from Polyak (1957), stimuli from the retina reach the visual cortex via the optic nerves, through the lateral geniculate nucleus, being crossed over in the optic chiasm. First, the stimuli reach the primary visual cortex, or area V1. This means that, even if each hemisphere of the brain receives information about the opposite eye, the visual field of the retina is mapped onto the Primary Visual Cortex, thus the topography of the visual scene is preserved in V1. Furthermore V1 is organised in cortical columns. Cortical columns are specific groups of neurons that extend from the peripheral cortex to

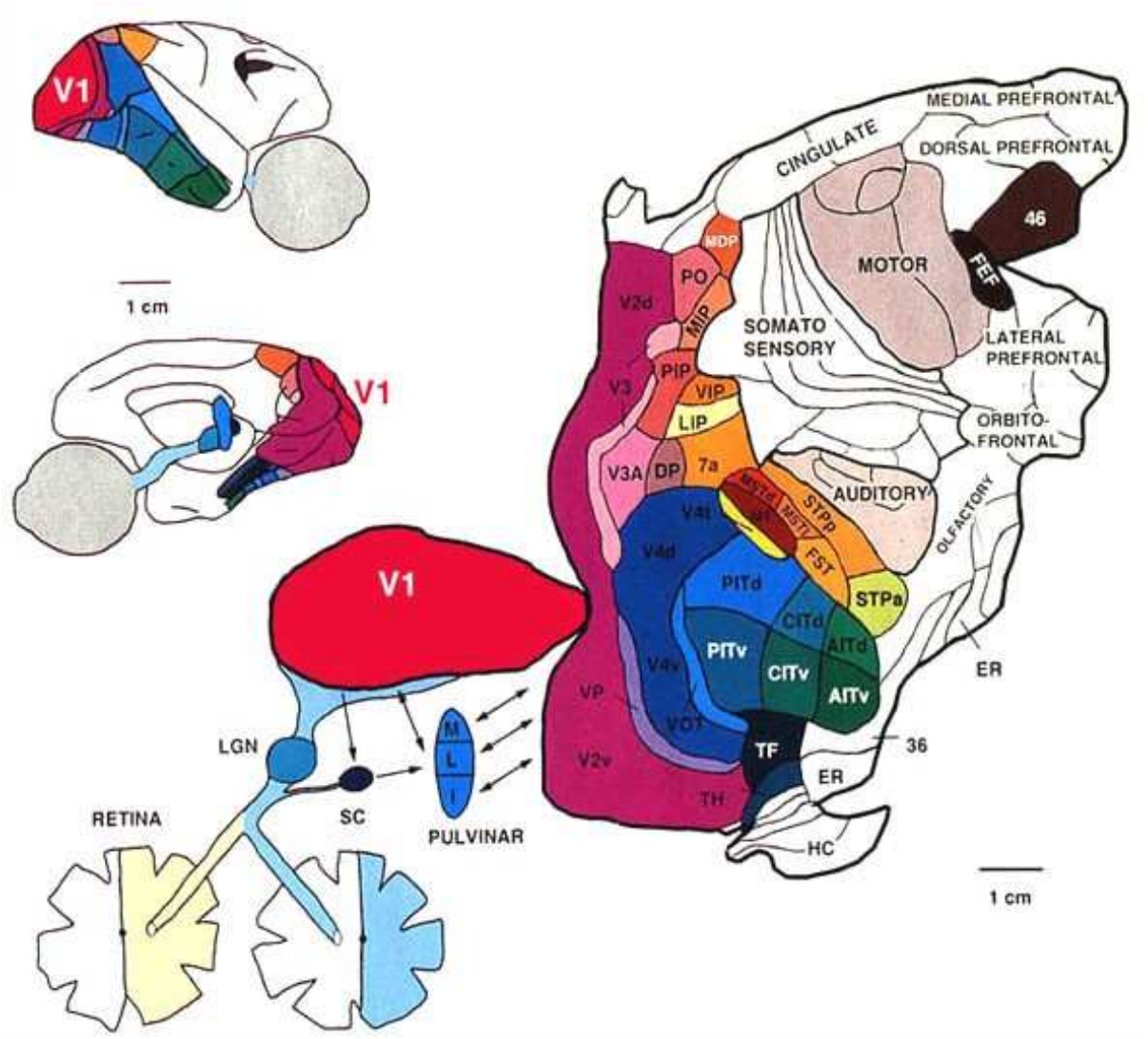


Figure 1.2: The visual cortex. The figures show areas V1 to V5 or MT, area LIP and even motor control areas. Adapted from Van Essen et al. (1992)

the white matter and are devoted to one type of stimuli.

From V1, the topography is not preserved any more, for more complex operations are processed. V2 and V4 seem to classify patterns of the visual scene, like colours, orientation or disparity, in order to prepare the representation of complex objects. V3 and V5 deal more with orientation and motion.

The primary visual cortex receives input from LGN and sends feedforward connections to the rest of the visual cortex, such as areas V2, V3, V5 or Middle temporal (MT), the adjacent Medial Superior Temporal (MST) area and even the Frontal Eye Field (FEF), as seen for example in Lund et al. (1975); Maunsell and Van Essen (1983); Ungerleider and Desimone (1986b,a); Fitzpatrick et al. (1994). It has also been shown that it receives many connections from such areas as V2, V3, V4, V5 or MT, MST, FEF, Lateral IntraParietal (LIP) area and inferotemporal cortex, as seen for example in Ungerleider and Desimone (1986b,a); Barone et al. (2000); Rockland et al. (1994); Suzuki et al. (2000).

Within this structure of the visual cortex, information is often believed to be processed in two main pathways : one, the ventral pathway, is concerned with the identification of an object, while the other, the dorsal pathway, is concerned with the positions and movements of this object. Those two pathways, though well differentiated, are neuronally linked to each other at various stages of the processing.

It has been proved, in fundamental work on the organization of the brain, that in certain parts of the brain individual neurons are highly sensitive to specific sensory inputs. These neurons are organized in groups that form topographic maps of the stimuli. For example, in the visual system, we find columns of neurons that detect specific directions of movement as seen in Mountcastle (1957) and Hubel and Wiesel (1962). In area MT, neurons detecting the same direction will be grouped together, as shown by Liu and Newsome (2003). Models have been proposed for this organization by Heeger et al. (1995) and by Simoncelli and Heeger (1998).

1.2.3 Higher order areas and motor control

Beyond the higher areas of the visual cortex, such as MT and MST, we find the Lateral IntraParietal area (LIP). It receives projections from the visual cortex. It has been shown that the activity of area LIP depends on the characteristics of the visual stimulus, and that on the other hand LIP activity is involved in saccadic eye movement preparation. In fact, a significant number of feedforward connections from area LIP project to the Superior Colliculus (SC), which is a spatially organised pre-oculomotor structure of the midbrain as

seen in Horwitz et al. (2004). Wurtz et al. (2001) show that there is a progression of neural code, from more stimulus dependent in the visual cortex to less stimulus dependent in the Superior Colliculus, via areas such as LIP and SEF (Supplementary Eye Field). This is also explained by Carpenter (1988); Carpenter and Williams (1995). More precisely, it has been shown by for example Wurtz and Goldberg (1989) that the SC encodes the metrics (amplitude and direction) of the Saccadic Eye Movements (SEM) in a topographic manner. These instructions are then projected to the three oculomotor nuclei situated in the brainstem. These three nuclei are the motor neurons that control the eye muscles, as shown by Robinson (1964).

1.3 A benchmark decision task based on visual stimuli

Extensive experimental studies have been carried out on visual decision processes of the macaque monkey, in particular by Newsome and colleagues (Britten et al. (1992); Shadlen and Newsome (1996, 2001); Zohary et al. (1994)). In this set of experiments, monkeys have to watch random dots stimuli called kinematograms and then make a deliberative eye movement that indicates a decision regarding the stimuli.

1.3.1 Kinematograms

The kinematograms consist of a succession of images of dots on a screen. The precise design of the kinematograms is described in Britten et al. (1992). Between two images, a percentage p of the dots is chosen to change position as a function of their previous position, giving the impression that they move coherently in one direction. The remaining dots change position randomly between two images. This design gives the experimenter a control over the clarity of the motion perception: the larger p is, the more obvious will the motion be. If $p = 0$, all dots randomly change position between two images. If $p = 100\%$, all dots move in one direction. So, as we can see in Fig. 1.3 and Fig. 1.4, adapted from Britten et al. (1992) and Britten et al. (1993) respectively, by simply varying the coherence, one can vary the difficulty of the task of discriminating the general direction of the kinematogram. We will not describe all the details of the kinematograms here, but one should notice that the $p\%$ dots that change place according to a coherent direction are chosen randomly between two images. That makes it impossible to evaluate the general direction by following the direction of one dot. One can see such kinematograms, with various values of coherence p on *http://monkeybiz.stanford.edu/research.html*.

1.3.2 Experimental set up

After several months of training, rhesus macaque monkeys are made to watch the kinematograms and to discriminate their general direction in exchange for rewards such as fruit juice or water. During this task neural activities of carefully chosen neurons in the monkeys' visual cortex are recorded at the same time as psychophysical performance is measured. Various experiments can be run within this paradigm. One can

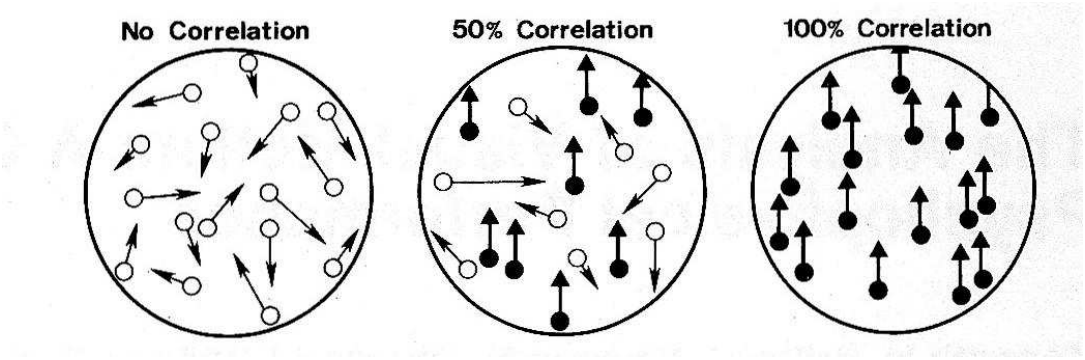


Figure 1.3: Illustration of kinematograms, adapted from Britten et al. (1992).

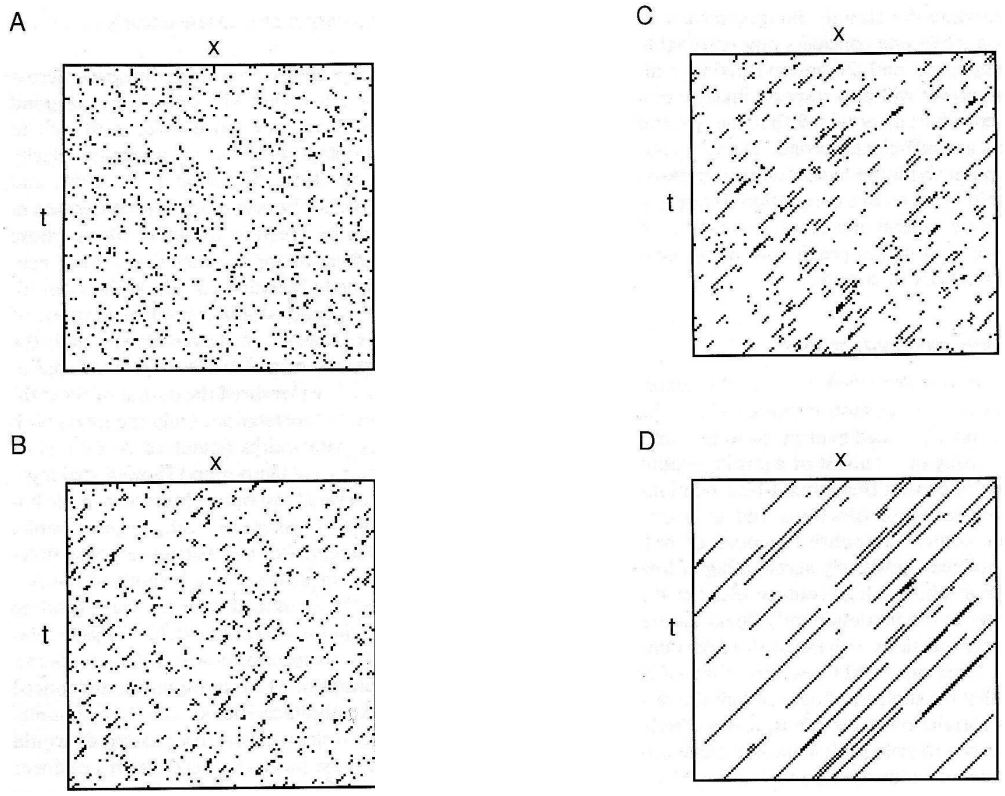


Figure 1.4: Illustration of kinematograms, adapted from Britten et al. (1993). In abscissa, the positions of the dots, in ordinate, the time of the snapshot. In that case, only the movements of dots along the x axis are represented. Panel A: coherence $= 0\%$*;Panel B: coherence* $= 30\%$*; Panel C: coherence* $= 60\%$*; Panel D: coherence* $= 95\%$*.*

either display the kinematogram as long as necessary until the monkey takes a decision, or only flash the kinematograms for a chosen period of time and ask the monkey to express its decision via a Saccadic Eye Movement only after a chosen delay. In the first case, it is possible to evaluate the Reaction Time (RT) of the animal, and in all cases one can measure their Error Rate (ER). This experimental set up is illustrated in Fig. 1.5, published by Shadlen and Gold (2004).

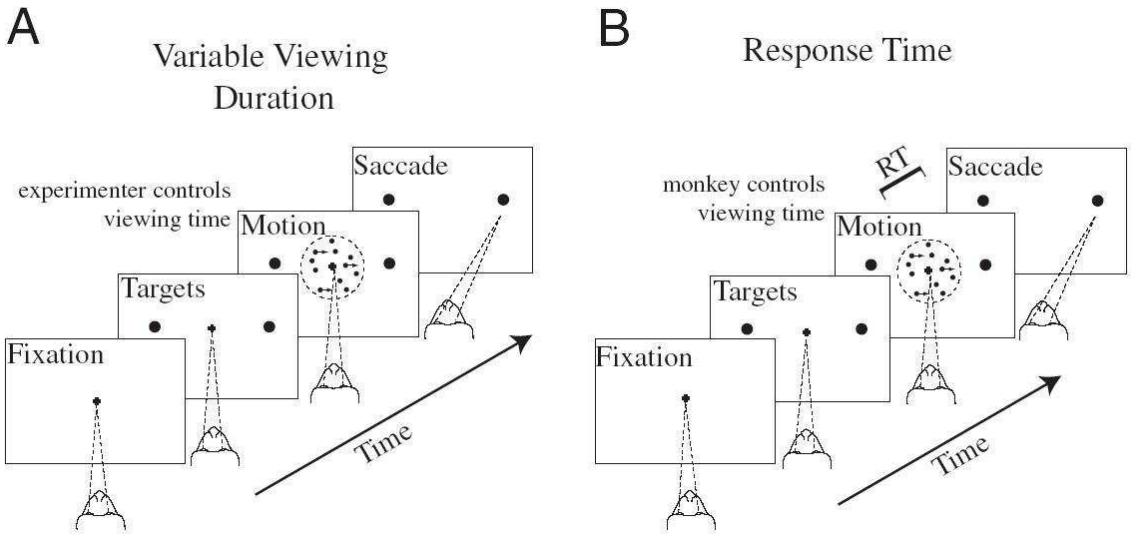


Figure 1.5: Illustration of two experimental set ups for the visual discrimination task involving monkeys, adapted from Shadlen and Gold (2004).

1.3.3 The results: from perception to decision

Discrimination area MT (V5)

In an early series of experiments, reported for example in Britten et al. (1992) and Zohary et al. (1994), neurons were recorded from one area of the cortex whose activity corresponded to the monkey's behaviour. First, it was proved, in Newsome et al. (1989), that "psychophysical judgements could be based on a relatively small number of neurons". More specifically, area MT was proved to be directly involved in the perception of motion: "These experiments revealed that MT neurons carry directional signals of sufficient precision to account for the psychophysical sensitivity of behaving monkeys (Britten et al. (1992)), and that

electrical micro stimulation of MT can influence choices in the discrimination task in a directionally specific manner (Salzman et al. (1992))... Taken together, these findings strongly support the idea that directional signals in the motion pathway, and specifically in area MT, contribute directly to the perception of motion", quoted from Britten et al. (1993).

In Britten et al. (1993), the behaviour of MT neurons as a function of the coherence of the kinematogram has been precisely described. They "found that many MT cells have an approximately linear dependence of mean Firing Rate on stimulus correlation". These linear responses, as opposed to non-linear responses to contrast observed in the case of grating displays led to the conclusion that the MT neurons are also submitted to inhibitory input coming from the stimuli. An intrinsic variability in neural responses of MT was also pointed out, and it was showed that it is not due to the variability of the stimuli. Furthermore, it is possible to adapt the kinematogram to the specific fields of neurons under study: as shown in Liu and Newsome (2003), neurons are "clustered according to preferred speed".

In summary, the recorded activity of carefully chosen MT neuron can serve as evidence for the person who runs the experiment to discriminate the direction of the kinematogram. The performance of such a set up would be comparable to the performance of the behaving monkey itself. What remain unexplained are the source and the role of the variability in neural responses in MT. In Chapter 3, we present a model of MT neurons that proposes a possible source of intrinsic variability, and shows its beneficial role.

Decision area LIP

Knowing where and how the perceived general direction of a kinematogram is encoded is not enough to explain how decisions based on this evidence arise, because this encoding is an instantaneous evaluation that varies with time. Which neurons are influenced by MT activity? Is the activity of these neurons related to the actual decision of the monkey? In order to accurately pinpoint the neural basis of the decision, one should find neurons whose activity is correlated to the final decision of the animal, but dissociated from the stimulus perception and from the actual motor movement expressing it.

Discriminating decision-related activity from stimulus-related activity has proved easy, using recordings during which the animal performs an erroneous decision: the kinematogram moves downwards, whereas the monkey performs an eye saccade that expresses his belief in an upwards movement, as explained in Shadlen and Newsome (2001). This is how we see that MT activity is not correlated to the building up of a decision. Shadlen and Newsome (2001) also could dissociate decision making from motor planning by introducing a delay period between motion viewing and the required eye saccade, or by measuring the certainty of the decision as a function of the strength of the stimulus, because this certainty is supposed not to be expressed in the execution of the saccade.

One area emerges as the most central for kinematogram related decision making: the Lateral IntraParietal area, or area LIP. This hypothesis is supported by Glimcher (2003b), Mazurek et al. (2003) and Huk and Shadlen (2005). As stated in a recent review by Sugrue et al. (2005) on decision making: "These results have prompted the proposal that LIP, in cooperation with other areas, implements the decision transformation in this task, converting a sensory representation of visual motion into a decision variable that is capable of guiding behaviour." Results supporting this hypothesis are published in for example Roitman and Shadlen (2002) and Shadlen and Newsome (2001). In Fig. 1.6, adapted from Shadlen and Newsome (2001), we can see the build up of measured LIP activity during decision formation, the storage of this decision during the delay period and the burst of activity during the saccade execution, followed by a reset to baseline activity.

In a recent paper, Huk and Shadlen (2005) analyse in more details the formation of the decision in area LIP. They show that the neural recordings fit well the model of integration of evidence over time. This integration mechanism is an idea that came to the field of decision making from psychology, with models proposed by for example Ratcliff (1978),whose model is a diffusion model that implies temporal integration. So, according to these results, neurons in area LIP integrate sensory evidence about the kinematogram from area MT. When their activity reaches a threshold, the decision is taken and the monkey performs the saccadic eye movement.

In brief, we have seen that we can describe with satisfying approximation this visual discrimination task as follows: area MT compiles sensory stimuli into a global perception of motion, and LIP integrates sensory evidence from MT into a decision that is forwarded to motor control areas such as the Superior Colliculus (SC) in the prefrontal cortex. Another important feature is that the decision can be stored for a while in the LIP area before the monkey acts upon it. What remains to be done is to formulate a hypothesis on how this integration of evidence is performed, how this higher LIP neural activity is maintained and re-initialised, as we see in neural measurements such as illustrated in Fig. 1.6. Furthermore, we will see below that aspects of this decision making chain remain unexplained.

1.3.4 Stimulus independent control over decision making

Distinguishing the formation of an abstract decision from the preparation of the motor command is not as obvious as it seems. In recent research, Horwitz et al. (2004) have used another experimental set up, the "Loose Stimulus Response Association" (LSRA) task, to distinguish decision related neural activity from motor control. In this set up, the monkeys have to keep their decision in mind for some time before the target for saccadic eye movement is displayed. In short, their reasoning is that, since the monkeys can not predict which movement they will have to do, the activity from which the experimenter can guess

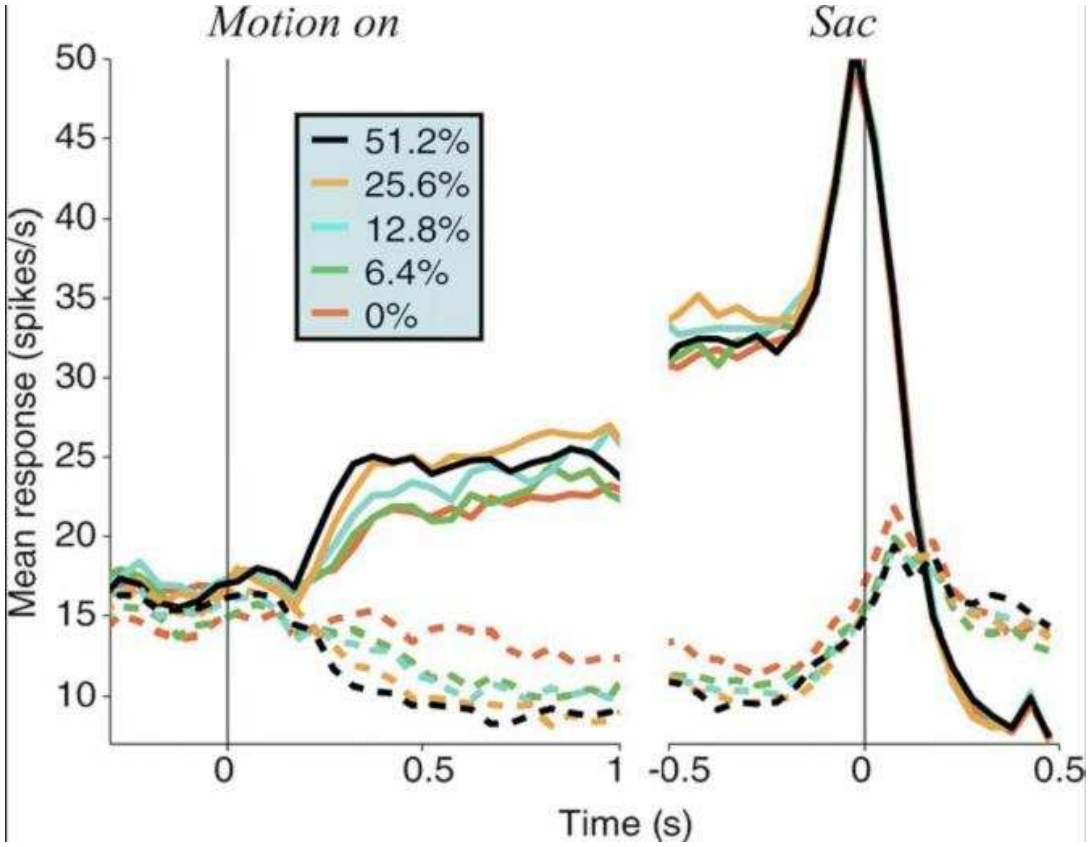


Figure 1.6: Reproduced from Shadlen and Newsome (2001). Population response from 104 LIP neurons during the direction discrimination task. The average firing rate is plotted as a function of time during the motion viewing and delay periods. Solid and dashed curves are from trials in which the monkey judged direction toward and away from the Receptive Field, respectively. The time origins of the graphs are set on motion onset and on saccade onset. We can see that before the saccade, during the delay period, the LIP activity is maintained.

the monkey's decision independently of the actual stimulus must be decision-related activity. To their surprise, they found that areas like the Superior Colliculus (SC), previously believed to be only involved in motor control, were also showing decision-specific activity: "That decision-related signals are present in a spatially organised structure like the SC raises the alternative possibility that decisions in the LSRA task are in fact coded with respect to space" in Horwitz et al. (2004). Decision-related neural activity has also been detected in the dorsolateral prefrontal cortex by Kim and Shadlen (1999) and Rorie and Newsome (2005).

On the other hand, in Shadlen and Newsome (2001), we see that the decision-related activity in LIP bears traces of the particulars of the stimulus. The less easy the discrimination task is, the less sure the monkey is of the decision, and this translates into LIP activity: the difference between the activity of the neurons whose receptive field projects to one saccade and those opposite is reduced with the uncertainty of the stimulus. This suggests that a network of cells belonging to various areas from low-level SC to higher level LIP cooperate to integrate all levels of the signal.

Moreover, as stated in Shadlen and Newsome (2001), area LIP receives other information besides sensory related input. This other information can be reward expectation or prior probabilities as seen in Platt and Glimcher (1999) or urgency as studied in Reddi and Carpenter (2000), where the authors explain that we can decide to take decisions with various speeds and accuracies, according to externally applied factors such as contextual urgency. In this last paper it is pointed out that the time between presentation of a stimulus and eye movement is much longer than a direct transmission of stimuli would suggest. This delay is explained by a process of deliberation happening in higher areas of the cortex, where the monkeys would evaluate what move they prefer to do, depending on the global context. In fact, as some authors point out (Salinas (2003)), even in the presence of specific stimuli, we do not systematically act upon them. It is not because we see a red light that we systematically execute the leg movements that correspond to pressing the brake pedal of a car. In that article, the author proposes background activity as a switch between different network states. The specific neural activity that corresponds to acting upon a stimulus is triggered by this background activity as much as by the presence of a stimulus.

1.3.5 Phenomenology of perception

We have seen that, up to current knowledge about this particular visual discrimination task, the chain from perception to action via decision making seems to be spread across various neural areas and various levels of connectivity, all contributing to making a decision in a manner that is very dependent on the contingencies of both the stimuli and the demanded motor response. On one hand, it is clear that this chain from perception to action is not a purely physical or deterministic reflex, as suggested by Descartes or Sherrington. On the other hand, it has been consistently proved that neural areas such as LIP, that account

for the higher cognitive processes involved in this decision making are never abstracted from either the sensory stimuli or the motor control.

This would be paradoxical in a traditional Cartesian vision of the human body and soul. The only way to explain these results is to use a philosophical paradigm that links the body stimuli and motor abilities to the higher cognitive processes. In *Phénoménologie de la Perception* Merleau-Ponty (1945, 1968) gives us this exact paradigm, drawing on the philosophical tradition known as Phenomenology. First, he shows that there exist no pure sensations, and that our perceptions are informed by our intentions. In other words, our seeing of a curve is not our sensory apparatus encoding the position of the infinite number of points that constitute it, it is rather giving these stimuli the meaning of curve. He then also refutes the intellectualist point of view in which, in short, we only perceive what our rationality tells us about the world. He goes beyond these two views and their refutations by showing that they eventually come back to the same misconception: *"En réalité, l'image d'un monde constitué où je ne serais avec mon corps qu'un objet parmi d'autres et l'idée d'une conscience constituante absolue ne forment antithèse qu'en apparence: elles expriment deux fois le préjugé d'un univers en soi parfaitement explicite"*[1] p.51, Merleau-Ponty (1945). Basically, he shows us that perception can neither be reduced to a sum of constant sensations or qualias as the empiricists do, nor can it be explained by pure reasoning. He explains perception through the motivation, where stimuli are perceived because they make sense in relation to the motivation of the subject. *"Il n'y a pas de raison, mais il y a un motif"* [2], p. 48.

This vision of perception sheds an interesting light on the questions we raised above: on the one hand, the decision-related neural activity of the monkey cannot be abstracted from stimulus perception nor from action planning and on the other hand it is influenced by its motivation or intentionality. This is exactly the point made by Merleau-Ponty in his *Phénoménologie de la Perception*: perception arises only when one makes sense of sensory stimuli according to one's intentionality. The decision-related activity of LIP neurons in the Rhesus Monkey can be seen as this process of making sense of stimuli in light of our motivations. The actual intentions of the monkeys will change the characteristics of their decision making: will the monkey make a decision about the stimuli, how fast, with what accuracy?

This phenomenology of perception was a precursor to the more recent field of enactive perception. In the fields of robotics and artificial intelligence, enactive perception is an attempt to study perception as an active behaviour of the subject instead of considering the perceiver as just a passive receiver of information, as

[1] In fact, the image of a constituted world where, with my body, I should be only an object among others and the idea of a absolute constituting consciousness are only apparently antithetical: they are a dual expression of a universe perfectly explicit in itself (translation by Colin Smith)

[2] There is no reason, but there is a motive

shown in for example Brooks (1991), or Thomas (1999). Drawing on these lines, O'Regan and Noe (2001), have attempted to give us a "Sensorimotor account of vision and visual consciousness", in which they apply to vision and visual awareness the phenomenological view of perception. This paper explains that "Vision is a mode of exploration of the world that is mediated by the relevant sensori-motor contingencies". Sensori-motor contingencies are the rules that enable us to predict the changes that will happen in our perception if we perform this or that movement. Mastering these sensori-motor contingencies requires training, after which we can sustain the illusion of a stable perception, and of representation of the external world inside the brain. This theory is supported by experimental evidence of sensory substitution and change blindness. After training, it has been proved that one can, for example, "see with one's ear". It is also well known that when somebody wears inverting glasses as reported in Stratton (1897), or Harris (1965), following a few hours of being incapacitated, one adapts in a few days and the world seems to turn back up after a couple of weeks of wearing the glasses. That supports the idea that our illusion of a stable world comes from our trained mastery of rules of how stimuli change according to which movements. More precisely, in order to study visual consciousness, they state that "To see is to explore one's environment in a way that is mediated by one's mastery of sensori-motor contingencies *and* to be making use of this mastery in one's planning ...." O'Regan and Noe (2001). In other words, it is making sense, in the light of one's motives, of actively collected stimuli. In the case that we study here, this happens when the monkey pays attention to the general direction of the kinematogram with the intention of performing the appropriate eye saccade. Our hypothesis is that, during this activity, the monkey makes the decision, stores it and plans a saccade in areas LIP or SC, whilst this whole process is controlled by his/her intentionality that translates in the global activity of the brain.

So, by rejecting the dualist view of Descartes on the body and soul, we have gained a conceptual framework that enables us to explain the fact that decision-related activity depends on stimulus characteristics, on higher level cognitive processes and is also found in motor control areas. The task remains to find a neural basis for the motives and intention of the monkey. In this work, we are trying to show that these motives can be mediated neuronally by various aspect of what is traditionally called neural noise.

1.4 Thesis Organisation

In order to motivate the work presented in this dissertation the introduction gives an overview of the brain anatomy and more specifically of the sensori-motor system of primate vision, from the inputs to the retina to the neural control of the eye muscles. In order to understand and discuss models of decision making, we need a more extensive knowledge of the biological processes involved, as well as a description of the modelling tools that have been developed.

1.4.1 Neurology and neuro-modelling of decision making

The first chapter will describe the knowledge in neural modelling and then focus on the neurophysiology of decision making. We will also describe the modelling approaches to specific decision making structures of the primate brain, following the pathway that goes from the visual stimuli to the Saccadic Eye Movement (SEM) generation. The last part of the chapter will describe some mathematical tools used to model dynamical systems perturbed by noise.

1.4.2 Input noise improves discrimination accuracy

In order to study decision making structures, the first step is to create a satisfying model of their input. How is the visual stimulus encoded in the visual cortex in order to provide suitable evidence on which to base sensory-related planning? According to the literature, some neurons in the extrastriate cortex (MT, MST) encode stimulus information into neural evidence for later decision making. We produced a model for movement encoding in the case of kinematograms. Despite its inherent flaws, this model sheds light on issues concerning the early stages of vision.

Then we assumed a simple neural encoding of the kinematogram directional information and modelled the translation of this inputs into decision-related evidence. This model mimics neurons of area MT. We call this part the discrimination stage of the pathway, because the output of the model can be used to discriminate the stimuli. It already shows the importance of noise in neural coding at this relatively low level of cognition. It shows how inhibitory and correlated inputs increase input noise to Integrate and Fire neuron models at the same time as they improve their discrimination accuracy, and it also stresses the importance of considering populations of neurons as opposed to single neurons.

1.4.3 Decision making induced by perturbing a dynamical system

As seen in section 1.3.3, the evidence produced by MT neurons is then integrated by LIP neurons until the primate reaches a decision. The second part of our research describes a mechanism that simulates this integration, with a model that uses recurrent inhibition between populations of Integrate and Fire neurons to generate a neural competition. The model shows that the dynamics of this competition can be controlled by the two first statistical moments of the low-level background activity of the brain. In order to model that the decision is only fully made when expressed by a bodily movement, we had to create a crude model of the preparation, in area SC, of a Saccadic Eye Movement (SEM).

1.4.4 Discussion

Here we discuss our work. We summarize the work presented in the whole dissertation. By comparing our models and results to others, we outline our contributions to the research in our domain, and stress the limitations of our work. We show attempts to create more precise models that can overcome the limitations inherent to the simplifying assumptions that we took during the course of this work, and lastly we outline the directions for longer term research projects that this dissertation is hinting at.

1.5 Published Work

Partial results of the research presented here have been published in various papers. In this paragraph we present the list of these publications. All our publications can be downloaded from: *http://www.cogs.susx.ac.uk/users/bg22/*

- Gaillard, B. and Feng, J. (2005). Modelling a visual discrimination task. *Neurocomputing*, 65-66:203209.
- Gaillard, B., Feng, J., and Buxton, H. (2006a). Neuronal model of decision making. In Feng, J., Qian, M., and Jost, J., editors, *Networks: From Biology to Theory*. Springer Verlag. In Press.
- Gaillard, B., Feng, J., and Buxton, H. (2006b). Population approach to a neural discrimination task. *Biological Cybernetics*. 94(3):180–191.

Chapter 2

Neurology and neuro-modelling of visual decision making

We are tackling a vast field: from vision to motor control via higher level cognitive processes like decision making. Furthermore, our subject of research extends over various level of complexity, from the single cell characteristics of one neuron to philosophical concepts such as attention and motivation. A detailed analysis of this whole domain is of course beyond the scope of this work. However, the aim of this research is to investigate the interaction between psychology and the sensorimotor chain. Thus, we need an understanding, if only partial, of the flow of information between stimulus and motor control and its relation to decision making and the global context. Only then can we produce a precise and contextually relevant research on our specific area of focus: neural processes involved in simple decision making. In this chapter, we will mainly focus on the level of neurons and populations of neurons, but we will refer to the basics of higher level concepts issued from the psychological fields.

2.1 Neurons are the basic units of the brain

Since the foundational work of Santiago Ramón y Cajal, more than a century ago, we know that the nervous system is made of billions of separate nerve cells. Ramón y Cajal proved that the basic units of the nervous system are individual cellular elements, which Waldeyer had already named ”neurons” in 1891. In Ramón y Cajal (1909), he laid the foundations of neuroscience research by providing a detailed description of cell organization in the nervous system, illustrated by renowned drawings, such as in Fig. 2.1, which are still reproduced in neuroscience textbooks. In addition, he defined a basic principle of the functioning of neural connections, ‘the law of dynamic polarization’, that states that the nerve cells are polarized, receiving information on their cell bodies (soma) and dendrites and conducting it to distant locations through axons.

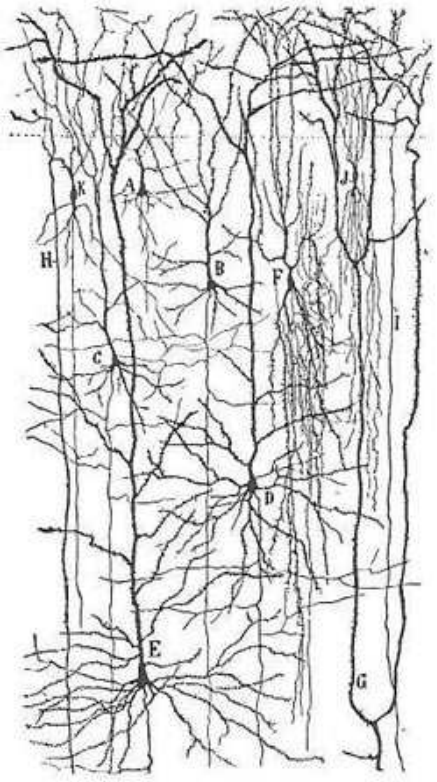
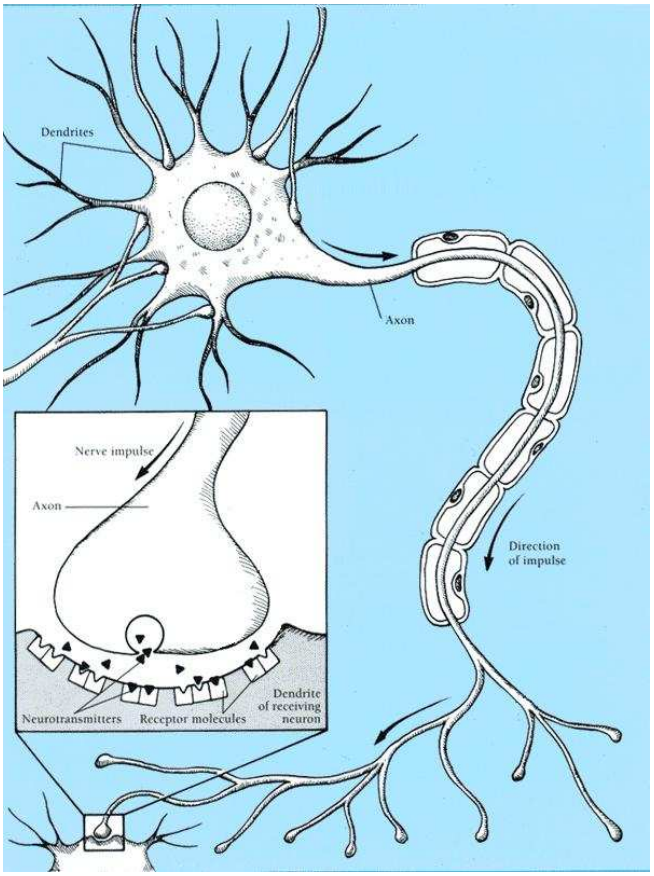


Figure 2.1: Superficial layers of the human frontal cortex drawn by Ramón y Cajal on the basis of Golgi impregnation. The main cell types of the cerebral cortex, i.e. small and large pyramidal neurons (A, B, C, D, E) and non pyramidal cells (F, K) (interneurons in the modern nomenclature) are outlined. Adapted from Ramón y Cajal (1909)

2.1.1 Biological neurons

Seeing a neuron as an information processing system, we can follow classic descriptions given by Ramón y Cajal (1909) or more recently by Gerstner and Kistler (2002). According to their widely accepted interpretation, neurons take inputs from the dendrites, process the information in the soma, and then generate and send the output through the axon. We can see these three parts in details in Fig. 2.2.

In Koch (1999), detailed descriptions of the various biophysical parts of the neurons are given. The membrane separates the inside from the outside of the neuron cell. The membrane potential is the difference



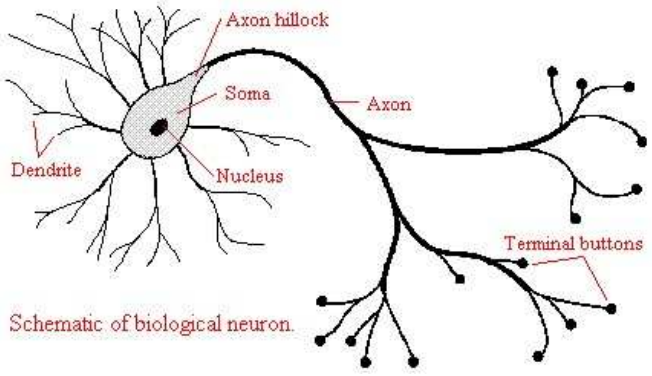


Figure 2.2: Two different schematic representation of a neuron and a synapse. Adapted from www.pfizer.com/brain/dlgame.html and from *www.lebenswissen.de*

of electrical potential between the inside of the neuron and the outside of the neuron. This potential varies because of various complex ionic currents flowing between inside and outside via ionic channels. When this membrane potential reaches a threshold, the neuron emits a spike and the membrane potential falls back to the resting potential.

The soma is the central part of the neuron. It contains the nucleus and other elements needed by every cell for their maintenance. The soma is separated from the axon by the hillock. When the membrane potential reaches a threshold, an action potential is generated. As described in Rieke et al. (1997), an action potential, or spike, is a brief (in the order of one ms) and localised increase of the membrane potential, followed by a refractory period during which the membrane potential is fixed at the resting potential. As seen in Koch (1999), spikes come in various shape depending on the type of the neuron, but one neuron will always produce the same spike. This spike propagates along the axon, away from the neuron. In order to keep the spike relatively unchanged over long distances (according to Braitenberg and Schutz (1991), an axon can be as long as 10 mm.), many axons are surrounded by a myelin sheath constituted of fat-containing cells that are called Schwann cells. This sheath has two important effects: it insulates the axon from the outside of the neurons, preventing electrical dissipation, and it significantly increases the speed of transmission of the spike. As mentioned in Izhikevich et al. (2004) and as shown in experimental measurements by Waxman and Bennett (1972) and Swadlow (1994), an un-myelinated axon will still propagate spikes, but at the speed of 0.15 m/s, whereas a myelinated axon will propagate spikes at 1m/s.

The succession of these spikes constitutes the neural code. The spike in itself, having a normalised structure, does not contain information. Its timing related to the timing of other spikes carries the information. This information is transmitted to the next neuron when action potentials reach the end of the axons, at the terminal buttons. There the arrival of a spike induces the release of neurotransmitters to the next neuron via the synapse. Synapses are the loci where the terminal buttons of the axons are in contact with the dendrites of the next neuron. A synapse is illustrated in Fig. 2.2. Upon receiving a spike, ions are released that modify the membrane potential of the next neuron at the dendrite.

The dendrites are filaments that branch out of the soma in many directions. They have an irregular surface that enables them to collect the action potentials from the other neurons. For each received spike, the membrane potential slightly increases if the spike comes from an excitatory neuron, or decreases in the case of an inhibitory neuron. The dendritic tree continuously receives spikes from various locations and combines the received information both spatially and temporally. As a first approximation one could compare this combination to a summation, but more detailed biological studies described in for example Koch (1999) have shown that other non-linear operations play a fundamental role in input computation. Let

us give two examples of such non-linear computations. Some spikes come from inhibitory neurons. Many of them have a linear effect, simply opposite to the effect of excitatory spikes. However, some will produce what is called *shunting inhibition*: situated at strategic locations of the dendritic tree, the reception of one of these shunting inhibition spikes will prevent the information from a whole part of the dendritic tree from reaching the soma. Another fundamental non-linear effect is the decay, or leak. When, due to the reception of some spikes, the membrane potential is higher than the resting potential, it tends to decrease if it does not receive more spikes. This is a fundamental characteristic that enables the neuron to compute rich features of incoming spike trains, such as correlation and variability.

There are various kinds of neurons. Some salient categories are pyramidal cells, stellate cells, granule cells, and Purkinje cells. Pyramidal cells are the most prominent cells in the cerebral cortex. Their name comes from their specific shape: the apex of the pyramid points toward the cortical surface. A large apical dendrite extends further upward toward that surface, while other dendrites arise from the corners and sides of the pyramid. The axon extends down into white matter from the base of the pyramid. Most pyramidal cells project association fibres to other cortical regions and to deeper nuclei of the brain. They come in various sizes, depending on which part of the cortex they belong to. A realistic picture of a pyramidal cell is shown in Fig. 2.3.

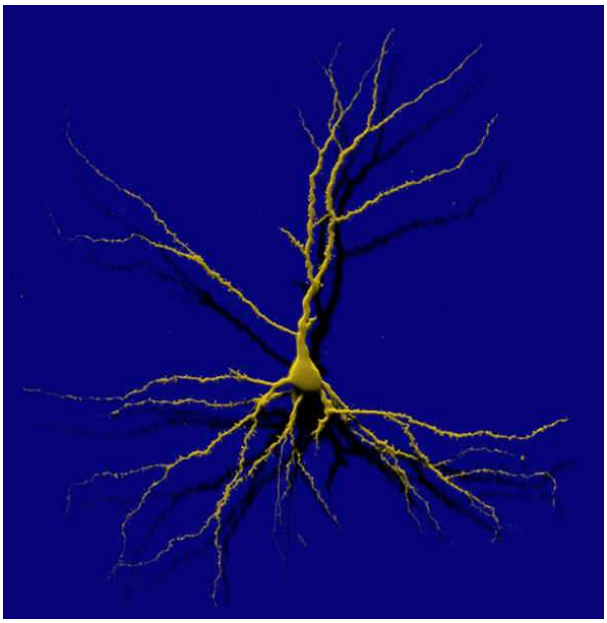


Figure 2.3: Three dimensional reconstruction of a pyramidal cell, based on laser microscopy. Adapted from $http://www.unimagdeburg.de/bio/Pictures.htm$

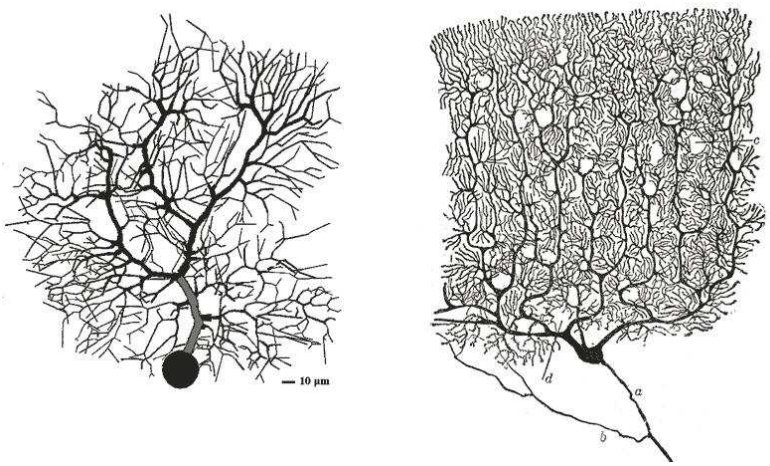


Figure 2.4: The cerebellar Purkinje cell in the guinea pig (left, adapted from De Schutter and Bower (1994)) and in the human (right, adapted from Ramón y Cajal (1909)). While the diagram on the left is a modern computer reconstruction of a fluorescent-stained P-cell and thus exhibits all details, the classical diagram on the right is a drawing based on Golgi silver stain preparation, and thus the actual Purkinje cell dendritic arbour could be even more complex than shown. Adapted from Simons and Pellionisz (2006) http://www.junkdna.com/fractogene/05_simons_pellionisz.html

Stellate cells and granule cells are neurons with short axons, and they are often also called local interneurons. Stellate cells owe their name to the star shaped distribution of their dendritic tree, whereas granule cell owe their name to the fact that they are very small and numerous. Lastly, Purkinje cells are cells from the cerebellar cortex, sending output to other regions of the brain. Thus they have long axons. Their size and beauty resulted in pictures of those cells being produced very often, such as Fig. 2.4

2.1.2 Neural models

Hodgkin-Huxley model of spike generation

Various neural models have been proposed in order to describe neural reactivity. The fundamental model of action potential generation is the Hodgkin-Huxley model (HH model), designed by Hodgkin and Huxley (1952) and based on extensive experiments on the giant axon of the squid. Their quantitative description models membrane potential dynamics, spike generation and spike propagation along the axon with an electrical circuit of four branches containing capacitances and conductances. This circuit is shown in Fig. 2.5. That their description accurately describes such a variety of observed membrane phenomena in terms of simple underlying variables is all the more impressive since the specific ionic channels underlying the membrane currents were not known at that time.

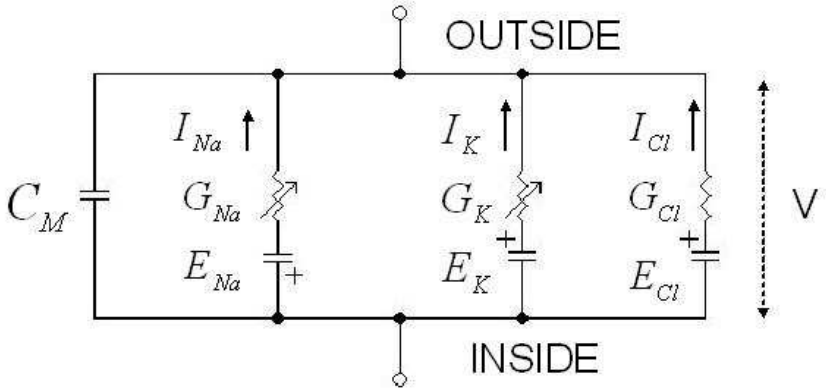


Figure 2.5: Electrical network modelling membrane potential dynamics, adapted from Hodgkin and Huxley (1952) and $http://sun.science.wayne.edu/jram/axon\_potential\_simulator.htm$. *$g_{Na}$ and g_K vary with time and with the membrane potential E_M. Originally, Hodgkin and Huxley named the illustrated chloride current, I_{Cl}, 'leakage current', I_{leak}. They had already recognized that it included a chloride component and had an equilibrium potential near that of chloride.*

The circuit presented on Fig. 2.5 is the schematic representation of the model that Hodgkin and Huxley postulated to describe the events underlying the generation of a spike. Let us describe step by step the elements of this model:

- The variation of the membrane potential variation $V(t)$ due to external current $I(t)$ is described by the following equation: $C_m \frac{dV(t)}{dt} = I(t) - I_{ionic}(t)$.
- The three main ionic currents involved are a sodium current, a potassium current and a leak current: $I_{ionic} = I_{Na} + I_K + I_{Cl}$.
- These ionic currents can be calculated via Ohm's law $I_i(t) = G_i(V(t),t) \cdot (V(t) - E_i)$, where I_i is the current of the ionic species i, E_i is the reversal potential given by Nernst's equation for the ionic species i and G_i is the conductance for the ionic species i.
- The conductances G_{Na} and G_K can be expressed as a maximum conductance multiplied by a coefficient that varies with time. Not knowing of the underlying ionic channels, Hodgkin and Huxley described these coefficients as functions of fictive gating particles that can be open or closed, depending on time and membrane potential. These variables control the dynamics of the model. In order to fit the observed kinetics of sodium and potassium currents, they described the conductances with these equations: $I_K = \overline{G_K} n^4 \cdot (V - E_K)$ and $I_{Na} = \overline{G_{Na}} m^3 h \cdot (V - E_{Na})$, with n, m and h being dimensionless measures of the probability of having respectively a potassium activation particle open, a sodium activation particle open and a sodium inactivation particle closed.
- The gating particles have first order kinetics, thus $\frac{dx}{dt} = \alpha_x(V) \cdot (1-x) - \beta_x(V) \cdot x$, x denotes the gating variable m, n or h.

Table 2.1: Kinetic parameters of gating particles in the Hodgkin and Huxley model.

x/parameter	n	m	h
$\alpha(V)$	$\frac{10-V}{100(e^{(10-V)/10}-1)}$	$\frac{25-V}{10(e^{(25-V)/10}-1)}$	$0.07e^{-V/20}$
$\beta(V)$	$0.125e^{-V/80}$	$4e^{-V/10}$	$\frac{1}{e^{(30-V)/10}+1}$

- α_x and β_x depend on V. We present the experimental evaluations made by Hodgkin and Huxley, as reported in Koch (1999), in Table 2.1.

They experimentally evaluated the other parameters that they introduced: $\overline{G_K} = 36mS/cm^2$, $E_K = -12mV$, $\overline{G_{Na}} = 120mS/cm^2$, $E_{Na} = 115mV$, $C_m = 1\mu F/cm^2$. They also measured the membrane leak $G_{Cl} = O.3mS/cm^2$ and chose its reversal potential so that when $V = 0$ the total membrane current is null: $E_{Cl} = 10.6mV$. As explained in Gerstner and Kistler (2002), the voltage values are empirical and need to be shifted by - 65 mV to be in accordance with what is accepted nowadays. It was later shown that the leak is mediated by chloride currents.

We thus reach the following equation for the complete Hodgkin and Huxley model of membrane dynamics:

$$C_m \frac{dV}{dt} = \overline{G_K} n^4 (E_K - V) + \overline{G_{Na}} m^3 h (E_{Na} - V) + G_{Cl}(E_{Cl} - V) + I \tag{2.1.1}$$

This model is still the basis of much contemporary modelling research, and has been the subject of countless pieces of literature. See, for example, Jack et al. (1975) for an extensive study of various aspects of this set of equations. These equations predict spike generation when the membrane potential reaches a threshold, spike propagation along the axon and sustained spiking activity at high frequency when the neuron is subjected to a step of current. Nowadays, we know more precisely what underlies ionic currents, and how ion channels open and close at the microscopic level. Furthermore, many other ionic currents have been found to influence the membrane potential behaviour, and can be added to improve the accuracy of the Hodgkin Huxley model. Nevertheless it is still the model on which most of spiking neuron simulations are grounded. Four dimensional non-linear differential equations can prove very complex and computationally costly to use in simulation experiments involving significant number of neurons. Consequently, simplified models have been developed that capture key features of neural functions more clearly.

Fitzhugh-Nagumo model

In the Hodgkin and Huxley model, n and $1-h$ have very similar dynamics. Therefore it is natural to simplify it by collapsing these two variables into one. As explained in Gerstner and Kistler (2002), we can

approximate the relation between $(1-h)$ and n as following the equation:

$$h = A - B \cdot n$$

where A and B are dimensionless and positive parameters evaluated to minimise the approximation error. We then replace n and $1-h$ by W, called recovery variable, that represents the position of the projection of the point of coordinates $(n, 1-h)$ on the line represented by the equation $h = A - B \cdot n$.

Furthermore, we can notice that m converges on its stationary value much faster than V, h or n change. We can thus perform the quasi-stationary states approximation that consists in replacing $m(t, V(t))$ by $m_\infty(V)$ in Eq. 2.1.1.

After a few steps, we reach the following two-dimensional model:

$$\frac{dV}{dt} = V - \frac{V^3}{3} - W + I \tag{2.1.2}$$

$$\frac{dW}{dt} = \phi(V + a - bW) \tag{2.1.3}$$

ϕ, a and b are dimensionless and positive parameters. The values of the parameters depend on the characteristics of the actually modelled neuron. These equations have been independently derived by FitzHugh (1961, 1969) and by Nagumo et al. (1999), it is thus called the Fitzhugh-Nagumo model (FHN model).

The interest of this model resides in phase plane analysis. This consists of representing the dynamical system that models the neuron with a vector field in two dimensions. In this plane, the coordinates of the points are the values of the two variables of the system, and to each point we associate a vector representing the derivatives of the variables at that point. This representation gives a good geometrical intuition to predict the system's behaviour starting from arbitrary initial conditions. Since analytical expressions of the solutions of differential equations such as Eq. 2.1.3 are not simple, this is invaluable.

In phase space analysis, much information can be grasped by representing the nullclines. Nullclines are curves that link all the points of the plane at which one of the derivatives is equal to zero. Thus, as illustrated in Fig. 2.6, along the nullclines, the vectors representing the derivatives are either vertical or horizontal. $\frac{dV}{dt} = 0$ implies that $W = V - \frac{V^3}{3} + I$, from Eq. 2.1.3. This is a cubic curve, a hyperbola which is simple to study. $\frac{dW}{dt} = 0$ implies that $W = \frac{V+a}{b}$, from Eq. 2.1.3. This is a line. At the intersection of the two nullclines, both derivatives are null: the system is unchanging, which means that we have an equilibrium point. This

equilibrium point is stable when the intersection is at the external branches of the hyperbola. However, it is unstable when the intersection is at the central part. In that case the system converges to a limit cycle that models the emission of a spike train. This is the powerful aspect of the model: it highlights the fact that the membrane potential, that can be reduced to V, increases with the synaptic input I, and that above a certain threshold, spike trains are generated.

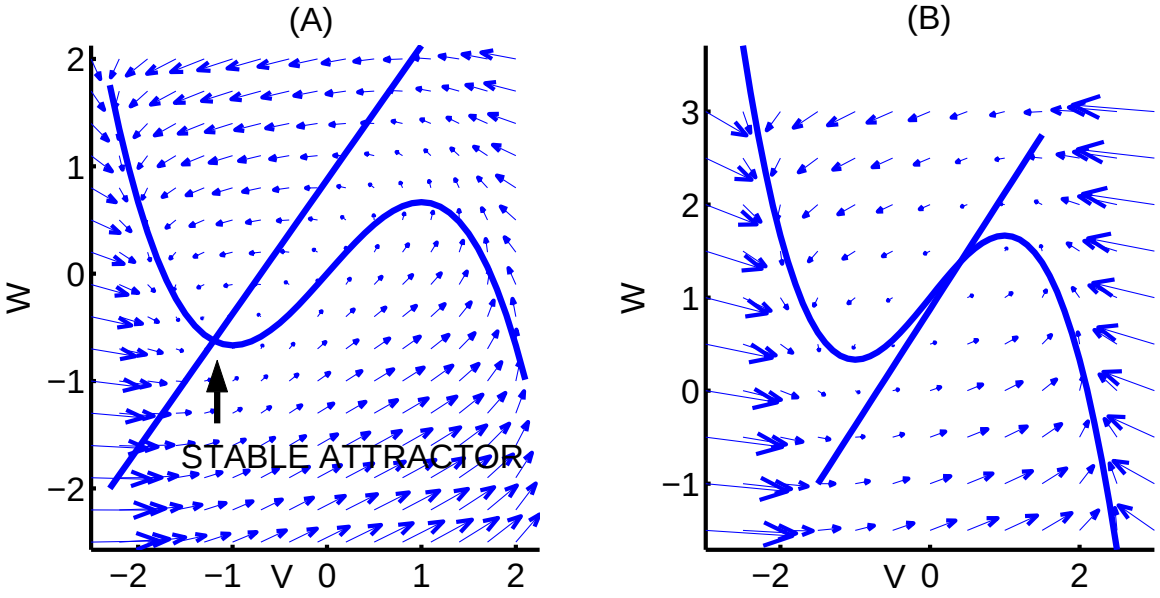


Figure 2.6: Phase plane portrait of the FitzHugh-Nagumo model, for $I = 0$ (A) and $I = 1$ (B). The parameters are the ones used in Cronin (1987) and Koch (1999): $a = 0.7$, $b = 0.8$ and $\phi = 0.08$. The straight line is the nullcline of the variable W, and the curve is the nullcline of V. At the intersection of the nullclines, both derivatives are zero, therefore we are at equilibrium. In both panels, the arrows show the flow $(\dot{V}, \dot{W})$, where the value of $\dot{W}$ is multiplied by 10, and $\dot{V}$ by 2, in order to give a more intuitive understanding that one derivative is zero on each nullcline. Note that this hides the sheer difference in dynamics between V (fast variable) and W (slow variable), which is at the source of the shape of a spike. In (A), the intersection is a stable equilibrium, whereas in (B) it is unstable. However, the system will not diverge but will fall in a limit cycle that models the generation of a regular train of action potentials emitted by a neuron under synaptic stimulation.

There are other two dimensional reductions of the Hodgkin-Huxley model, for example the Morris-Lecar model, that allows the generation of low frequency spike trains.

The Integrate and Fire model

The Hodgkin and Huxley model (HH model) and its two dimensional reductions can precisely describe the generation of the spiking behaviour of the neuron. The spikes that they generate are stereotypical courses of the neuron variables. We have already pointed out that the shape of a spike is not relevant to the information transmitted from one neuron to another. What matters is the relative timings of these spikes. Consequently, in order to simplify neural modelling, we can give up describing the course of neural variables during spike emission and only register when a spike was produced. We have seen in the HH model and its reductions that the generation of a spike occurs when the membrane potential reaches a threshold. What remains fundamental is to model membrane dynamics leading to threshold and spike generation.

Classically, people have used the Integrate and Fire model (IF model). This model was first studied by Lapicque (1907, 1926), even before the neural mechanisms of spike generation were known. Its simplicity (it only involves one variable) and versatility have made it famous, and it has been widely used and studied throughout history Stein (1967) until nowadays Feng (2001), Tuckwell (1988). In Gerstner and Kistler (2002), we see how it can be mapped to the more complete HH model, and generalised to fit biological feature more realistically. One fundamental amelioration of the IF model is the Leaky Integrate and Fire model (LIF), or Forgetful Integrate and Fire model, in which a leak is introduced in the membrane. Simple generalisations have also been made, such as the non-linear LIF seen in for example Abbot and Van Vreeswijk (1993), or the introduction of a time dependent conductance to the model, which induces adaptation, such as in Wehemeier et al. (1989). The Integrate and Fire model is described by a membrane potential equation and a reset mechanism.

In the Integrate and Fire model the membrane equation is:

$$C\frac{dV}{dt}(t) = I(t) \tag{2.1.4}$$

Then, the reset mechanism is stated as follows: when V, the membrane potential reaches V_θ, the neuron sends a spike and V is reset to $V = V_{rest}$.

In the case of the Leaky Integrate and Fire model, the membrane equation follows the equation Eq. 2.1.5

$$C\frac{dV}{dt}(t) = -\frac{V(t) - V_{rest}}{R} + I(t) \tag{2.1.5}$$

The reset mechanism is exactly the same as in the case of the non-leaky Integrate and Fire neuron: when V reaches the threshold potential V_θ, a spike is generated and the membrane potential is reset to the resting

potential, which is the stable state in which the LIF comes back when not submitted to any current ($I = 0$). Then, possibly after a refractory period during which the membrane potential is not influenced by the input currents, the membrane potential resumes its dynamics as defined by Eq. 2.1.4 or Eq. 2.1.5.

There are four parameters: V_θ the threshold potential, V_{rest} the resting potential, R the membrane resistance that controls the leak (or the rapidity at which the LIF forgets that a spike occurred) and C, the membrane capacity. To these, one often adds τ_{ref}, the refractory period. The resting potential has been experimentally measured to be negative. However, in order to make the notations simpler, and without changing anything in the dynamics, the LIF model is often rescaled so that $V_{rest} = 0$. The three other parameters can be chosen to fit the dynamics of the neurons under study. Similarly to the Hodgkin and Huxley model, one can map these dynamics and express them with electronic circuits. In the case of the Integrate and Fire neuron, the circuit simply involves a capacitance C, whereas in the case of the Leaky Integrate and Fire model, we have a resistance and a capacitance in parallel. This is well illustrated in Koch (1999), p. 336. If we assume a constant current I, we can easily integrate Eq. 2.1.5 and express $V(t)$, for $V(t) < V_\theta$.

$$V(t) = I \cdot R(1 - e^{-t/RC}) + V_0 e^{-t/RC} \tag{2.1.6}$$

V_0 is the initial value of the membrane potential, at time $t = 0$. In most cases we will have $V_0 = 0$, because either the neuron is initially at equilibrium at the resting potential $V_0 = V_{rest} = 0$, in the absence of input current, or it has just spiked and then the membrane potential is reset to the resting potential. We can calculate the time T_θ needed for V to reach V_θ, resulting in the generation of a spike. Posing $V(T_\theta) = V_\theta$ and $V_0 = V_{rest} = 0$ in Eq. 2.1.6 we have:

$$V_\theta = I \cdot R(1 - e^{-T_\theta/RC}) \tag{2.1.7}$$

Resulting in

$$T_\theta = -R \cdot C \cdot ln(1 - \frac{V_\theta}{I \cdot R}) \tag{2.1.8}$$

When $V = V_\theta$, a spike is generated, V is reset to V_{rest}, the neuron undergoes a refractory period of τ_{ref}, and then charges again following Eq. 2.1.5. Thus the InterSpike Interval (ISI) is $ISI = T_\theta + \tau_{ref}$. From the ISI, we define a measure of the neural activity, the Firing Rate (FR): $FR = \frac{1}{ISI}$. In the LIF model, we have an explicit expression of FR:

$$FR = \frac{1}{\tau_{ref} - R \cdot C \cdot ln(1 - \frac{V_\theta}{I \cdot R})} \tag{2.1.9}$$

Eq. 2.1.9 holds only for $I > \frac{V_\theta}{R}$, so that $1 - \frac{V_\theta}{IR} > 0$ and we can define its logarithm. Let us define $I_{min} = \frac{V_\theta}{R}$. When $I \to I_{min}$, $I > I_{min}$, $ln(1 - \frac{V_\theta}{IR}) \to -\infty$, $FR \to 0$. This comes from the fact that if I is too small, Eq. 2.1.7 does not have a solution. If I is too small, the leak induced by the non-infinite resistance R prevents the input current from raising the membrane potential to the threshold. This reproduces the phenomenon observed in the Hodgkin-Huxley and Fitzhugh-Nagumo models, illustrated in Fig. 2.6.

Eq. 2.1.9 shows us that the FR is an increasing function of I. Furthermore, when $I \to \infty$, $\frac{V_\theta}{IR} \to 0$, then $ln(1 - \frac{V_\theta}{IR}) \to 0$, so $FR \to \frac{1}{\tau_{ref}}$, which is the upper limit of the Firing Rate of the neuron. These features reproduce the Hodgkin-Huxley model behaviour, and experimentally observed neural behaviour.

In Fig. 2.7, we illustrate the relation between I and FR.

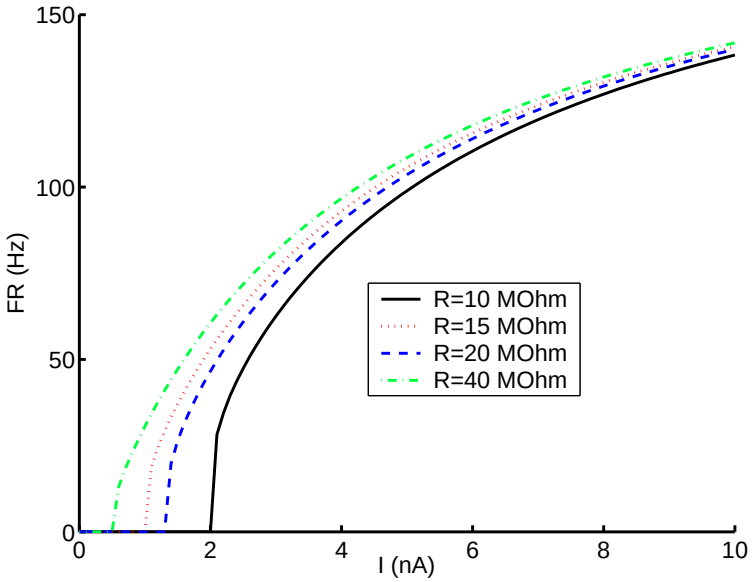


Figure 2.7: Output Firing Rate of the LIF model, as a function of the input intensity. We used classical parameters, in agreement with published results, such as in Shadlen and Newsome (1994), Feng et al. (2006), or Koch (1999). $V_\theta = 20mV$, $C = 1nF$, which makes the time constant of the neuron, $\tau = RC$, vary from 10 to 40 ms. $\tau_{ref} = 5$ ms so the upper limit of FR is 200 Hertz. We can see I_{min}, the lower limit of I at the intersections between the curves and the abscissas axis. As R decreases, the membrane leak increases, so I_{min} increases.

Synaptic processing of spikes

So far, we have modelled neural input as an externally applied current $I(t)$. In practice, neurons are submitted to spikes coming from the axons of other neurons. They have an effect on the membrane potential

at the dendrite, via the synapse. In this section we describe how we model spiking inputs into the current $I(t)$.

The generally accepted synapse description relies on the work done by Katz (1969). When an action potential reaches an axonal button, the neurotransmitters present in the button are released outside of the neuron into the synaptic cleft. The cleft is a space a few nanometres wide between the pre-synaptic neuron's axonal button and the post-synaptic neuron's dendrite. The neurotransmitters then bind to receptors on the surface of the post-synaptic neuron's dendrite, starting a chain of chemical events that results in a change in the post-synaptic membrane potential.

This chain of post-synaptic events is very varied, comprising a multiplicity of ionic channels, and has been the subject of an extended literature. See for example Hille (1992) for a detailed study . For the purpose of this thesis, we will only describe the simplest model, without giving the details of all the separate chemical species involved, even if each of these generates a different kind of dynamics. When the neurotransmitters penetrate the post-synaptic neuron, they force the opening of ionic channels, thus modifying the membrane conductance which in turn creates an input current:

$$I(t) = -g(t) \cdot (V(t) - E) \tag{2.1.10}$$

where $g(t)$ is the membrane conductance, $V(t)$ the membrane potential, E the ionic reversal potential and $I(t)$ the post-synaptic input current. As seen in Gerstner and Kistler (2002), the conductance $g(t)$ is often modelled as an alpha function of the type $g(t) = Ate^{-Bt}$. It is important to notice that, as shown by Gulyás et al. (1993), synaptic transmission is not a deterministic process and that the probability for a synapse to release neurotransmitters upon spike arrival can be as low as ten percent. We will come back to synaptic stochasticity in more detail in subsequent sections. Lastly, we see in Eq. 2.1.10 that the generated current depends on the relation between the membrane potential and the synapse reversal potential. If E is much higher than V_{rest}, then I will be positive, we have an excitatory synapse. Conversely, if E is less than V_{rest}, then I will be negative, we have an inhibitory synapse.

It has been justified in Gerstner and Kistler (2002) that we can take the approximation of stereotypical Excitatory or Inhibitory Post Synaptic Current Pulses (I-EPSP, also known as Excitatory or Inhibitory Post Synaptic Potential). In that case, we do not consider I per se. Instead, we look at the quantity of electric charge transmitted via the ionic channel for each spike. Thus we pose $I(t) = \frac{dQ(t)}{dt} = \sum_{i=1}^{n} \frac{dq_i(t)}{dt}$, where $q_i(t)$ is the total charge transferred through the membrane at time t at synapse number i, and n is the number of synapses. Each time a spike is transmitted to synapse i, at time t_i, the corresponding ionic channels open and a small quantity of charge flows in the soma for a small amount of time: $dq_i(t_i) = g_{i,open}(t_i) \cdot (V(t_i) - E)dt$.

The dynamics of the channel opening $g_{i,open}(t) \cdot (V(t) - E)$ that last dt, are modelled as an instantaneous current by a value called the synaptic weight $w_i = g_{i,open}(t) \cdot (V(t) - E)$ ascribed to each synapse i. At times others than when spikes are transmitted, no charge is transmitted. Consequently we have: $\frac{dq_i(t)}{dt} = w_i \delta(t - t_i)$, where δ is the dirac function. If we neglect the spatial distribution and the dendritic delays, we eventually reach the following simplified expression for the post-synaptic current:

$$I(t) = \frac{dQ(t)}{dt} = \sum_{i=1}^{n} w_i \delta(t - t_i) \tag{2.1.11}$$

where t_i is the time of arrival of one spike at synapse i. Only one spike can arrive at one synapse within the infinitesimal time interval $[t, t + dt]$, because, by the rules of differential calculus, dt is infinitely small. Then, in order to simulate an LIF model receiving spike trains, we only need to feed Eq. 2.1.11 in Eq. 2.1.5:

$$C\frac{dV(t)}{dt} = -\frac{V(t) - V_{rest}}{R} + \sum_{i=1}^{n} w_i \delta(t - t_i) \tag{2.1.12}$$

2.2 Networks of neurons: the brain's comparative advantage

We have so far studied the dynamics of one neuron receiving an electrical input. However, this is not satisfying enough, because the aim of studying such neural dynamics is to understand the behaviour of many neurons put together. In fact the impressing abilities of a brain rely on the interaction of vast numbers of neurons. Studying a single neuron model can also be detrimental to the bigger picture, because one might quickly run out of computing power to simulate too many, too precise models of neurons.

2.2.1 Neural coding

In the previous section, we have already made some progress toward modelling a network of neurons. We can now simulate the dynamics of a neuron whose input is expressed in terms of outputs of other neurons. We have a relatively simple description of how the synapses translate every incoming spike into membrane potential variations. Is that enough? The answer to this question is not straightforward and depends on the type of simulations one desires to perform. We will see further on in this thesis that we have addressed our neural modelling issue from various levels of precision. However, for now let us review several approximations that one can make about the spike trains that constitute neural coding.

Deterministic codes

As we have seen with the dynamical models of spike generation, whether it is the Hodgkin-Huxley or the FitzHugh-Nagumo model, neurons exhibit very distinct behaviours. We even have seen in the HH model

that, within a certain range of parameters, either they emit trains of spikes at a relatively high frequency or they are silent. This is probably why, besides the simplicity of the idea, the first approach to neural networks only considered binary neurons. Either a neuron was activated (state 1) and emitting spikes, or it was silent (state 0). The description of the possibilities of such a network in McCulloch and Pitts (1943) is considered by many as the opening of the fields of neural networks and artificial intelligence.

However, we have seen previously that basic experimental data shows that neurons have a much more rich panel of behaviours. Instead of modelling neurons as all or nothing units, we can at least see them as machines that generate an analogue output. One of the most classic ways to describe a spike train with a continuous measure is to evaluate the frequency of the spikes: the Firing Rate (FR).

Based on the FR, numerous models of neural units have been designed that enable us to study their behaviour in networks. Besides the biologically inspired models presented in the previous section, various simple and computationally efficient neural unit models have been tried. The simplest unit that we can think of is probably a neuron whose output FR would be the linear summation of all the input FRs:

$$FR_{out} = \sum_{i=1}^{n} w_i FR_i + b \tag{2.2.1}$$

where the index i represents the synapse i, and the synaptic weights are dimensionless. b, called the bias, is a constant value, independent from the FR, that is often added to the model. The interest of this model is that it enables linear representation of neural processing. Linear representations are very useful, because on one hand they are more precise than binary representations, they are continuous, but on the other hand they are still relatively easy to understand, model, and analyse. The power resides in the fact that we can represent the synaptic connections of a network in a matrix W, where the elements $W_{i,j}$ are the synaptic weights between neuron j and neuron i, and that in order to predict the behaviour of the system, one only has to iterate the multiplication of the vector of FRs by the matrix W. Through the use of the powerful algebraic tools of matrix calculus, we can easily predict the behaviour of such a linear dynamical system. We will not retrace here the whole history of engineering and linear calculus, but we can mention as an example the use of the eigenvalues of the connection matrix as a tool to determine the convergent or divergent nature of the system and its stable states. In computational neuroscience, since we are trying to understand networks that comprise large number of neurons, being able to predict the dynamics of such networks without having to simulate them, merely by looking at the values in the matrix (synaptic weights) is invaluable, and that is why initially people focused intensely on such linear models. This is a classic approach to difficult domains, and people shift to non-linear modelling only if the linear approximation is too far from reality.

In order to have a more realistic description, or a more ad-hoc neural function, a classical method is to apply

what is called an activation function to the linear summation of Firing Rates. The neural equation then becomes:

$$FR_{out} = F(\sum_{i=1}^{n} w_i FR_i + b) \tag{2.2.2}$$

We have represented in Fig. 2.8 a few classical activation functions, such as the sigmoid. According to Eq. 2.1.12 and Eq. 2.1.5, Fig. 2.7 represents the activation function that models the Leaky Integrate and Fire model.

These models have been very useful in artificial intelligence and engineering applications, and we will illustrate some of these applications in subsequent sections (2.2.2).

Noisy spike trains: statistical coding

So far the coding of the messages between neurons has been assumed to be deterministic, whether it was a binary code or an analogous value such as the FR. However, neurons are intrinsically noisy. As we pointed out in section 2.1.2, even the transmission of a signal through a synapse upon arrival of a pre-synaptic spike is a probabilistic event. Furthermore, the inputs to neurons are not steady nor predictable except in vitro. Thus a spike train is much more accurately described as a stochastic process: a family of random variables indexed on time. Classically the variable used to describe the statistics of spike trains is the InterSpike Interval (ISI), as seen in Gerstner and Kistler (2002) for example.

In order to describe the spike train as a stochastic process, we study the InterSpike Intervals as random variables. If we do not know the expression of the probability distribution of the ISIs, we can evaluate its mean and its central statistical moments. The central moment of order r of a random variable is defined as follows:

$$M_r(X) = E((X - E(X))^r) \tag{2.2.3}$$

where E is the mean of the random variable. The second order central moment is called the variance of the random variable:

$$Var(X) = \sigma^2(X) = E((X - E(X))^2) = E(X^2) - (E(X))^2 \tag{2.2.4}$$

σ so defined is called the standard deviation of the random variable.

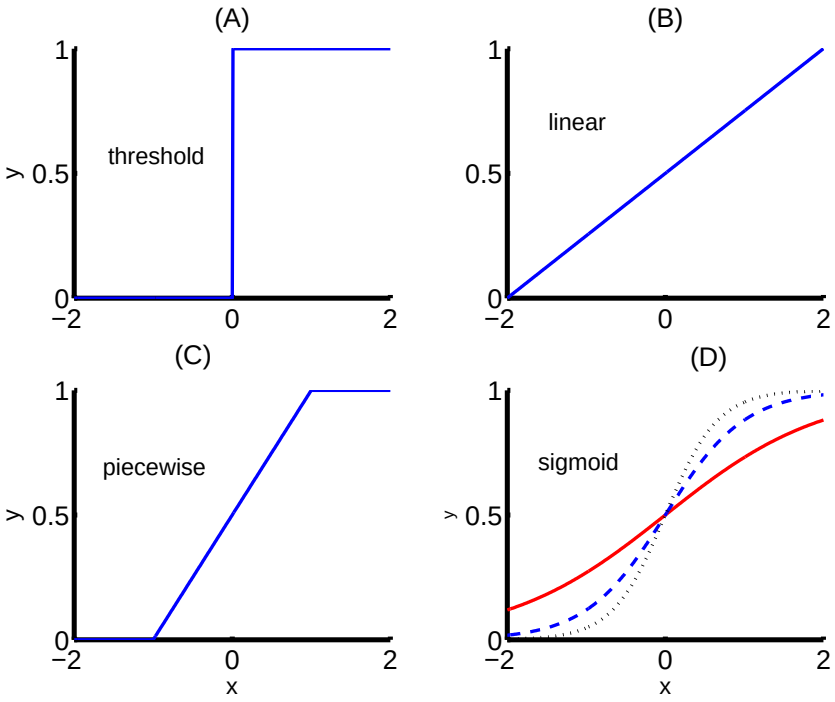


Figure 2.8: Illustration of various activation functions. x and y are normalised variables that respectively represent the linear summation of FRs and the output FR. In (A), we have a threshold function that in fact simulates binary coding. In (B), we have the linear function. In this case, the activation function can be the identity. This is not a very realistic function, because it might induce negative or arbitrarily high output FRs. Therefore, in order to comply with the limitations seen in biology and keep a stable system we can use the piecewise function shown in (B). Between two limit values, it is a linear function, and outside of them it is constant, either to the maximum Firing Rate or to a null activity. Lastly, in (D), we present a smoothed version of the piecewise function: the sigmoid. Its equation is $y = \frac{b}{1+e^{-ax}}$, where $b = 1$ and a equals 1, 2 or 3 (Respectively red plain line, blue dashed line or black dotted line).

An example: Gaussian random variables

A random variable is said to have a Gaussian distribution, or normal distribution, of mean μ and standard deviation σ if its density of probability is:

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{2.2.5}$$

It is a very useful model, because it is easy to manipulate: let X and Y be two random variables with Gaussian distributions of respective means μ_x and $\mu_y$and standard deviations σ_x and σ_y. Let us also assume that they are correlated with a correlation coefficient c. The correlation coefficient between two random variables is defined in 2.2.9. Then the variable $Z = X + Y$ has a Gaussian distribution of mean

$$\mu_z = \mu_x + \mu_y$$

and of standard deviation

$$\sigma_z = \sqrt{\sigma_x^2 + \sigma_y^2 + 2c\sigma_x\sigma_y}.$$

This easy rule for adding Gaussian distribution makes it a powerful modelling tool.

Furthermore, Gaussian distributed variables are often used for modelling because of the theorem central limit: let X_1, ... X_n be n independent random variables from the same probability distribution with finite mean μ and variance σ^2. Then the random variable

$$Y_n = \frac{1}{\sigma\sqrt{n}} (\sum_{i=1}^{n} X_n - n\mu)$$

converges in distribution to the Gaussian distribution of zero mean and unity variance. This is very powerful because it enables modelling the combination of variables coming from an unknown distribution as a Gaussian distribution that is very easy to use.

The Poisson Process

A frequently made assumption is to consider that the probability of emitting a spike at a precise time only depends on the timing of the last spike and on the input to the neuron. This neglects the observed fact that a neuron adapts so that its FR decreases with time when submitted to a strong input. In this condition, the spike train can be considered as a renewal process. Renewal processes, and in particular the Poisson process are well adapted to describe sequences of events such as client arrivals at a help desk, telephone calls to call centres, or in our case spikes arriving at synapses. They have been extensively studied in, for example Cox (1962) or Cox and Isham (1980). Up to now most of the research in neural networks that uses statistical models of spiking has been done under the assumption that the spike trains are Poisson processes. See for

example Feng (2003a). The Poisson process assumes that the mean of the number of event occurring in a period of time Δt is equal to its variance:

$$E(N(\Delta t)) = \lambda \Delta t = Var(N(\Delta t))$$

λ is called the rate of the Poisson process. It is also the Firing Rate of the spike train. Let us try to find the probability distribution of the InterSpike Intervals, $f_{ISI}(t)$. First we define the survivor function, which measures the probability that the neuron does not emit any spike within an interval t after the last emitted spike at time t_{sp}:

$$S(t|t_{sp}) = 1 - \int_{t_{sp}}^{t} f_{ISI}(u)du \tag{2.2.6}$$

In Eq. 2.2.6, the integral is the probability of emitting a spike between t_{sp} and t. $S(t_{sp}|t_{sp}) = 1$, because no spike can be emitted in the same time as the previous one, and $S(t|t_{sp})$ tends to zero when t tends to infinity, because probabilistically a spike will be emitted at one time or another. We then define the decay rate, or hazard rate of the survivor function:

$$\rho(t|t_{sp}) = -\frac{\frac{dS(t|t_{sp})}{dt}}{S(t|t_{sp})} \tag{2.2.7}$$

The hazard rate describes the chance or the hazard of emitting a spike at a time t, it is the intensity of the process. In the case of a Poisson process of rate λ, this intensity is constant equal to λ.

By integrating Eq. 2.2.7 and introducing the result in Eq. 2.2.6, we obtain the following definition for the distribution of InterSpike Intervals:

$$f_{ISI}(t) = \rho(t|t_{sp}) exp[-\int_{t_{sp}}^{t} \rho(u|t_{sp})du]$$

Since $\rho(t|t_{sp}) = \lambda$, we get the exponential law for the distribution of InterSpike Intervals:

$$f_{ISI}(t) = \lambda e^{-\lambda t}$$

Thus, the mean ISI will be:

$$< ISI >= \int_{+\infty}^{0} \lambda t e^{-\lambda t} dt$$

Via integration by parts we reach

$$< ISI >= \int_{+\infty}^{0} e^{-\lambda t} dt = \frac{1}{\lambda}$$

If we define the Firing Rate to be $FR = \frac{1}{<ISI>}$, then we reach:

$$FR = \lambda \tag{2.2.8}$$

Similarly, the variance of the ISI is

$$Var(ISI) = \frac{1}{\lambda^2}$$

Correlation

We have so far described statistical models of single spike trains, but the neural input is constituted of hundreds if not thousands of simultaneous such spike trains. Experimental data of Zohary et al. (1994) shows that besides their individual statistical signatures, their mutual correlation even when it is weak has significant effects on the output of the neuron. The same paper and other such as Sheth et al. (1996) or Romo et al. (2003) show that in biology many neurons tend to send correlated spike trains, especially if they are close to each other or perform the same function.

The effect of correlated input on simple neural models has been studied by Salinas and Sejnowski (2001), and more specifically on the LIF model by Feng (2003b) and Feng and Brown (2000). As seen in, for example Feng and Brown (2000), the standard measure of the correlation between the inputs to two synapses i and j receiving spike trains with InterSpike Intervals ISI_i and ISI_j is:

$$c(i,j) = \frac{< (ISI_i(t) - < ISI_i(t) >)(ISI_j(t) - < ISI_j(t) >) >}{\sqrt{< (ISI_i(t) - < ISI_i(t) >)^2 >< (ISI_j(t) - < ISI_j(t) >)^2 >}} \tag{2.2.9}$$

where $< x >$ denotes the mean of x for any random variable x. This correlation coefficient is then used to compute the total synaptic input with more precision than a mere linear combination of FR as we described in section 2.1.2.

The Moment Neural Networks

The assumption of a Poissonian spike train has been very helpful as a tool to study statistical neural networks. However, assuming that the mean spike count is equal to its variance is oversimplifying biological data. Furthermore, as shown in Feng et al. (2006), the output of a LIF model submitted to Poissonian spike trains is not exactly a Poisson process itself. Thus, recently, in Feng et al. (2006), Moment Neural Networks (MNN) have been developed. The aim of the MNN is to characterise the input-output relationship of the neural units of the network, as has been done before by for example Abbot and Van Vreeswijk (1993), but

releasing the Poisson assumption. It creates a LIF input/output mapping of the two first moments of spike trains seen as renewal processes. It is a much more precise and powerful scheme than the Poisson process approach and "should be considered as an attempt toward a theory of computation with stochastic systems" Feng et al. (2006). In this section we present some of the theoretical tools that they have developed for the Leaky Integrate and Fire feedforward network.

The output of a neuron in layer $k+1$ depends on the inputs from layer k. There are $p(k)$ neurons in layer k. w_{ij} is the synaptic strength from neuron j of layer k to neuron i of layer $k+1$. Spike trains are characterized by the two first moments of the statistical distribution of the InterSpike Intervals (ISI).

The membrane dynamics of the LIF neuron is described by Eq. 2.1.5, and can be reformulated as follows:

$$dV_i^{(k+1)}(t) = -\frac{1}{R \cdot C} \cdot (V_i^{(k+1)}(t) - V_{rest})dt + I_{i,syn}^{(k+1)}(t) \tag{2.2.10}$$

where R and C are the membrane capacitance and resistance, $V_i^{(k+1)}(t)$ the membrane potential of neuron i in layer $k+1$, V_{rest} the resting potential. For further notation, we pose $L = \frac{1}{R \cdot C}$, the leak. $I_{i,syn}^{(k+1)}(t)$ is the synaptic input defined by:

$$I_{i,syn}^{(k+1)}(t) = \sum_{j=1}^{p^{(k)}} w_{ij}^{(k)} dN_j^{(k)}(t) \tag{2.2.11}$$

where $w_{ij}^{(k)}$ is the magnitude of the postsynaptic potentials incoming from neuron j of layer k to neuron i of layer $k+1$, $N_j^{(k)}$ is the renewal processes that represent the incoming spike train arriving from the j_{th} neuron of layer k, $p^{(k)}$ the number of synapses from neurons of layer k to neuron i of layer $k+1$. If a synapse is inhibitory, then its weight is negative: $w_{ij_{Inh}}^{(k)} \leq 0$. As soon as the membrane potential of a neuron goes above the potential threshold V_θ, it is reset to V_{rest}.

In order to describe the efferent spike train from neuron i of layer k, Feng et al. (2006) introduce two variables. The first one is the mean intensity of the output:

$$\mu_i^{(k)} = \frac{1}{< ISI_i^{(k)} >} \tag{2.2.12}$$

where the brackets $\langle\rangle$ denote the expectation of the random variable between them.

The second variable is the 'noise' of the output, defined as follows:

$$(\sigma_j^{(k)})^2 = \frac{\left\langle (ISI_j^{(k)})^2 \right\rangle - \left\langle ISI_j^{(k)} \right\rangle^2}{\left\langle ISI_j^{(k)} \right\rangle^3} = \frac{Var(ISI_j^{(k)})}{\left\langle ISI_j^{(k)} \right\rangle^3} \tag{2.2.13}$$

Then they use theorem 1 of Feng et al. (2006), called the usual approximation scheme (UAS). The UAS theorem is expressed as follows:

Let ISI_1, ISI_2... be a series of independent and identically distributed random variables (for example Inter-Spike Intervals), denoted by ISI, of which the three first central moments exist and are noted λ (mean ISI), α^2 (variance of ISI) and λ_3 (third moment), and let $\{N_t : t \geq O\}$ be the corresponding renewal process. Then the usual approximation takes the following form:

$$\begin{aligned} dN_t &\approx \frac{dt}{\lambda} + \frac{\alpha}{\lambda^{3/2}} dB_t \\ B_0 &= \frac{\alpha^2 - \lambda^2}{2\lambda^2} \end{aligned} \tag{2.2.14}$$

where B_t is the standard Brownian motion. The proof of this theorem is beyond the scope of this thesis, but is outlined in Feng et al. (2006). We can then use this result, the definition of central moments as seen in 2.2.3, and the two variables introduced in 2.2.12 and 2.2.13 to make the following approximation that states that the renewal process that models the spike train coming from neuron i at time t in the time interval dt can be modelled by a continuous process as follows:

$$dN_i^{(k)} \approx \mu_i^{(k)} dt + \sigma_i^{(k)} dB(t) \tag{2.2.15}$$

In that case, we can read $\mu_i^{(k)}$ and $\sigma_i^{(k)}$ as the mean and standard deviation of the approximated continuous input from neuron i of layer k.

In order to combine all the approximated continuous inputs from the neurons of layer k, we need to multiply each of them by its corresponding synaptic weight w_{ij}^k, that can be negative or positive. The postsynaptic inputs can not have a negative standard deviation, thus we write:

$$w_{ij}^k \cdot dN_i^{(k)} \approx w_{ij}^k \mu_i^{(k)} dt + \left| w_{ij}^k \right| \sigma_i^{(k)} dB(t) \tag{2.2.16}$$

Since $B(t)$ is the Brownian motion, at every moment t_0 $dN_i^{(k)}(t_0)$ has a normal distribution of mean $\mu_i^{(k)}$ and standard deviation $\sigma_i^{(k)}$, so we can add these correlated Brownian motions by following the rules for Gaussian distributions presented in 2.2.5. We thus obtain the following synaptic input, by combining, Eq.2.2.11 and Eq.2.2.15:

$$I_{i,syn}^{(k+1)}(t) = \bar{\mu}_i^{(k+1)} dt + \bar{\sigma}_i^{(k+1)} dB_t \tag{2.2.17}$$

where

$$\bar{\mu}_i^{(k+1)} = \sum_{j=1}^{p(k)} w_{ij} \mu_j^{(k)} \tag{2.2.18}$$

$$(\bar{\sigma}_i^{(k+1)})^2 = \sum_{m,n=1}^{p(k)} c_{m,n}^{(k)} |w_{i,m} w_{i,n}| \sigma_m^{(k)} \sigma_n^{(k)} \tag{2.2.19}$$

where $c_{m,n}^{(k)}$ is the correlation coefficient between the incoming spike train from neuron m, layer k and the spike train from neuron n, layer k, as defined in Eq. 2.2.9.

Then we derive, using Siegert's expression as in Feng (2003a), the central moments of the output interval distribution. This derivation is beyond the scope of this thesis, but we give their results below. The first moment of the output of neuron i of layer $k+1$ is:

$$\left\langle ISI_i^{(k+1)} \right\rangle = \frac{2}{L} \int_{A_i^{(k)}(V_{rest})}^{A_i^{(k)}(V_\theta)} g(x) dx \tag{2.2.20}$$

Limits of the integral:

$$A_i^{(k)}(y) = \frac{yL - \bar{\mu}_i^{(k)}}{\bar{\sigma}_i^{(k)} \sqrt{L}} \tag{2.2.21}$$

$L = \frac{1}{R \cdot C}$ is the time constant of the neuron, and the function g is defined as:

$$g(x) = e^{x^2} \int_{-\infty}^{x} e^{-u^2} du \tag{2.2.22}$$

The second moment, or variance of ISI $\left\langle (ISI_j^{(k+1)})^2 \right\rangle - \left\langle ISI_j^{(k+1)} \right\rangle^2$ is expressed as follows:

$$Var(ISI_i^{(k+1)}) = (2 \cdot R \cdot C)^2 \int_{A_i^{(k)}(V_{rest})}^{A_i^{(k)}(V_{th})} e^{x^2} \left\{ \int_{-\infty}^{x} e^{-u^2} g^2(u) du \right\} dx \tag{2.2.23}$$

These two central moments of the output of neuron i in layer $k+1$ is all we need to analytically compute the output of the next layer. We thus have a mapping that preserves the first and second order statistics of the spike trains. In Fig. 2.9, we present a graphical illustration of the mapping, for a correlation between spike trains clamped to $c = 0.1$. This illustration shows the striking importance of the second moment of incoming spike trains for the two first moments of the output.

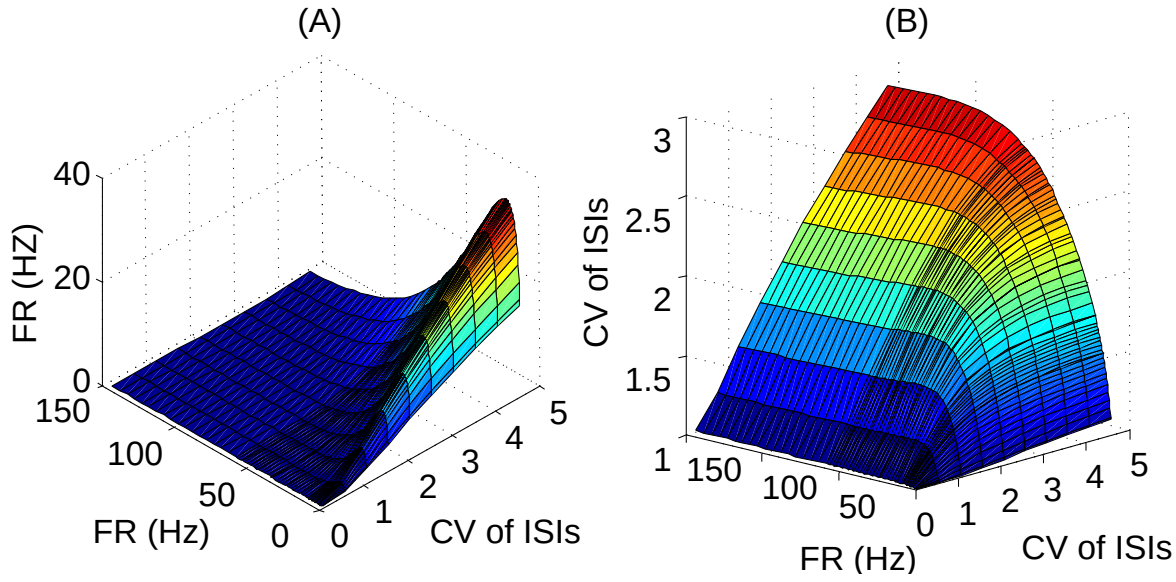


Figure 2.9: Graphic representation of the moment mapping, in the homogeneous case, with 300 excitatory neurons and a ratio between the total weight of inhibitory inputs and the total weight of excitatory inputs $-\frac{\sum_{i=1}^{p_I} w_{ij}^I}{\sum_{i=1}^{p_E} w_{ij}^E} = 3$, *where* p_I *is the number of inhibitory inputs,* p_E *the number of excitatory inputs,* w_{ij}^I *negative synaptic weights* w_{ij}^E *positive ones. Panel (A): output FR as a function of the input FR and the coefficient of variation (CV) of the input ISIs. Since the inhibitory inputs are stronger than the excitation, the output FR is low for the high input FR and high input CV: in Eq. 2.2.20,* g *is positive, increasing. If the input FR is large, then* $\bar{\mu}^{(k)}$, *defined in Eq. 2.2.18 is large and negative thus the bound of the integral in Eq.2.2.20 are large and positive, where* g *is large. So* $< ISI >$ *is large and thus the FR is small. However, a large CV means a large input standard deviation, thus a large* $\bar{\sigma^{(k)}}$, *which leads to a small difference between the bounds of the integral in Eq. 2.2.20. This leads to a smaller* $< ISI >$, *i.e. a larger output FR, as we can see for the lower input FRs. Panel (B): output CV of the ISIs as a function of the input FR and the CV of the input ISIs.*

Spike coding

One of the pending issues of statistical processing of spike trains is the time involved. Since the measured FRs rarely exceed the order of 100Hz, the time needed to gather 100 ISIs, a minimum data sample to evaluate statistical moments of a spike train, would be one second. Living beings reactions require a much faster processing. In Vanrullen (2002), a model is proposed that takes into account biological requirements of very fast processing and structural understanding of visual inputs: humans can recognize a means of transportation from an animal in roughly 150 ms. Bearing in mind the number of cortical layers involved in such high level visual interpretations, and the time characteristics of neuronal firing, he builds a strictly feedforward model in which very few spikes per neurons are used. This model challenges the classical statistical approaches and claims to use precise spike timing.

More recently, some have also started to put the statistics aside and model large networks of neurons spike by spike. They compared the strengths of various spiking models in terms of their capacity to reproduce biological features as well as in terms of computing cost in Izhikevich (2004), and designed their own model of a spiking neuron in Izhikevich (2003) that they use to study the stability of firing patterns in a sparsely connected network in Izhikevich et al. (2004) and Izhikevich (2006).

Population coding

Seeing the impressive number of neurons that send inputs to one neuron, it is natural to study the neural code at the level of a population. Furthermore, there are time constraints that in many cases do not allow a neuron to fire more than a few spikes during the observed phenomenon. This prevents any reliable statistical processing of the spike train issued from one neuron. However, doing statistics over populations of neurons enables us to consider only very short time windows. Various population coding schemes are presented in Gerstner and Kistler (2002). It has been formally proved that an infinite population of either perfect or leaky IF units uniformly distributed in phase will respond instantaneously to any supra-threshold stimulus.

Let us illustrate the fundamental idea: the FR of one neuron is defined as the number of emitted spikes per second. To practically measure it, we count the number N of spikes produced within a sufficiently long time window Δt, and divide one by the other. Δt needs to be long enough to get a sufficient number of spikes to produce a reliable statistical measure. If we count the spikes produced by n neurons, we will on average in a homogeneous population get n more spikes during Δt. Thus, the time needed will be divided by n. If, at $t = 0$, the membrane potential of the neurons are randomly distributed following the distribution of potentials that corresponds to the stationary state, then, when the number n of neurons tends to infinity, the size of the time window $\frac{\Delta t}{n}$ tends to zero. With large enough homogeneous populations of neurons, we can do statistical measures in arbitrarily short measuring times. A homogeneous population of

neurons is a population of similar neurons that receive the same inputs and whose initial states are drawn randomly from the same distribution. They might only be connected to the same external input, or also be laterally connected. However, in that case the lateral connection strength should be all the same: usually, for n neurons we have $\forall i, j \in [0,n], w_{i,j} = \frac{J_0}{n}$, where J_0 is the total strength incoming to one neuron and is arbitrarily decided.

2.2.2 Network structures

Learning

Learning capacities are a prerequisite for systems that aspire to model intelligence. A neural network without learning or at least adaptive capacities would not have much interest in terms of explaining long term behaviour of living beings or in terms of providing pioneering ideas for artificial intelligence. Thus, in Haykin (1999) it is stated that what structures a neural network is its learning rule. Let us describe some of the most fundamental learning rules of neural networks.

The fundamental idea of learning in neural networks has been formulated by Hebb, in Hebb (1949). He stated a learning rule that was localised at the junction between neurons, and did not depend on the global picture: if the post-synaptic activity is correlated to the pre-synaptic activity, then the connexion between the two neurons is strengthened (potentiated). In his own terms:

"When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic changes take place in one or both cells such that A's efficiency as one of the cells firing B, is increased".

At the time, this was a postulate, not grounded on biophysical evidence. Only much later Kelso et al. (1986) has managed to experimentally exhibit long lasting synaptic changes induced by joint activity of pre and post-synaptic neurons that resemble the Hebb rule.

There are various ways to adapt Hebb's learning rule. One of the main additions is to define a decrease of the influence between neuron A and neuron B if A fires when B does not or vice versa. Furthermore, it is usually assumed that the modification happens via synaptic plasticity. A classic mathematical expression of it is expressed in Sejnowski (1977) as follows:

$$\Delta w_{i,j} = \eta(V_i - < V_i >)(V_j - < V_j >))$$

where $\Delta w_{i,j}$ is the change of synaptic weight, η the learning coefficient and $< V_i >$ the average neural

activity over a suitable time. This is the formal expression of the rule that increases the synaptic weights in case of positive correlation between pre and post synaptic activity, and decreases them in case of a negative correlation. In terms of spikes, learning rules could be formulated as follows: if a synapse receives a spike a short time before the post synaptic neuron emits a spike, then the synapse is potentiated. If it receives it a short time after, it is depressed. If the time difference is large, then nothing is changed. Such rules are called Spike Timing Dependent Plasticity (STDP) and have recently proved powerful and promising, as Rao and Sejnowski (2001) or Di Paolo (2002) show. The figure Fig. 2.10 illustrates the STDP rule.

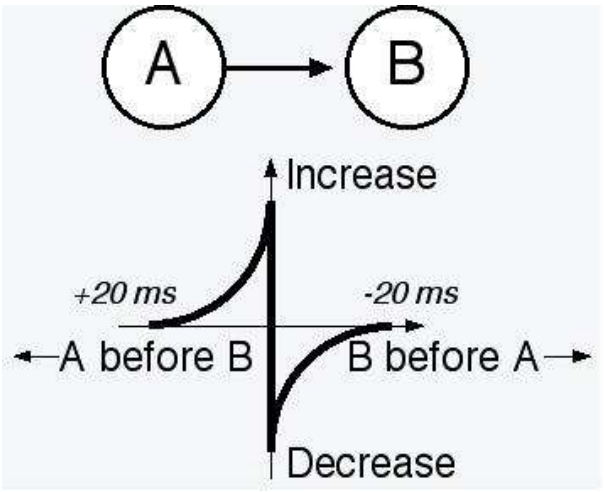


Figure 2.10: Illustration of STDP. Adapted from the website of Per Jesper Sjostrom www.dendrite.org/ jesper/what/What.html

On a more 'engineering' side of things, a significant amount of learning rules has been derived under the gradient descent method. They are called supervised learning methods because an observer is supposed to check the output of the network against an expected input (both depending on the input, of course) and measure the error. Then it is possible to know the partial derivatives of the error against the parameters of the system (gradient of error), in our case the synaptic weights. Modifying these parameters in the local direction of the negative gradient will lead the system into a minimum of error. The problem of local minima can be overcome by adding noise to the gradient descent. Fig. 2.11 illustrates the gradient descent method.

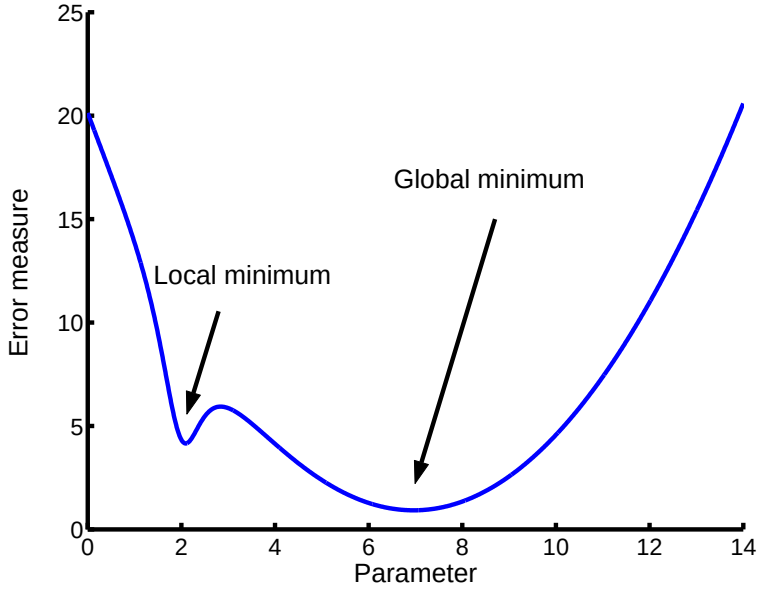


Figure 2.11: Illustration of the gradient descent method. If we follow the line of the steepest gradient we will end up in a local minimum. Adding noise will ensure that we reach a global minimum.

Feedforward networks

A feedforward neural network is constituted of successive layers. Neurons receive inputs from the previous layer and communicate their outputs to the next layer, although there can be lateral connections. Initially neural descriptions of the sensory system involved feedforward networks. Feedforward networks have been very successful at introducing neural networks in machine learning and artificial intelligence. Their great strength is their ability to discriminate categories within a data set and to generalise so that they can discriminate new data. One of the most famous discoveries was the perceptron, formalised by Rosenbaltt (1958). The perceptron illustrates very simply the gradient descent inspired learning rules. We have a set of inputs represented as vectors of neural activity $x(i)$the size of the number of synapses of the neuron. The vectors $x(i)$ are associated with corresponding desired output $d(i)$. The output of the neuron is $y(i) = x^T(i)w$ where w is the vector of synaptic weights of the neuron. The error is $e(i) = d(i) - y(i)$, and we take $E(i) = e^2(i)$ as cost function. The gradient is thus $\frac{\partial E}{\partial W_k} = -2x_k e_k(i)$. Thus, normalising the rule in order to change the weights in the steepest descent direction, we get the update rule, at the i_{th} step:

$$\Delta W(i) = \eta e(i)x(i)$$

.

The perceptron was later proved by Minsky and Papert (1969) to be limited to linearly separable problems. Furthermore, they casted doubts into the possibilities of generalising it as we see in the following quote from their book:

"The perceptron has shown itself worthy of study despite (and even because of!) its severe limitations. It has many features to attract attention: its linearity; its intriguing learning theorem; its clear paradigmatic simplicity as a kind of parallel computation. There is no reason to support that any of these virtues carry over to the many-layered version. Nevertheless we consider it important to elucidate (or reject) our intuitive judgement that the extension to multilayer systems is sterile."

Their intuition was later proved wrong with algorithms such as back propagation or radial-basis functions. Rumelhart and McClelland (1986) published a book that is symbolic of the regain of interest in multilayer perceptrons. In parallel, unsupervised learning structures like the Self Organising Maps have been developed by Kohonen (1995). These structures learn how to find categories in the presented data through training and competition.

Recurrent networks

Contrary to feedforward networks, in recurrent networks there can be feedback connexions coming back from a posterior layer to earlier neurons. Currently people are starting to realise the fundamental importance of those feedback connexions. Classical layered architectures with feedback include the Simple Recurrent Network (SRM) described by Elman (1990) and the Recurrent Multi Layer Perceptron (RMLP), as used by Puskorius et al. (1996). In the RMLP, each layer has a delayed recurrent loop to itself, enabling the network to not only learn static patterns of input but also patterns evolving with time.

Generalising the feedback connexions, we can let aside the notion of successive layers and feedback connexions, and see a neural network as a population of connected units with no hierarchical order. In Edelman (1987), this is called a re-entrant network. If the network is made of N neurons, we can describe the state of the network by the N dimensional vector x whose elements are the activities of the neurons. Assuming that a first degree differential equation describes the membrane dynamics of each neuron, like in the LIF model, we get the following description for the membrane potentials of the network:

$$C_j \frac{dV_j(t)}{dt} = -\frac{V_j(t)}{R_j} + \sum_{i=1}^{N} W_{j,i} x_i(t) + I_j(t) \tag{2.2.24}$$

where V_j is the membrane potential of the neuron j, $W_{j,i}$ is the synaptic weight from neuron j to neuron i, I_j an external current applied to neuron j, C_j and R_j parameters of the neuron j.

An activation function ϕ is applied to the membrane potentials in order to define the output activity of the neuron: $x = \phi(V)$. ϕ can take many shapes. It can be a Heavyside function, by virtue of which if $V > 0$ then $x = 1$, else $x = 0$ or even -1 as described in the first version of the famous Hopfield network model in Hopfield (1982). ϕ can also be a sigmoid as in Hopfield (1984) or even the translation of the membrane potential course into spikes Hopfield (1995). Of course, such a spiking model can be based on the LIF neuron. In fact, in Amit and Brunel (1997) a network of such Integrate and Fire neurons is studied extensively and shown to have the ability to stabilise at a low level spontaneous activity, under certain reasonable parameters settings. We will see later on in this thesis how we intend to use such a network to model decision making. Furthermore, it has been shown that after learning such networks can develop local attractors of higher activity to which they will converge if presented with an input similar to the learned pattern. This is an idea that has been uncovered in simpler networks by Hopfield as we will see in the next paragraph.

Hopfield showed that a re-entrant neural network under certain conditions (W symmetric, for all i, $W_{i,i} = 0$) can act as a content addressable memory: the network activity is contained in a finite domain and will converge from any initial condition to a local minimum of a specific energy function, one of the attractors of the network. This energy function is also called a Lyapunov function, because its existence and properties enable us to guarantee the stability and convergence of the network, according to the Lyapunov theorem. By adapting the synaptic weights, one can decide which states of the Hopfield net will be attractors. An initial condition close to one of these attractors will most likely be in its basin of attraction. If a state of the network is associated to a juxtaposition of letters, one of the attractors is a specific sentence, and we initiate the network on an incomplete version of this sentence, the network will converge to the stable state representing the whole sentence. This is why Hopfield nets are considered to be content addressable memories.

The Hopfield nets are still limited because they use a deterministic representation of neural activity. In order to go beyond the Firing Rate characterisation of neural activity, people soon began to look at synchronicity. The notion of synfire chains, in Abeles (1991) and Abeles (2002), whereby neurons of a same group fire precisely and repeatedly at the same time was a major step in understanding the brain via recurrent network modelling. Recently, attention has been given to transmission delays. According to Izhikevich (2006), in re-entrant networks these delays give rise to polychronicity. Polychronicity is a pattern whereby a group of neurons fire repeatedly in a precise order and with surprisingly precise timing, but not necessarily simultaneously. In their work, these polychronous patterns are however transient.

2.3 Decision making

2.3.1 Neurology of decision making

In order to make a decision, the organism first needs to translate the stimuli into arguments in favour of one or the other decision. In the case of the kinematogram task, experimental data shows that the area MT or V5 is fundamental in this process. As seen in Britten et al. (1993), it is believed to produce the evidence on which the decision is based. In their paper, it is shown that the response of neurons in MT increases with the coherence of the kinematogram in its preferred direction and decreases in the opposite direction. They found that many MT cells have an approximately linear dependence of mean firing rate on stimulus correlation. This led Wang (2002) to assume a linear relationship in his decision making model, as we will see in the next section.

In the literature about the visual, kinematogram-based decision task, it has been consistently shown that the experimenter can guess, with acceptable accuracy, the decision of an animal that is performing the decision task, merely from reading the activity of specific neurons of area MT across a period of time. For example, in Britten et al. (1996), it is experimentally shown that neurons from area MT act as if their receptive fields were centred on kinematograms displaying a certain direction.

However, the MT activity does not constitute the decision making itself because the animal's actual decision can contradict the evidence produced by the MT neurons. Furthermore, the evidence produced by MT neurons fluctuates in time, following the instantaneous kinematogram display, but the neural activity involved in the decision making must be able to keep in memory the quantity of evidence for one or the other alternative, despite the stimulus variations. The neural processes by which the animal actually takes a decision must consequently accumulate evidence related to the alternative decisions. The accumulation of evidence must be implemented in complex networks of neurons that exhibit specific dynamics, whereby neural activity correlated to one alternative grows on a long timescale, taking into account evidence produced over a period of time, and remains stable as working memory.

Such neural variables that measure and accumulate the quantity of evidence in favour of one decision have been experimentally exhibited. In Huk and Shadlen (2005) it is shown that activities of neurons in area LIP accurately reflect a simple model of integration of evidence like the one shown in Fig. 2.14. As we can see in the results of recording experiments, reported in Shadlen and Newsome (2001), Smith and Ratcliff (2004), or Mazurek et al. (2003), the activity of LIP neurons reflects a competition. On the onset of the stimulus, their activities increase together until the neurons supporting one decision take over whilst their opponents' activity decreases to a very low level. This is illustrated in Fig. 2.12. Furthermore, Platt

and Glimcher (1999) have shown that activities of LIP neurons that accumulate evidence for decisions are influenced by the prior probability of specific answers. It is thus established that neurons in area LIP accumulate evidence over time and keep it stable for extended period of time even if the stimulus fluctuates or disappear. What is not known is how this accumulation happens, what is the connectivity of the network underlying this pattern of neural activity. Some hypothesis have been formulated, modelled, and compared to experimental results, as we will see in the next section.

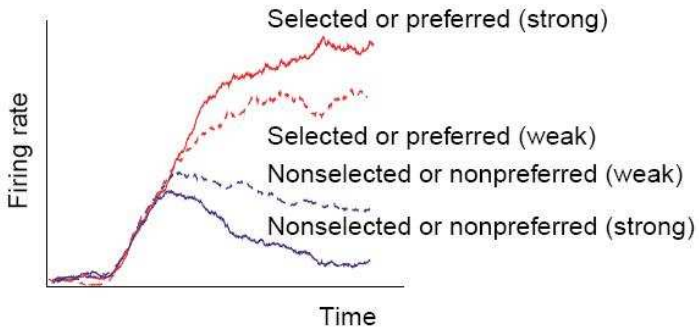


Figure 2.12: Neural activity associated with stimulus selection has been recorded in area LIP and in prelude or build-up neurons in areas SC. Early stimulus-linked activity does not discriminate between decision alternatives. Later, cells associated with the selected stimulus or the preferred direction of motion show an increased or maintained level of firing. Cells associated with the non selected stimulus or the non preferred direction show a decreased level of firing. The growth of discriminative information represented by the difference in firing rates occurs more rapidly for easily discriminated stimuli (strong) than for less easily discriminated stimuli (weak). Adapted from Smith and Ratcliff (2004)

2.3.2 Modelling decision making

Diffusion, accumulator models an neural integration

The study of the Reaction Time (RT) of animals and humans asked to perform simple two-choice perceptual discrimination task has a long history. As seen in Zohary et al. (1994), the stimuli are translated into evidence that is highly variable in time. Thus a decision cannot rely on this evidence at any instant. A decision mechanism has to find a way to average out the variance. One of the first approaches to modelling such a decision mechanism is the accumulation of evidence. According to Smith and Ratcliff (2004), two models accounting for this accumulation have originally been devised in the field of mathematical

psychology. The first one is the random walk model, as seen in Laberge (1962). In the random walk, a variable combines, over time, the difference between the quantity of evidence in favour of each of the two alternatives. When the variable reaches a positive (respectively negative) threshold the positive (respectively negative) decision is taken. This is the relative evidence method. Another way to model this accumulation is to use two variables, one for each alternative. The decision is taken when the first variable reaches a threshold. This is the absolute threshold method, and has been studied by Townsend and Ashby (1983).

In the framework of accumulation of evidence, the evolution of the decision variable or variables is modelled as the movement of a particle in a gas. The difference between two successive positions will have a Gaussian distribution around a mean value that is called the drift rate. This drift is null if there is no feature in the stimuli that can be used to argue in favour of one or the other alternative. In the case of the kinematogram, the drift will be null if there are no dots moving consistently in one direction. However, the drift rate of one decision variable will be positive if the stimulus produces evidence in favour of the corresponding alternative. In this model, the value at time T of the decision variable is its initial value plus the integration, from the time of the stimulus onset to time T, of the difference between two positions. Let us study a simplistic example, in order to understand how this accumulation of evidence method is in fact an integration process. Let ν be a decision variable, ν_0 its initial position, $d\nu(t) = \mu dt + \sigma dB_t$, μ the drift rate and σ the standard deviation of the Gaussian distribution. Let us assume that $\sigma = 0$. This is the unrealistic but simplistic case of perfect evidence extraction, with no perturbation. Each time step is dt, infinitesimally short. Then we have $\nu(T) = \nu_0 + \int_0^T \mu dt$ which means $\nu(T) = \nu_0 + \mu \cdot T$.

In Fig. 2.13, we can see an illustration of the diffusion model for the integration of evidence. It has been very successful at explaining the variations of RT as a function of the quantity of evidence, as shown for example in Ratcliff et al. (1999) and Ratcliff and Rouder (1998).

As explained is Smith and Ratcliff (2004), models that use two independent decision variables and take the decision according to which variable reaches the threshold first cannot explain some measured data, such as error responses being faster than correct responses. Usher and McClelland (2001) solved this problem by adding a leak in the accumulator, variability in the starting point and inhibition between each competing variable.

Competing attractors

The model of Usher and McClelland (2001) is not only a leaky accumulator. It is different from the diffusion and race to threshold models because it is a "Leaky *Competing* Attractor Model". They introduced

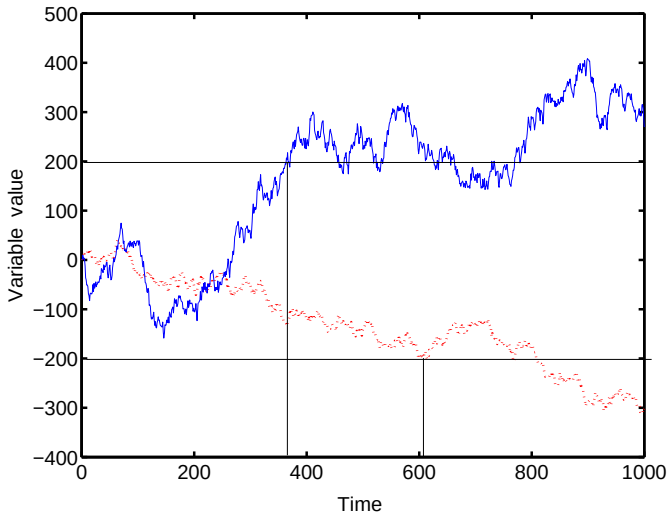


Figure 2.13: Illustration of the diffusion process, for two different inputs, one positive (blue continuous line) and one negative (red dotted line). Horizontal lines are decision barriers, vertical lines decision times.

an accumulating variable for each of the N possible decisions, with a leakage, and linked all these variables (typically Firing Rates of populations of neurons) with lateral inhibitions. In fact, we know that, in the visual cortex, eighty percent of inhibitory inputs to cortical cells are lateral or top-down inhibitions as argued in Usher and McClelland (2001). The authors of the leaky competing attractor model explain that lateral inhibitions implement a Winner-Takes-All strategy that is very useful in computational neuroscience. First, it gives a lot of flexibility regarding the number of choices available in one particular experiment. The previous models, that we described previously, used two alternative accumulating variables, and the decision was to be taken with either of this two methods: in the relative evidence strategy, a decision is taken when the difference of activity of the two accumulators reaches a threshold. In the absolute evidence method, the decision is taken when the first accumulator reaches a threshold of activity. Generalising the second method to an arbitrary number of alternatives is easy, but as argued in Usher and McClelland (2001) and Smith and Ratcliff (2004), models based the first method involve only one variable, the relative evidence, and produce results that accommodate better with the experimental data and have faster reaction times. In the leaky competing attractor model, there is one accumulator of evidence for each of the arbitrary numerous alternatives, and the first accumulator to reach a threshold will generate a decision in favour of its alternative, like in the absolute evidence method. However, the relative evidence in embedded in the model because each accumulator inhibits the others. Consequently, when one accumulator has a higher activity, it will reduce the activity of the other ones and eventually take over the activity, *"so that at the end the states*

of the accumulators will reflect the relative amount of information accumulated in comparison to the other accumulators." Furthermore, lateral inhibition is a biologically plausible hypothesis, as seen in Usher and McClelland (2001), p. 10, Fig. 2, and more generally in recordings of cells involved in selective activities. Their model also involves recurrent self excitation for each accumulator, in order to generate a system that settles in stable states. In this article, the accumulators are modelled as populations of neurons, and are simplified into units that have a threshold linear function. Since then more detailed models of the same idea of Winner-Take-All strategy for decision making have been presented that fit with neurological data and are based on Integrate and Fire neurons.

The first step was a mathematical analysis of large networks of Integrate and Fire neurons whose spike trains are Poisson processes, carried out by Amit and Brunel (1997). In their study, they proved that a realistic network of Integrate and Fire neurons can realistically achieve the dynamical behaviours necessary for the Winner-Takes-All strategy. They showed that such a network could converge to a stable low activity global attractor, and that, after adequate Hebbian learning, the network could exhibit local, transient high activity. These network features have been used to fit the leaky competing accumulators into a more general and more biological framework of non-linear dynamical systems, applied to our perceptive decision task: Wang (2002) implemented a probabilistic decision making system with slow reverberation that neuronally enables slow integration of evidence.

His model involves a recurrent network composed of competing units that inhibit each other in the same time as they integrate instantaneous evidence coming from the MT neurons. For the sake of simplicity, the model focuses on a two-alternative choice based on the kinematogram stimuli, so it involves only two competing units. Each of the units is a network of excitatory Integrate and Fire neurons. A third population of neurons is constituted of inhibitory interneurons. These interneurons receive excitatory inputs from each of the two competing subpopulations and send inhibitory activity to the same competing units. Therefore, we see that this implements the lateral inhibition introduced by Usher and McClelland (2001). Furthermore, in each of the units the neurons are strongly connected with recurrent excitation. This enables the convergence to a stable state of activity in which one unit has a high activity and maintains it even when the stimuli stop, whilst the other unit is inhibited to a near zero activity.

Wang's model is specifically interesting because he introduced two original features. First of all he modelled the synaptic inputs to the Integrate and Fire neurons in detail, which enabled him to model the effect of the time constants of the ionic currents on the dynamics of the network. In this thesis, we have mentioned ionic channels, but we quickly simplified them and we usually consider that one incoming spike on the synapse generates an instantaneous flow of electric charge in the postsynaptic neuron. In Wang's paper the ionic

channel dynamics have an important role. Upon arrival of a spike a quantity of these channels open, and then they close over time. The speed at which they close is called decay and is modelled by a time constant. Wang showed in his paper that if the time constant of the recurrent excitation is long enough (about 100 ms), then the network that models decision making integrates evidence over long times, whereas if it is too short it latches too quickly onto one attractor, not taking enough evidence into account. The second interesting feature is the background noise to which all the neurons of his network are submitted. It enables the network to take a decision even in the case of two balanced stimuli. Even if both the competing units are submitted to an important background activity that has the same statistical signature, the instantaneous realisation of this background activity varies from one unit to the other. That in turn induces the network to converge, at a random time, to one or the other attractor with 50% probability. However, when the stimuli are unbalanced the convergence is generated by the evidence expressed by the discrepancy of the two stimuli.

Other dynamical models such as the ones that we have just described (Wang (2002), Usher and McClelland (2001) or Mazurek et al. (2003)) can fit the biological data well, as seen in Heekeren et al. (2004), Sugrue et al. (2005), or Rorie and Newsome (2005), and introduce finer considerations such as prior probabilities. In Mazurek et al. (2003) there is an implementation of a model of the simple visual decision making chain that we are studying and that involves neural integrators for accumulating relative evidence in area LIP, following an idea that was already suggested by Shadlen and Newsome (1996). We can see an overview of their model in Fig. 2.14.

Since, furthermore, it has been shown that there are many inhibitory connections in such areas as LIP in the cortex, we see that using recurrent excitation and lateral inhibition is a plausible neural implementation of the accumulation of evidence necessary to perform decision making based on visual stimuli. We have also seen that the random background activity, or background noise, is more important than a mere perturbation. It has a role that influences the dynamics of a network that integrates evidence and competes for a decision. In Wang (2002), the characteristics of this noise are fixed, and a decision is systematically taken, even in the case of balanced stimuli. Therefore, in order to further the study of such relevant models we will investigate, further on in this thesis, the influence of the statistical signature of the background activity on the dynamics and on the actual performance of a decision modelled by such networks.

2.3.3 Higher level models of decision making

The neural models for decision making that we described above are appealing in terms of biological plausibility, but some issues about decision making remain, to which they do not provide answers. These models do not explain how we or the monkeys decide to take a perceptual decision about the moving dots

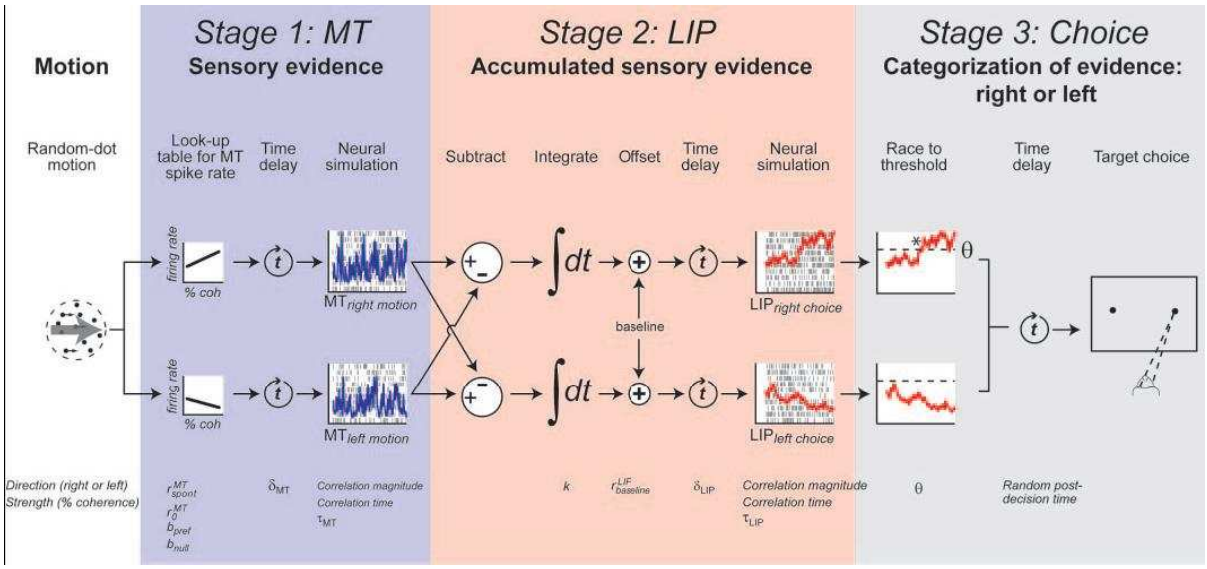


*Figure 2.14: Model of the decision process. The input to the model is a two-parameters description of the random dot motion stimulus shown on one experimental trial: direction and strength (coherence). The model predicts the behavioural choice (left or right) and the response time. It also represents neural activity which is intended to simulate the responses of direction selective neurons in area MT and decision related neurons in area LIP. The grouping into stages (coloured background) corresponds to sensory representation, accumulation of evidence for two competing hypotheses, and comparison of the accumulated evidence to a threshold. Labels below the elements show the variable used in the model. This specific example demonstrates a trial with strong rightward motion. Each neural simulation block shows simulated spikes from 10 neurons (tick marks) and the smoothed average spike rates from the full ensemble of 100 simulated neurons (solid curves). The race to threshold elements show the smoothed average rates from the two LIP ensembles racing against each other toward the decision threshold; * indicates the time of decision in the winning LIP. Adapted from Mazurek et al. (2003)*

or not to care about the moving dots at all. Furthermore, the role of urgency, and how we end up making quicker or slower decisions according to the context is not satisfyingly explained. The idea developed in these models is to lower the threshold above which the neural activity must grow for a decision to be taken, or to modify the network connectivity so that the dynamics of the attractor network are faster. This is not precise enough, because in the first case we do not know how the threshold is neurally implemented and in the second case, modifying the network connections is a long term process, more akin to learning than to adaptability to contingencies. In order to understand decision making in a way that can answer these questions, we need to take into account the whole subject taking a decision, and maybe abstract ourselves for a while from the neurological level.

Probabilistic decision making

From a formal point of view, and assuming a rational and well informed subject, one can model a decision as a probabilistic choice. As explained in Glimcher (2003b), this type of modelling is based on work formalised by Blaise Pascal, a mathematician of the enlightenment period. According to him, a decision made by a rational thinker in an uncertain environment will be based on the expected value of the choice, obtained by taking the reward expected from a choice, if correct, and multiplying it by the probability that the choice will be correct. Describing subjects as people who try to maximise the utility of their actions is a classic method used by economists and psychologists to try to understand people's behaviour in social environments.

Observing that one also tends to avoid risk, Bernoulli, another mathematician, refined the theory with the concept of utility. In his framework, instead of multiplying the probability of correct choice by the value of the reward, the subjects multiply it by a function of reward value. This function is smaller than the real value of rewards, and of positive decreasing derivative: the higher the reward, the larger the difference between the objective value of the reward and the subjective value. This theory of expected utility has been applied to economics for a long time, as we can see in Kreps (1990) and is starting to have neurobiological interpretations as seen in Glimcher (2003a). For example, he found out by experimenting on monkeys that *"the FR of LIP neurons were correlated with relative expected utility..."*. Correlating neural activities with variables traditionally used by economists is a new trend in the study of decision making, called neuroeconomics. In a recent review, Sugrue et al. (2005) introduce neural currencies and choice values to model decision making and choice behaviour. Linking analytical probabilistic choices to their neural basis is an interesting development in the theory of decision making. However it is beyond the scope of this thesis, in which we study the neural basis of a very simple perceptual decision making. As we will see in the next paragraph, we still have to answer questions about attention before we try to understand the neurology of probabilistic reasoning.

Somatic states and attention

In Damasio (1994), we find a global theory of reasoning that involves emotions and the body and that can answer the questions of when and how the subject chooses to act upon a stimulus (a kinematogram in our case) and takes a decision. His theory is about judgment, rationality and decision making in general, in social or emotional contexts. However, he explains that judgements and rationality can be considered as choosing between alternatives. He does not deny that decisions involve analytically comparing arguments in favour of a few competing alternatives, using probability measurements and maximising the expected utility of one's action as in the models mentioned in the previous paragraph, but he argues that this analytical step is a final step and that the alternatives have been pre-processed beforehand. He grounds his argumentation on clinical cases in which patients had a part of their cortex severely damaged. In the cases that he describes, the patients did not lose any faculties such as language, mathematics or even working knowledge. However they were incapable of both feeling emotions and of making appropriate judgements about their lives, their social surroundings or to make commitments.

In fact, at every moment in life, we are confronted with a myriad of choices between an almost infinite number of alternative courses of action. We can choose between these courses of action because we can imagine them and some of their consequences through "*Potential Representations*". Potential Representations are neural structures that are learned throughout life, through the repetition of actions, strategies, and their consequences. According to Damasio, the majority of these alternative potential representations are inhibited before they become activated, or even before they start being evaluated. What inhibit them are "*somatic markers*". Somatic markers translate body sensations associated with certain actions, experienced or imagined. If they are negative, the potential representation is not activated. This way, only a few alternatives are activated and compete for the decision, involving more precise evaluation of the evidence in favour of each alternative. More generally, he argues that the somatic state, which means the state of the body in relation with its environment will act as a global somatic marker that will control attention and working memory. He explains that through learned strategies for maintaining the homeostasis, the body state will translate into patterns of neural activity that, through increasing the activity in certain cortical areas, will prioritize the corresponding potential representations. Furthermore, he argues that in a similar way the somatic state enables working memory to be activated. In the introduction, in section 1.3.5, we have seen that perception is a motivated activity. We have used the theory of the phenomenology of perception to understand that we only perceive in light of our aims and what we called motives. In Damasio's theory, these motives are inscribed in the brain as somatic states: if we do not need to make a decision about the directions of kinematograms, i.e. if we do not have the motivation for it, than the potential representations that underly the evaluation of evidence in favour of one or another alternative will not get activated, and we

won't perceive the kinematogram's directions.

This theory on decision making not only fits the conceptual framework of this thesis, but is also coherent with the models that we described earlier on in this review. First of all, Damasio mentions Newsome's Winner-Takes-All strategy model in the case of kinematograms as an example of the biological relevance of his theory. Secondly, his theory gives a chance to answer the questions of attention and urgency that they could not address. Each unit in Wang's competitive network represents an alternative choice of interpretation of the direction of the kinematograms. In terms of Damasio's theory, it is a potential representation. The network could be activated, in which case a decision will be taken as described in the article Wang (2002), which means that a unit will reach a higher activity: the potential representation is activated. On the other hand if the network is not activated the two potential representations retain their low level undifferentiated activity. The decision will be taken only if the monkey pays attention to the kinematogram's direction. According to Damasio, this attention is controlled by the somatic state or motives in Merleau-Ponty's words. This somatic state could be translated into the background activity that all cells receive and was introduced but not used in Wang's model. So, we will see some studies about background activity and the effects of noise in dynamical system further on in this review, but first let us finish the study of the sensorimotor chain and describe how the monkey expresses a decision with an eye saccade.

2.4 Saccadic eye movements

Saccades are very fast eye movements, their angular speed reaching several hundreds of degrees per second. However, it has been shown that the time between a stimulus presentation and the performance of the corresponding saccade can reach several hundred milliseconds, because of the many neural computations and cognitive processes that must happen in order to prepare the saccade. Wurtz and Goldberg (1989) give a review of saccadic eye movements.

Using electro stimulations in the cortex of monkeys, Gold and Shadlen (2001) experimented on neurons in another area of the cortex, the Frontal Eye Field (FEF). This area has been shown to be involved in the motor control of the eye. They proved that the rate of increase of their FR is correlated to the rate of increase of the probability that one direction the answer of the subject performing the kinematogram direction decision task: experimenters could guess the direction of an eye saccade by measuring the FR of FEF neurons, which much higher accuracy than when measuring the FR of neurons of MT or LIP.

More precisely, two specific neural circuits (the pontine and mesencephalic burst circuits) for motor control give the signals to the eye muscles during a saccade. These neural circuits are controlled by a cortical

region called the Superior Colliculus (SC). The SC receives input from the early visual cortex and from extrastriate areas such as area MT and LIP. This is how the decision about making or not a saccade is conveyed. Furthermore, The Superior Colliculus receives a powerful inhibition which must be suppressed before it can drive a saccade.

The role of the SC in preparing saccades has been known for a long time, as shown for example by Sparks and Le (2004). It is explained in their paper that the neurons in SC are organised in movement fields. Movement fields are the motor system equivalent of the receptive fields of the sensory system. However, it has been shown recently that the activity in SC is not purely related to saccade preparation, as it can be correlated to a decision that is independent of the characteristics of the saccade, as seen in Horwitz et al. (2004). Classically, people model this control of SC upon saccade generation as an integrating device. In such models, similarly to the LIP integration occurring for decision making, SC gathers inputs from many areas including LIP until one movement field reaches a threshold. Then the powerful inhibition is released and the motor control neurons mentioned before generate the saccade.

One finds in the literature that, when modelling the reaction time, one adds a couple of dozens of milliseconds for stereotypical saccade realisation after the SC activity reaches the threshold. In fact in many decision making models, the authors add a fixed processing time to the reaction time, because they do not simulate the complete neural chain from early visual processing to motor control of saccade generation. In Mazurek et al. (2003), for example, the reaction time is the sum of the model's decision time and of 300 ms. These offset parameters are chosen to fit experimental data to the model results, but such an offset does not change the qualitative explanatory value of a model. In their own words: "The delay before visual information arrives in MT and LIP is based on estimates from neural recordings in these areas. These scaling factors and delays are mainly aesthetics. They allow us to compare real data to the response trajectories predicted by the model. They do not affect the model predictions in a substantial way."

2.5 Role of random neural activity

2.5.1 Synaptic noise and correlation

Correlation between neural activities is often considered detrimental to the signalling capacity of a pool of neurons because it reduces the signal to noise ratio of the output, as seen in for example Zohary et al. (1994). However Romo, in Romo et al. (2003) shows that under certain conditions of positive correlation, even if the noise increases, the discrimination accuracy of neurons remains stable, or improves. This happens when neurons have opposite reactions to the same change of stimulus. In the

case of their article, the firing rate of some neurons (minus neurons) decreases with the strength of the stimulus, while some other neurons' firing rate (plus neurons) increases with stimulus strength. These neurons *"with opposite slopes"* generate a more accurate pooled signal when their outputs are positively correlated then when they are uncorrelated because of the pooling method. In order to discriminate stimuli from the output of these neurons, one will measure the FR difference between the plus and minus neurons. Consequently we see that, because of the positive correlation the noise will be subtracted out by this operation whilst the signal variations will be enhanced thanks to the "*opposite slopes*" of the neurons.

In a more theoretical approach, recent research shows that negative correlation has a subtractive effect on the non-informative parts of a signal. This phenomenon has been neglected but is now being studied by Durrant and Feng (2006). In that paper it is shown that lateral inhibition within a column in which neurons are meant to have similar behaviour improves the accuracy of a population of neuron that performs a task based on its pooled activity. Cortical columns are an important structural unit in the brain, and operate on the principle of population coding, where the activity of the pool as a whole is taken to be the signal, rather than the firing rates of the individual neurons. The principle of this improvement is stated as follows: lateral inhibition between neurons will create negative correlation whereby if one neuron's firing rate increases, it reduces its neighbour's firing rate: one neuron's firing reduces the probability of other neuron's firing. Let us look at the simple case in which we model the activity of individual neurons as a firing rate to which is added normally distributed noise. We pool the activity of two neurons n_1 and n_2, whose firing rates are μ_1 and μ_2 and whose added noise is normally distributed with standard deviations σ_1 and σ_2, which are correlated by a negative correlation coefficient c. Then the pooled activity of these two neurons will have, as seen in 2.2.5, a firing rate of $\mu = \mu_1 + \mu_2$ and an added normally distributed noise of standard deviation $\sigma = \sqrt{\sigma_1^2 + \sigma_2^2 + 2c\sigma_1\sigma_2}$. Since $c < 0$ then the signal to noise ratio (μ/σ) is larger than if the signals were uncorrelated ($c = 0$). The noise gets subtracted out by the correlation. In Durrant and Feng (2006), we find a model that shows that the performance of a network Integrate and Fire neurons on a dot tracking task is improved by lateral inhibitory connections that induce negative correlation.

2.5.2 Stochastic low level background activity of the brain

Decision making is also controlled by our context, our motivations, and our history. This idea has been addressed by Salinas (2003): he uses "background synaptic activity as a switch between dynamical states in a network". The decision making is dependent on the characteristics of the neuronal activity of the brain.

Synaptic background activity has been observed to significantly influence neural outputs by numerous authors such as Shadlen and Newsome (1994), by Wolfart et al. (2005), by Chance et al. (2002) and Softky and Koch (1993). It has been used to model cognitive control on decision making by Salinas (2003), and

it has been modelled by Amit and Brunel (1997) who managed to model a population of spiking neurons in stable spontaneous activity of approximately 5 to 10 Hz, as observed in the brain. Furthermore, in Wang (2002), all the neurons involved in the competition leading to a decision receive a large amount of Poisson inputs, fire spontaneously at a few Hertz, and Wang argues that this is the dominant source of variability.

2.5.3 Noise in dynamical Systems

Pitchfork bifurcation

In Berglund and Gentz (2002), there is a very interesting study of the influence of noise on the dynamical system that presents a bifurcation upon the slow change of one parameter. Let us describe the dynamical system:

$$\frac{dx}{dt} = \lambda x(t) - (x(t))^3 \tag{2.5.1}$$

The nullclines are defined by the equation:

$$\lambda x = x^3$$

$x = 0$ is one solution, and if λ is positive, $x = \sqrt{\lambda}$ or $x = -\sqrt{\lambda}$ are other solutions. When λ is negative, the equilibrium at $x = 0$ is stable: if $x > 0$, then $\frac{dx}{dt} < 0$ thus x tend to decrease and get closer to $x = 0$, whereas if $x < 0$, then $\frac{dx}{dt} > 0$ thus x tend to increase and get closer to $x = 0$. On the contrary, for λ positive, the equilibrium at $x = 0$ is unstable: if $x > 0$, x close to zero, then $\frac{dx}{dt} > 0$ and vice versa. However, the equilibriums at $x = \sqrt{\lambda}$ or $x = -\sqrt{\lambda}$ are stable.

In order to create a pitchfork bifurcation situation, we make the parameter λ slowly vary with time, from negative values to positive values. In the absence of noise, the trajectory of x is predictable: from any initial value, as long as λ is negative, it will tend to $x = 0$ but not reach it, thus x will not change sign. Then, when λ becomes positive, the $x = 0$ equilibrium becomes unstable, so x will diverge to the stable equilibrium of its sign. The time necessary for this divergence is called the bifurcation delay. It is defined as the first exit time from a strip around the solution $x = 0$. The closer x initially was to zero, the longer the bifurcation delay is. This deterministic bifurcation set up is described in Fig. 2.15.

The interest of the work of Gentz resides in the study of this system under the presence of noise. The equation Eq. 2.5.1 becomes:

$$dx = \frac{1}{\varepsilon}\lambda x(t)dt - (x(t))^3 dt + \frac{\sigma}{\sqrt{\varepsilon}} dW_t \tag{2.5.2}$$

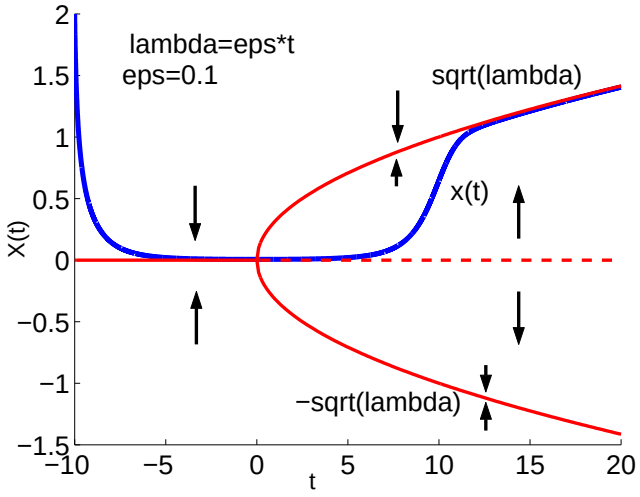


Figure 2.15: Illustration of the deterministic pitchfork bifurcation. $\lambda = \varepsilon \cdot t$. *Simulation realised in Matlab, with the following parameters:* $\varepsilon = 0.1$; *Time step for integration* $dt = 0.003$. *The red plain lines are the stable attractors, and the red dashed line is the unstable attractor. The arrows show the local sign of the derivative of* x, *thus they show the stability of the local equilibrium trajectory.*

where dW_t is a Wiener process. She has obtained qualitative results about the influence that ε and σ have on the dynamics of the system. First of all, the particular solution is very unlikely to leave a neighbourhood of the unstable zero solution before $t = \sqrt{\varepsilon}$, if σ is small compared to ε. Secondly, toward the end of the convergence to an attractor, the convergence is exponential.

The bifurcation delay is random, but is very likely to lie in the interval $\left[\sqrt{\varepsilon}, C\sqrt{\varepsilon log(\sigma)}\right]$.

We have three parameter regimes. If $\sigma \geq \sqrt{\varepsilon}$, the definition of the bifurcation delays loses sense, and the results obtained in their paper do not hold any more. This is because the noise has a preponderant importance around the bifurcation and the path no longer depends on the initial condition. We need $\sigma << \sqrt{\varepsilon}$. If $e^{-1/(\varepsilon p)} \leq \sigma << \sqrt{\varepsilon}$, for some $p < 1$, the bifurcation delay is called microscopic because $C\sqrt{\varepsilon log(\sigma)}$ is very small. If $\sigma \leq e^{-1/(p\varepsilon)}$ for some $p > 0$, then the bifurcation delay is macroscopic, because the upper bound of the interval is of order one.

Stochastic resonance

One of the classic uses of noise in dynamical systems is the phenomenon of stochastic resonance. Stochastic resonance happens when a specific, non-minimal value of input noise optimises the signal to noise ratio of

the output of the non-linear system. This phenomenon was observed a long time ago, and used a lot in many areas of science and engineering. In neuroscience, it has been pointed out by Benzi et al. (1981) and used for interpreting neural results in Usher and Feingold (2000).

2.6 Decision making: from bio-chemistry to embodiment

So, in this chapter we have seen that neurons are the basic units of the brain and that their organised communications are constituted of spikes propagating along axons from one neuron to another. The strength of the neural network that constitutes the brain and even more specifically the cortex lies in the sheer number of neurons and connections between them and on the fact that neural activity is highly distributed. The neural code is so complex and depends on so many detailed phenomena such as the dynamics of ionic channels and gating variables that we need to use statistical tools to capture its richness. This way we can sometimes simplify neural activity into smooth continuous input out mapping of neural activity, using measures such as the FR. This analog, or continuous approach is adapted to the modelling of sensations and motor control. However, in order to reason, we need some digital structures, whereby some groups of neuron can maintain discrete patterns of activity that correspond to the "ones" or "zeros" states of digital, computer like organisation. In the decision making process that we are investigating, experimental measurement have shown that, while the activity of neurons in area MT fluctuates analogically with the stimulus, the activity of neurons in LIP ramps up to a higher level that is maintained stable in time in order to express a decision. This pattern of LIP activity is an example of the digital code that can appear in some parts of the brain and has successfully been modelled by networks involving recurrent excitation and lateral inhibition.

Nevertheless, the picture is not as simple as it seems. In order to have a satisfying decision making modelling, we need to understand why and how the dynamics change, why and when we pay attention and take a decision at all, how we can perform more analytical probabilistic evaluations of the expected rewards, and to acknowledge that LIP is only one of many structures of the brain in which researchers have found activity correlated with decision making. Therefore we need to see the neural processes under our scrutiny as embedded in a global system, involving both the brain and the body. After studying various models of decision making and one theory of judgement and rationality (Damasio (1994)), we have formulated the hypothesis that what some call neural background activity could be the neural expression of somatic states or motives, of the particulars of the body and education of each individual and thereby control the dynamics of judgements, in our case the dynamics of perceptual decision making.

Therefore the interesting research that remains to be done is to investigate the interactions between this digital processing realised in our case by competitive units converging to one or the other alternative and

the analog, statistical neural processing arising from our somatic states, our motives. Our hypothesis is that the synaptic noise and what is traditionally considered to be noisy low level background neural activity are actually expressions of the somatic states. Consequently, in what follows, we will investigate the influence of the statistical signature of various types of neural noise on the decision making process modelled as a Winner-Takes-All competition between attractors. These models will focus on the kinematogram based perceptual decision task. To begin with, we will investigate in the next chapter the influence of synaptic noise on the production of evidence by the visual cortex.

Chapter 3

Input noise improves discrimination accuracy

We have seen that in the course of decision making the stimulus is analysed by the early visual cortex, evidence for the decision making is extracted by MT neurons and encoded in their output Firing Rate. In order to have a complete model of visual perception, we would have to play out the kinematograms on a screen and model a retina with light receptors, or maybe more simply with a camera. In this work, we do not go that far in detail. Our electronic simulation of a kinematogram is composed of a hundred dots. A dot is represented by two coordinates in a plan. Thus the kinematogram is made of two vectors of a hundred elements: the abscissas and ordinates of the hundred dots.

In this chapter, we present an attempt at modelling how the early visual system interprets this model of a visual flow into movement information. From this modelling attempt we decide which assumptions to use for the input to the second stage, the discrimination performed by MT neurons. We then present a model of this discrimination task, constituted of leaky Integrate and Fire neurons. We analyse the performance of this model under the influence of input noise.

3.1 Reading out the Stimulus

The first step was to model the encoding of the retinal input into neuronal data that carries the information about the directions of the dots. The difficulty was that there were a hundred dots moving simultaneously: if we model vision as a discrete process (this idea his supported by VanRullen and Koch (2003)), then we have to analyse a succession of images of a hundred different dots, modelled as a succession of vectors containing a hundred coordinates each. The dots are not different from one another, so we cannot

know which coordinates of two successive images are two successive coordinates of the same dot. The task of associating a dot to its origin between two successive snapshots is called the correspondence problem.

In order to model the perception of a continuous movement, in Gaillard (2002), we took the successive images along very small time steps, and associated them with successive neuronal grids. Consequently, the position of a moving dot within one time step hardly changed. That allowed us to solve the correspondence problem by using what we called 'The nearest point assumption'. When we had two successive images of a hundred dots that have moved, the assumption was that the origin of one dot P was the dot in the previous image that was the closest to P.

3.1.1 Receptive fields of successive grids

In Gaillard (2002), we modelled the solution of the correspondence problem with the use of receptive fields in the early visual system, as illustrated in Fig. 3.1. In that work, the retina is considered as a two dimensional input space. Each dot of the kinematogram is thus modelled as a ray of light that hits the retina at one point, defined by its two coordinates : its abscissa and its ordinate, ranging from 0 to $1cm$, in a square.

The retina is associated with several grids of neurons, in a two-dimensional topographic organisation. There are n^2 neurons in each grid, referred as neuron (i, j) which means neuron of the i^{th} line and the j^{th} column. Each neuron corresponds to one area of the retina, by the means of a receptive field. For neuron (i, j), the receptive field is centred on the point of the retina whose coordinates are :

$$abscissa = i \times res$$

$$ordinate = j \times res,$$

res being equal to $\frac{1}{n}$. This means that the distance between two receptive centres is res and that they cover the whole retina : no point in the square [0,1]x[0,1] cm of reception of the retina is further to the centre of a receptive than $\frac{res}{2}$. The first grid receives inputs from the retina, and sends outputs to the second grid and to area V5 or MT, the part of the visual cortex that computes motion. The second grid receives inputs from the first grid, and outputs to V5. Each grid expresses the positions of the dots at a precise time : the first grid at time t, the second grid at time $t - dt$, where dt is the time lapse between to snapshots of the kinematogram. That is why we call the first grid, grid0, the second grid, grid1.

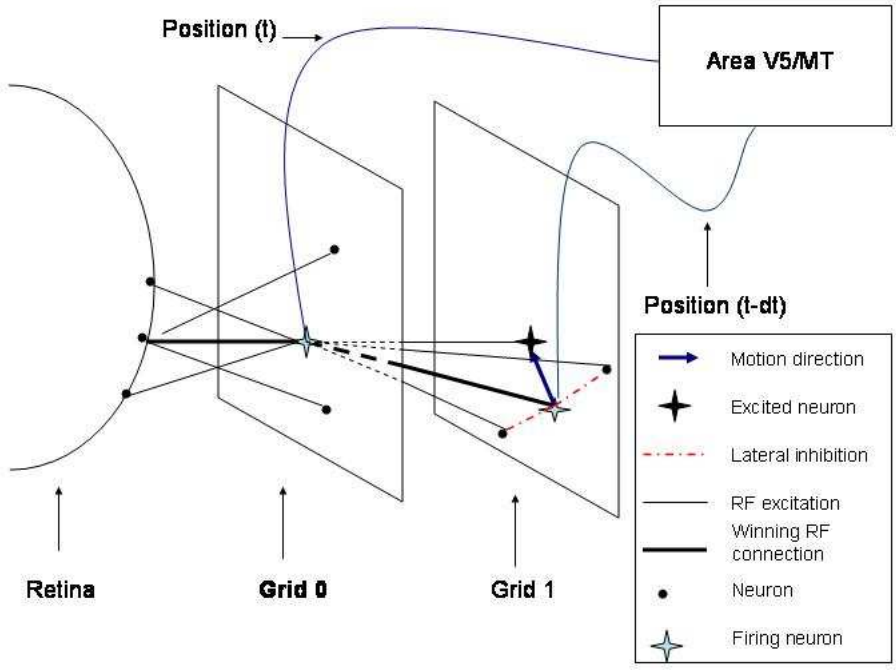


Figure 3.1: Illustration of the grid model for computing the directions of the dots. There are no lateral connections in grid 0, but we can see in red the lateral inhibition that induces Winner Takes All type of competition in grid 1. The other connections, between the retina and grid 0 and between grid 0 and grid 1 are synaptic connections whose weights correspond to a Gaussian shaped receptive field. The blue stars represent neurons firing at time t, expressing the positions of the dot at time t in grid 0 and $t - dt$ in grid 1. The black star represents the grid 1 neuron excited by the firing of grid 0 at time t and that will fire at time $t + dt$. Receiving simultaneously the position of the dot at time t and $t - dt$, V5/MT can compute motion information.

Receptive fields of grid0

For each new time step and for each dot, each neuron (i, j) of the first grid (grid0) gets an instantaneous activation input corresponding to its Gaussian shaped receptive field :

$$V_{i,j}(t) = V_{i,j}(t - dt) + \sum_{k=1}^{N_{dots}} G_{i,j}(x_k, y_k) \tag{3.1.1}$$

where $V_{i,j}(t)$ is the membrane potential of neuron (i, j) at time t, and

$$G_{i,j}(x_k, y_k) = e^{-\frac{\|centre(i,j)-(x_k,y_k)\|}{\mu}} \tag{3.1.2}$$

Where i and j are the number of the line and of the column of the particular neuron under consideration. This neuron has a receptive field centred at $abscissa = i \times res$ and $ordinate = j \times res$, $0 < i, j \leq n$, res is the distance between the receptive fields of two neurons along one coordinate, n^2 is the number of neurons, x_k and y_k are the coordinates of the k^{th} dot at time t, μ is a parameter that controls the sharpness of th receptive field.

Only the neuron that receives the most activation input fires. The model requires that the neurons need to be previously excited in order to be able to fire accordingly to an input. So here, we assume that all the concerned neurons of this grid (grid0) are excited. In short, we mean by 'excited' the state of an IF neuron whose membrane potential is close to firing threshold. In this set-up, the neurons of grid0 whose centres are the nearest to the input dots fire : Their positions are the output that goes to grid1 and V5. This assumes that the receptive fields of the neurons of grid 0 are sharp enough, so that only the neurons that is the closest to the dots receives enough inputs to fire, whereas its neighbouring neurons receive negligible inputs from the same dots. In the programming of this model, we have artificially evaluated the highest excitation of the local neurons and assumed that the other neurons receive zero excitation from the dot. There is no lateral connection inside grid0. In order to be more biologically realistic, we could model the same system with local lateral inhibition, so that the neuron that fires would in the same time inhibit its neighbours. This inhibition would compensate the slight increase in membrane potential resulting from the smaller inputs that neurons whose receptive are centred close to the dots may receive.

Winner-Takes-All strategy with spike timing in grid 1

In a term paper (Gaillard (2003)), we described an attempt at storing the positions of the dots in grid1, and retrieving the origins (positions at time $(t - dt)$) of the dots presented to grid0 at time t, so that, by taking the difference between two successive positions of the same dot, we could compute the direction of the dot's movement. For that, we used a Winner-Takes-All strategy in grid 1, via lateral inhibitory

connections between the Integrate and Fire neurons of the grid. Taking into account the requirement for rapid processing, we associated a spike timing code to the receptive fields structure. We started with a perfect Integrate and Fire model and we highlighted the interest of adding a leak to it. The main idea behind the model was to tune the receptive fields of the neurons so that the position of a dot was remembered as a high membrane potential, close to firing threshold, in the neural grid corresponding to time $t - dt$. Upon excitation from the grid implementing the positions of the dots at time t, the closest remembered position would cross threshold, fire, and inhibit the surrounding neurons, hereby implementing nearest point assumption by using the Winner-Takes-All strategy. We see from this description that the model fundamentally relied on the precise control of the membrane potential of specific neurons. The nearest point assumption is explained and evaluated in the next section.

The connections between grid 0 and grid 1 are feedforward connections, and take the form of Gaussian receptive fields. Neurons of grid 1 receive synaptic connections whose strength decrease with the distance like in grid 0, according to the equation 3.1.2, except that x_k and y_k are not the coordinates of a dot, but the coordinates of neuron k in grid 0.

The receptive fields between grid0 and grid1 are realised by very sharp Gaussian functions (with a small μ coefficient), so that only the neurons corresponding to the positions of the dots will get an excitation that raises their membrane potential very near the threshold potential. The others will receive a much smaller, decreasing with the distance, excitation so that they can fire if they were previously excited, but are negligibly excited if they previously had a zero potential.

Furthermore, there are lateral inhibitory connections in grid 1, in order to ensure that only one neuron per dot fires at each time step. The only neurons that can fire in grid 1 are neurons that are excited, which means that they encode the positions of various dots at time $t - dt$. The competition is twofold: the nearest excited dot will receive a stronger input because of the shape of the receptive field. Furthermore, even in the case of another competing neuron receiving enough excitation to also fire, the nearest one will fire first, because it will receive its spike before its opponent, because the distance from the firing neuron in grid 0 is shorter, thus the connection delay is smaller. The firing, winning neuron will then send inhibitory spikes that will inhibit the opponents before they can fire. The strength of the inhibition also decreases with the distance.

The lateral inhibition plays a second important role, besides the competition: It reduces the membrane potential of the excited neurons, who store the dots positions at time $t - dt$ but that do not fire at time t. This way they are not in the excited state anymore at time $t + dt$ which would disrupt the motion computation. A good memory system needs to also forget. In addition, we used the leak of the Integrate and Fire model

to un-excite the neurons with time. As a development of this model, we could use a third grid that would store the positions of the dots at time $t - 2dt$. this could be useful for motion predictions.

Nearest point assumption

The second step of the process is to estimate the position the dot had at time $t - dt$. We assume that the continuous time that characterises the natural world is discretised into very small time steps dt, in comparison to the speed of the dots. The only knowledge we have of the previous positions of the dots in the image is the state of grid1. This is a distributed memory : the 'excited' neurons, whose membrane potentials are close to the firing threshold, correspond to the locations of the dots in the previous image. (We have seen that in the primary visual cortex the topography of the visual scene is preserved). The membrane potential of these neurons got increased to a near firing level, because the centre of their receptive field was the neuron of grid0 who fired at time $t - dt$ due to the presence of the dot: The neurons of grid 1 store in their membrane potentials, as a distributed memory, the positions of the dots. So, the task is now to make fire the excited neuron of grid 1 that corresponds to the input from grid 0.

Here we use the fact that dt is small. We have made the assumption that the kinematograms are represented into area V1 by a succession of images of the dots' positions. The lapsed time between two successive images is dt. We assume that dt is very small, so that the dots hardly have moved between two images. This enables us to use the nearest point assumption in order to solve the correspondence problem. The nearest point assumption is expressed as follows: Let us have two successive images of the positions of the N dots in the kinematogram. Since the dots are non distinguishable, the task is, for each dot D of the image of time t, to find which dot in the image of $t - dt$ is the origin of D. To solve this task, we assume that the origin of D, in the image of $t - dt$, is the dot whose position is the closest to the position of D in the image of t. Let us quantify the scope of validity of this assumption, as was done in Gaillard (2002).

Let N be the number of dots in the kinematogram, V be the velocity of a dot, and thus $dL = V \cdot dt$ the distance covered by a dot during a timestep dt.
As seen in section 3, the dots move in a square, and their positions are described by their two-dimensional coordinates that vary from 0 to 1 cm.
For any dot D in the square, we want dL to be less than the distance from D to any other dot in the square. This is obviously not always the case because two randomly chosen positions in a square can be arbitrarily close to each other. Thus, the nearest point assumption is not perfect. However, we only want the nearest point assumption to be reasonable, whilst the remaining errors can be accounted for by the input noise to the modelled neurons, further down in the model, for example in section 3.2. A reasonable criterion is to have dL smaller than the expectation of the distance between two randomly chosen dots.

In order to calculate this expectation, we first calculate the probability $P(r)$ for one dot to be in the centre of a circle of radius r that does not contain any other point. This probability is the product of the probabilities for every dot not to be inside the circle, because the dots' positions are independent from each other. Since the dots' positions are equiprobable on the whole square, we subtract the area of the disk from the area of the square and we reach the following expression:

$$P(r) = (1 - 2\pi r)^N \tag{3.1.3}$$

Let R be the radius of the largest empty circle centred on D. R is a random variable whose probability function is:

$$F_R(r) = 1 - (1 - \pi r^2)^N \tag{3.1.4}$$

thus, its density of probability is:

$$f_R(r) = 2\pi r N(1 - \pi r^2)^{N-1} \tag{3.1.5}$$

In that case, the expectation of R is:

$$E(R) = \int_0^\infty 2\pi r^2 N(1 - \pi r^2)^{N-1} dr \tag{3.1.6}$$

In, fact we limit the possible radius of the circle to $R = 0.5$, because otherwise the circle is wider that the square in which the dots move. This limitation is reasonable because it is highly unlikely to have no dots in a circle of radius larger then $r = 0.5$ around D in the case of our kinematograms.

$$E(R) = \int_0^{0.5} 2\pi r^2 N(1 - \pi r^2)^{N-1} dr \tag{3.1.7}$$

$E(R)$ is the expectation of the radius of the largest empty circle around D. It is equivalent to the expectation of the distance between two randomly chosen dots. If the distance between two successive positions of a dot is shorter than a fraction of $E(R)$ than the nearest point assumption holds. For $N = 100$, $E(R) = 0.049$, and for $N = 50$, $E(R) = 0.07$. So, if we take the reasonable modelling values $dt = 0.005s$ and $V = 0.01m.s^{-1}$, then $dL = Vdt = 0.005cm$, which is a tenth of $E(R)$. Thus, with an update rule of 5 ms, even if the dots cross the screen very fast (one second), we are largely within the scope of validity of the nearest point assumption.

Computing the dot's directions

Still only focusing on one dot's trajectory, the V5 area receives as inputs two consecutive positions of a dot : It only needs to compute the difference between these two positions to get the direction of the dot, encoded as a two dimensional vector. From the vector that represents a direction, we want to get a number, the rate of a Poisson process to be used in a discrimination task with an Integrate and Fire model of a neuron, as done in Deng et al. (2003).

If we do not focus on the norm of the vector (the speed of the dot), we can represent its movement by the angle argument of the vector. This is the angle between the vector (1,0), and the vector direction of the dot. For instance, if the vector direction of the dot is (x,y), and x and y are strictly positive, this angle is $arctan(\frac{y}{x})$. In Gaillard (2002), we assumed that V5 computes it perfectly, and used an artificially coded programme for it.

Discussion

This model was only partially successful: it could track the movement of one dot, but made too many errors when many dots were shown simultaneously in the kinematogram. In the own words of our conclusion:

We tried to build a biologically possible model of the primary visual cortex that would, out of two successive images of the positions of a hundred dots, output their hundred directions. The idea we developed seemed promising, but we have only been able to perform one part of the task: it works when only one dot moves. With several dots that move at the same time, we have more difficulties. Each new dot that appears in the same time step will interact with the other dots and modify the membrane potentials of all the grid, either by increasing the potentials all around it, or by firing the neuron of itss previous position, which will decrease the membrane potentials around.

So this model could only partially perform the task at hand. This was nevertheless an interesting work, because it proposed a method for solving the correspondence problem with a distributed memory system based on a network of biologically inspired neurons. Except for some computing shortcuts, all the processing described by the grids only relied on Integrate and Fire neurons. This research highlighted the difficulties that one faces when trying to finely tune a distributed system of many neurons in order to apply it to a task that is typical of real world situations. We have seen that the limitations of this model were due to the fact that we did not manage to control precisely enough the dynamics of the membrane potentials of the many neurons when submitted to the influence of many dots. The system was too sensitive to external influences such as other dots. This was without adding errors, noise or perturbations typical of real life

tasks. This might be due to a non-optimal model, or to non-optimal parameters. However, in my opinion, this model suggests that it would be very difficult to find parameters for a robust system that would perform such a complex task and whose output would depend on the production of one spike for one moving dot, requiring the membrane potential of one neuron to have a very precise value at a precise time. This type of encoding seems very fickle. So, analysing the successes and limitations of the model made us think that in order to model the neural basis of real life perceptive tasks in a robust and reliable way, it would be very beneficial to consider noise and errors as an important part of the system, to incorporate neural variability and unpredictability of a complex stimulus and to distribute the information across numerous neurons.

Therefore the subsequent work bypasses the modelling of early visual cortex functions and makes assumption directly at the level of the inputs to area MT. These assumptions, nevertheless, are informed by this modelling attempt in so far as they incorporate as an essential feature the stochasticity of neural communication.

3.1.2 Assumptions in use at the next level

The output of the early visual system is encoded as a vector containing the evaluated hundred directions corresponding to each moving dot. In neural terms, each element of the vector contains a Firing Rate (FR) that is proportional to the angular direction of one of the 100 moving dots on the screen, and normalized between 0 and 100 Hz. For example, if one dot moves rightwards, the corresponding FR will be 0, if it moves downwards, the FR will be 25 Hz, leftwards, 50 Hz and if the dot moves upwards, the FR to the corresponding synapse will be 75 Hz. So, for n_c dots that move coherently n_c values are equal, namely $FR = 25$ Hz for a downwards movement, and $FR = 75$ Hz for an upwards movement.

This suggestion of coding the direction of a dot with a FR in a linear way draws on well established experimental data that shows that there are neurons in the higher levels of the visual cortex that have directionally sensitive receptive fields. This is supported by for example Reid et al. (1987), in another benchmark perceptive experiment during which the activity of neurons in the cat's visual cortex were recorded during gratings displays. Furthermore, in for example Wang (2002) the response of MT neurons is modelled as a FR that is linearly dependent on the direction of the stimulus. In Heeger et al. (2000), it is shown that some neurons in area MT of monkeys and in its homologous area MT+ in humans show an increased FR when the subjects are presented with a kinematogram in their preferred direction, show a reduced FR in the opposite direction, and little change in their FR in an orthogonal direction. All this evidence convinced us that encoding the direction of dots with FRs was reasonable, even if our encoding here is different in one aspect from the encoding in the models that we have just mentioned. In these, the authors encode the kinematogram's direction with neurons' FRs that decrease linearly with the angular

difference between the stimulus and the direction they are tuned for. Thus, in order to discriminate, they use at least two populations of neurons tuned to two different directions. We only use one population of neurons whose FR increases linearly with the direction of the stimulus. We will discuss this difference in subsequent chapters.

The kinematogram stimuli are now coded into vectors of motion directions. This vector of directions is now the input of neurons that discriminate the overall direction, as illustrated in Fig. 3.2.

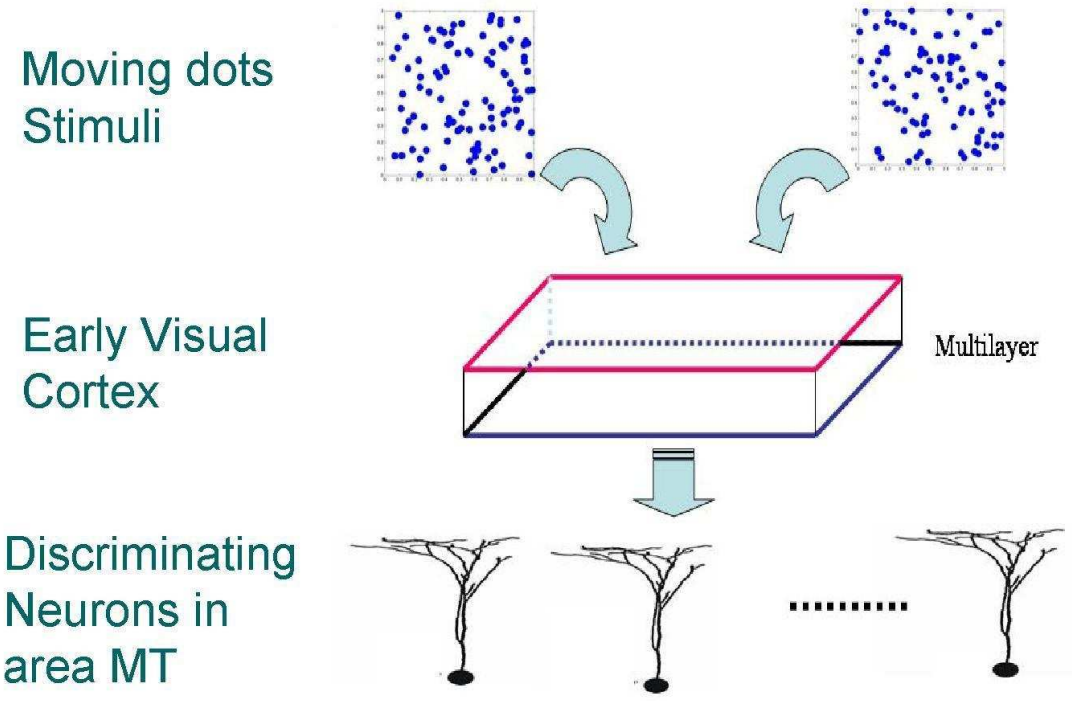


Figure 3.2: A schematic illustration of the model we use in the rest of this chapter. The 'multilayer' part represents the early visual cortex, from the retina to MT. It is treated as a 'black box' that gives a pre-processed input to the represented discriminating neurons of area MT.

3.2 Model of MT neurons

In the previous section, we modelled the encoding of the stimulus into a vector of directions of moving dots. This vector is the input of the next stage of the decision making chain. As seen in the first chapter, there is strong evidence that area MT contains discriminating neurons (Britten et al. (1993)). It has been shown that groups of MT neurons show an activity that is higher when the direction of the kinematogram is closer to their receptive direction, and when the coherence level is higher. By reading the activity of such neurons, an observer can guess the direction of the kinematograms: this part of the chain is the evidence producing stage. The rest of this chapter describes a model for it. It reformulates, extends and gives more precisions to the papers that we already published in Gaillard et al. (2006b) and Gaillard and Feng (2005).

3.2.1 LIF model and approximations

The discriminating neuron model used here is the Leaky Integrate and Fire (LIF) model, as used in Gerstner and Kistler (2002), Tuckwell (1988) and Feng (2001). The dynamics of the membrane potential V below threshold are defined as follows:

$$dV = -\frac{V - V_{rest}}{\gamma} dt + I_{syn}(t)$$

where γ is the decay coefficient. The synaptic input is described as a linear summation of incoming Poisson processes, as in Eq. 2.1.11.

$$I_{syn}(t) = \sum_{i=1}^{p} a_i E_i(t) - \sum_{j=1}^{q} b_j I_j(t) \tag{3.2.1}$$

$E_i(t)$ and $I_j(t)$ are Poisson processes with rates $\lambda_{i,E}$ and $\lambda_{i,I}$, p and q the numbers of excitatory and inhibitory synapses, and a_i and b_j are the positive magnitudes of the Excitatory Post-Synaptic Potential (EPSP) and Inhibitory Post-Synaptic Potential (IPSP) at the synapses i and j. For the sake of simplicity we considered the synaptic input to be pooled from a large number of synapses. Of course, one neuron in area MT has many more synapses than 100, more than one synapse receive information concerning one moving dot, and synapses have various weights (generally, $a_i \neq a_k \neq b_j \neq b_l$ for different i, j, k, l) and furthermore, not every synapse is activated all the time. Our aim was not to exactly describe the synaptic connectivity of MT neurons, but to evaluate their discrimination capacity, so we grouped similar incoming inputs to a number of synapses of various weights and summarised them into one incoming Poisson rate and one synaptic weight. Therefore, we modelled each neuron to have 100 excitatory synapses: $p = 100$, because it corresponds to the number of moving dots, and we consider that all excitatory (respectively inhibitory) synaptic weights are the same, a (respectively b). The input Poisson processes are the outputs of

the modelling of previous section. We suppose that, for each direction, the neuron receives both excitatory and inhibitory inputs of the same intensity: for all i, $\lambda_{i,E} = \lambda_{i,I}$, and that they are not correlated.

The main parameter that we study is the ratio between inhibitory and excitatory synaptic inputs, we call this ratio r: $r = \frac{\lambda_{i,I}}{\lambda_{i,E}}$, from Eq. 3.2.1, and we pose $p = q$. We can then rewrite Eq.3.2.1 as:

$$I_{syn}(t) = a(\sum_{i=1}^{p} E_i(t) - \sum_{j=1}^{p} I_j(t)) \tag{3.2.2}$$

with $\lambda_{i,I} = r\lambda_{i,E}$. Out of the $p = 100$ synaptic inputs that correspond to 100 moving dots, n_c synapses correspond to the coherently moving dots. The inputs from these n_c synapses are correlated with a correlation coefficient c, whereas the rest of the inputs are not correlated. This correlation comes from the fact that neurons whose receptive fields are centred on a particular direction will be close to each other and tend to fire together, in a correlated manner, as justified in Feng (2001) and Zohary et al. (1994).

We then use the Usual Approximation Scheme, as presented in section 2.2.1, in the equations 2.2.14 and 2.2.15 and we obtain the following expression for the synaptic input from one excitatory synapse:

$$E_{i,k} \approx \mu_{i,E}dt + \sigma_{i,E}dB(t) \tag{3.2.3}$$

where $\mu_{i,E} = \frac{1}{<ISI_{i,E}>} = \lambda_{i,E}$ and $\sigma^2_{i,E} = \frac{Var(ISI_{i,E})}{<ISI_{i,E}>^3} = \frac{1}{<ISI_{i,E}>} = \lambda_{i,E}$, because of Eq.2.2.13 and $Var(ISI) =< ISI >$, since the incoming inputs are Poisson processes.

We can therefore rewrite Eq. 3.2.2 as follows:

$$I_{syn}(t) = a(\sum_{i=1}^{p}(\lambda_{i,E}dt + \sqrt{\lambda_{i,E}}dB_i(t)) - \sum_{j=1}^{p}(\lambda_{j,E}dt + \sqrt{r\lambda_{j,E}}dB_j(t)(t))) \tag{3.2.4}$$

As seen in Eq. 2.2.17, we combine such continuous processes according to the following rule: the sum of two normal distributions of means λ_1 and λ_2 and standard deviation σ_1 and σ_2, correlated by a correlation coefficient c is a normal distribution of mean $\lambda = \lambda_1 + \lambda_2$ and variance $\sigma^2 = \sigma_1^2 + \sigma_1^2 + c\sigma_1\sigma_2$.

In the case of our model, we assumed that the inhibitory and excitatory inputs are not correlated ($c = 0$), thus we can rewrite Eq.3.2.4 as:

$$I_{syn}(t) = a(\sum_{i=1}^{p}((1-r)\lambda_{i,E}dt + \sqrt{(1+r)\lambda_{i,E}}dB_i(t)) \tag{3.2.5}$$

and we then combine the sum of these processes into:

$$I_{syn}(t) = a(\bar{\mu}dt + \bar{\sigma}dB(t)) \tag{3.2.6}$$

where

$$\bar{\mu} = \sum_{j=1}^{p}(1-r)\lambda_j \tag{3.2.7}$$

$$\bar{\sigma}^2 = \sum_{m,n=1}^{p} c_{m,n}(1+r)\sqrt{\lambda_m}\sqrt{\lambda_n} \tag{3.2.8}$$

We have defined the correlation such that for $i, j \leq n_c$, $c_{i,j} = c$; for i or j different from n_c, $c_{i,j} = 0$; and of course $c_{i,i} = 1$.

We thus reach the following expression for $\bar{\sigma}$:

$$\bar{\sigma}^2 = \sum_{j=1}^{p}(1+r)\lambda_j + \sum_{i,j=1,i\neq j}^{n_c} c(1+r)\sqrt{\lambda_i\lambda_j} \tag{3.2.9}$$

where we can see that only n_c synapses receive correlated inputs with correlation coefficient c and the rest of them are independent. We finally reach the following expression for the dynamics of the membrane potential:

$$dV = -\frac{Vdt}{\gamma} + a(\bar{\mu}dt + \bar{\sigma}dB(t)) \tag{3.2.10}$$

Expressing the membrane potential with this continuous process approximation is powerful because it gives us an interpretation of the membrane dynamics in terms of three meaningful components: a deterministic leak ($-\frac{Vdt}{\gamma}$), a drift ($\bar{\mu}dt$) and a noise term($\bar{\sigma}dB(t)$), defined by the well known and simple Brownian motion, that has been studied extensively by physicists. Furthermore, it gives us a simple update rule for numerical simulations, as modelling every incoming Poisson process would be too complicated. In order to simulate the Brownian motion, we took, at each time step dt, a value from the normal distribution $\mathcal{N}(0, \bar{\sigma})$. During the simulations, we used the following parameter values:

- The ratio between inhibitory inputs and excitatory inputs: r was variable.

- The number of synapses (corresponding to the number of dots in an experiment): $p = 100$.
- $a = 0.5$ mV: The EPSP and IPSP.
- $\lambda(j)$, the rate of the Poisson process incoming to the j^{th} synapse, is proportional to the direction of the j^{th} dot. 0 Hz$< \lambda <$ 100 Hz, the coherent signals being $\lambda = 25$ Hz or $\lambda = 75$ Hz.
- The time decay parameter: $\gamma = 20$ ms.
- The time step for the integration: $dt = 0.01$ ms.
- N_{spikes} is the number of output spikes that we used to measure the output FR.
- The correlation coefficient between coherent motions: $c = 0.1$.
- The number of coherent inputs: $n_c \leq p$ is variable. Coherent inputs are dots that move consistently in one direction.
- The resting membrane potential: $V_{rest} = 0$ mV.
- The threshold membrane potential: $V_\theta = 20$ mV.

When the membrane potential V crosses the threshold V_θ, a spike is generated and the membrane potential is reset to $V_{rest} = 0$ Hz. In this model, we did not use any refractory period.

3.2.2 Experimental methods

We repeatedly presented the model with kinematograms, and measured the output Firing Rate (FR) of the MT neurons. For the simulations presented further in this chapter, we did not use the early visual cortex model presented in the previous subsection. We directly produced the vector of Poisson rates corresponding to the output of the model: If a certain percentage of the dots moved coherently, the same percentage of the FR had a constant, equal rate, and the rest was randomly updated. Thus, even the drift term in Eq. 3.2.10 varied with time and between simulation. Furthermore, we evaluated the FR by measuring the time t needed to produce a hundred spikes: $FR = \frac{100}{t}$, which is not a perfect measure. Consequently, the measured FR changed from one simulation to another and had a distribution. So, we considered it as a random variable, measured its value over a sample of simulations and we assumed that it had a Gaussian distribution:

$$G(x) = \frac{1}{2\pi \cdot \sigma^2} exp^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

where σ^2 is the variance of the sample of measured FRs and μ is its mean. This model for the FR repartition is illustrated in Fig. 3.3.

So far, we have only considered a single MT neuron. However, as explained in section 2.2.1, it is more realistic to measure the FR over a population of neurons. To evaluate the FR, we used the same method as for the single neuron: We measured the time needed to produce 100 spikes. Our population of discriminating neurons was homogeneous, not laterally connected and received inputs coming from the same kinematograms: The vectors of Poisson rates were the same. However, we did not assume that all the neurons received the same spike trains, and consequently the time courses of their membrane potentials were modelled as different realisations of the process of Eq. 3.2.10. We simulated the discrimination tasks with various numbers of neurons, but the main results that we present in the current work were obtained with a population of a hundred units.

3.2.3 The Total Probability of Misclassification

The discrimination boundary

Given one moving dot's input, the model gives us an output FR. From this FR, we want to be able to tell the direction of the general motion of the kinematogram. The criterion will be the comparison of this FR to a value called the discrimination boundary: If the FR is greater than this value, then the movement is upwards, otherwise downwards. How do we define the discrimination boundary? This value is the point where the two Gaussian curves that model the FR distributions for the upwards and downwards motions intersect. This value can easily be formally determined as follows.

First, after storing the 100 FRs for an downward motion and the 100 FRS for a upward motion, we measure they respective means (μ_1 and μ_2) and standard deviations (σ_1 and σ_2), from which we derive the two equations of the normal distributions (Gaussian curves) that model they repartition ($y = G_{\mu_1,\sigma_1}(x)$ and $y = G_{\mu_2,\sigma_2}(x)$). We then find the intersections of the two curves by solving the equation expressing the equality between their values: all x such that

$$G_{\mu_1,\sigma_1}(x) = G_{\mu_2,\sigma_2}(x) \tag{3.2.11}$$

and we choose the only intersection that is situated between the two means ($\mu_1 < x_0 < \mu_2$). x_0 is then the abscissa of the discrimination boundary.

The discrimination boundary is illustrated in Fig. 3.3. In this figure, we present two histograms that list the output FRs of a hundred repetitions of the discrimination experiment, as described in the previous section. We can create this illustration for every different value of the parameters, especially for a ratio between inhibitory and excitatory input strength varying from $r = 0$ to $r = 1$. However, for $r = 0$, the discrimination performance is lower as we will see in the following sections, so the histograms overlap more, thereby

serving the illustrative purpose of the figure better, even if the absence of inhibitory inputs is less biologically realistic.

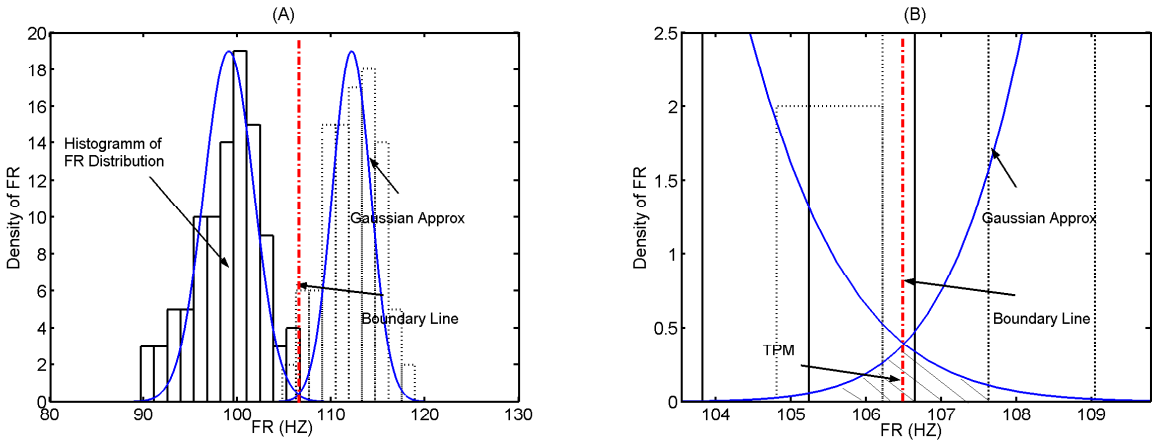


Figure 3.3: Graphical illustration of how the Total Probability of Misclassification(TPM) is calculated from the boundary line between two Gaussian distributions of FR, as explained by Eq. 3.2.11 in 3.2.3. (B) is a detailed zoom of (A). The histograms represent the distribution of FR measured for a hundred repetitions of the same experiments, upwards and downwards. The simulations were run for with a single neuron, for a coherence of 22%, the FR was measured on hundred spikes, and $r = 0$. We see that the Gaussians fit the distribution of FR reasonably well. Adapted from Gaillard et al. (2006b)

Calculation of the Total Probability of Misclassification

The Total Probability of Misclassification (TPM) is the number of misclassifications divided by the total number of classifications.

Using the Gaussian model for the output FR, these numbers will be areas: if we classify an upwards input as downwards, it means that the FR is inferior to the boundary, but belongs to the distribution of the larger FRs: this is the area of the right Gaussian curve that is on the left of the boundary. Therefore, the number of errors will be the surface of the right Gaussian curve on the left of the boundary, plus the surface of the left Gaussian curve on the right of the boundary.

The total number of classifications is the sum of the two surfaces under the Gaussian curves.

$$TPM = \frac{\int_{-\inf}^{x_0} G_{\mu_1,\sigma_1}(x)dx + \int_{x_0}^{\inf} G_{\mu_2,\sigma_2}(x)dx}{\int_{-\inf}^{\inf} G_{\mu_1,\sigma_1}(x) + G_{\mu_2,\sigma_2}(x)dx}$$

.

where $x_0, \mu_1, \sigma_1, \mu_2, \sigma_2$ are defined around Eq. 3.2.11.

3.2.4 Time

A fundamental performance measure in decision tasks is the Reaction Time. Before a decision can be taken, the subject needs to accumulate enough evidence, as we have seen previously. This part of the model produces evidence, in terms of spike trains. To evaluate its performance, it is important to know how fast this evidence is produced. The measure we took during these simulations was the time needed to produce a hundred spikes. This time is of course greatly reduced in the case of the population coding.

We also measured the loss of accuracy when processing time is artificially reduced. In this set up, we simulated the MT neurons for a predefined period of time, and evaluated the FR by dividing the number of spikes collected by that time. This FR evaluation, in turn, allowed us to measure the influence of urgency on the discrimination accuracy.

3.2.5 Controlling input noise with inhibitory inputs

r is the ratio between inhibitory and excitatory inputs. If $r = 0$, the MT neuron exclusively receives excitatory signals, whereas if $r = 1$ it receives as much excitatory as inhibitory inputs. We can see in Eq. 3.2.10 that the stochastic perturbation is proportional to $(1 + r)$ whereas the drift is proportional to $(1 - r)$. We call input noise the term $\bar{\sigma} dB(t)$. What is the effect of increasing the strength of this input noise on the efferent FR distribution over successive simulations? We have measured how this input noise increases with r and c, the correlation value between the inputs generated by the coherently moving dots, and how it affects the noise to signal ratio of the efferent signal. We present the results in Fig. 3.4. The post-synaptic input is the quantity defined by $I_{syn}(t)$ in Eq. 3.2.6. The data was generated, for each value of r, by generating the stimulus 100 times, measuring the parameters σ and μ of the post-synaptic input, and taking the mean value of $(\frac{\sigma}{\mu})$ over these 100 repetitions. For the measure on the outputs, (curve (B)), we generated a sample of FRs, one output FR value being generated by one simulation of the input presented to the neurons. After one hundred repetitions of this simulation, we obtained one hundred FRs and measured the mean and standard deviation of the sample. We repeated the operation for each value of r and c that are presented in the figure. In order to measure the output FR of the neuron, we measured the time T needed for the neuron to produce 100 spikes and computed $FR = \frac{100}{T}$.

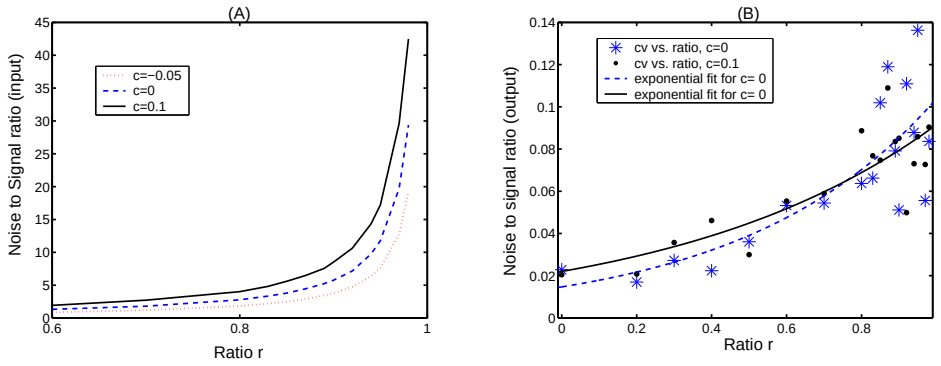


Figure 3.4: In (A) we show the noise to signal ratio ($\frac{\sigma}{\mu}$) of the post-synaptic input to the neuron as a function of r. The post-synaptic input is the quantity defined by $I_{syn}(t)$ in Eq. 3.2.6. This ratio increases dramatically with the strength of inhibitory inputs (i.e. r), and with synaptic correlation. c is the value of the correlation coefficient of the coherent inputs. (B) shows the signal to noise ratio (standard deviation divided by mean) of the sample of output FRs measured on 100 experiments, with and without synaptic correlation ($c = 0$ or $c = 0.1$). The curves are exponential fits with two parameters. The standard errors of the regression (Root Mean Squared Errors of the fits) are respectively $RMSE = 0.022$ ($c = 0$) and $RMSE = 0.014$ ($c = 0.1$). In (B) we see that the signal to noise ratio of the neuron's output is hardly affected by the synaptic correlation, only by r.

As explained in Feng (2001), increasing the coefficient of variation of the input will increase the coefficient of variation of the efferent spike train of the neuron. This increased variability affects the width of the histograms that represent a set of measures of the output FR. Thus, intuitively, it should be more difficult to discriminate between two inputs from their efferent FRs. However Deng et al. (2003) have formally proved that the discrimination is easier when the coherent inputs (those upon which we discriminate) are correlated. They proved, in the general case of an Integrate and Fire neuron discriminating between mixed inputs, a set up that corresponds exactly to our current model, the following theorem (theorem 2 of the paper): *When the correlation coefficient c is greater than 0, the output histograms become more separated than the input histograms; when $c < 0$, the output histograms are more mixed than input histograms. $c = 0$ is the critical case.*

This is a fundamental result, because it enables us to have a better discrimination performance thanks to the correlation, despite the fact that correlation increases the input noise, as we see in Fig. 3.4. In most of the simulations presented below, we used a correlation coefficient of $c = 0.1$ for synapses that receive the coherent input. This is justified by the organization of area MT: Neurons detecting the same direction will be grouped together and it has been shown by Zohary et al. (1994) that in area V5 of the visual cortex of the monkeys, the level of correlation is $c = 0.1$.

3.3 Results

The aim of this model was to study the effect of input noise on the discrimination performance of the LIF model of MT neurons. The ratio between inhibitory and excitatory synapses increases the post-synaptic input noise but also improves the discrimination accuracy. We can see an illustration of this result in Fig. 3.5.

On the other hand, we see in Eq. 3.2.10 that when the number of inhibitory synapses increases (increasing henceforth the input noise) the drift term decreases. Ultimately, when $r \to 1$, the drift tends to zero. Thus the membrane potential increases much slower leading to very low FRs. Very low FRs in turn induce long processing times for the neuron to produce hundred spikes, which is detrimental to the overall performance of the model.

In the following subsections, we reproduce parts of our published work that analysed this trade-off between time and accuracy, and compare the performance of a population of neurons to the performance of a single neuron.

3.3.1 Increasing the input noise improves the discrimination.

As we can see in this section, adapted from our paper Gaillard and Feng (2005), increasing the input noise improves the discrimination. Furthermore, even if we measure the FR over the first hundred spikes with no regards to time, the population performs better than the single neuron.

Extensive simulations over the range of r and over the range of input coherence (percentage of coherently moving dots) produced the following results, summarized in Fig. 3.6:

- Obviously, the TPM decreases when the coherence increases: the more separated the inputs are, the easier the discrimination task is.
- The TPM decreases non-monotonically when r increases. For the single neuron, the better performance achieved by increasing the input noise occurs only for $r > 0.7$.
- The population performs much better, for almost one order of magnitude, than the single neuron and its TPM decreases steadily with r.

The following adapted from our paper Gaillard et al. (2006b) explains why the population discriminates better, even if the observer waits for a hundred spikes, with no regards to processing time.

The better performance of the population can be explained as follows: The efferent FR of the population is not exactly the same as that of a single neuron because the assumption of ergodicity does not hold. In the

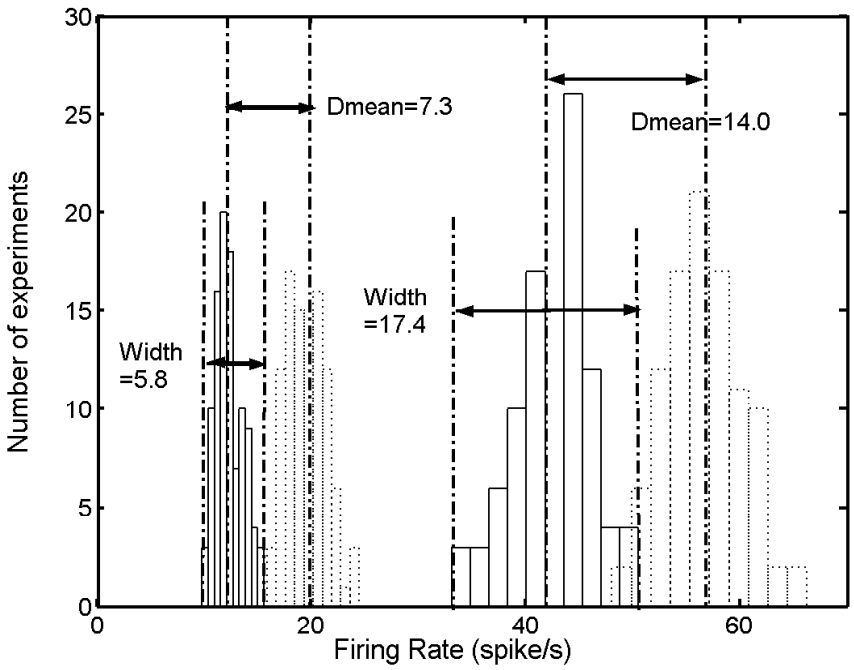


Figure 3.5: Illustration of the improvement of the discrimination accuracy with r. Each histogram is a graphical representation of the distribution of the efferent FRs over a hundred repeats of the same experiment. From left to right: $r = 0.98$, downward; $r = 0.98$, upward; $r = 0.7$, downward; $r = 0.7$, upward. The coherence of the kinematogram was 15%. We see that the difference between the means of two histograms decreases much less than the width of each histogram, when r increases. Whilst the difference between the means of upwards and downwards input is approximately halved ($\frac{D^{0.7}_{mean}}{D^{0.98}_{mean}} = 1.9$), the difference between the width of the histograms is almost divided by 3: In the downwards case, ($\frac{D^{0.7}_{width}}{D^{0.98}_{width}} = 3.02$). That leads to the separability of the output (the histograms do not overlap) when $r = 0.98$, whilst the responses are not separable (histograms do overlap) when $r = 0.7$. This performance improvement is quantified by measuring the TPM: $TPM_{r=0.98} = 0.01$ whilst $TPM_{r=0.7} = 0.02$. Adapted from Gaillard et al. (2006b)

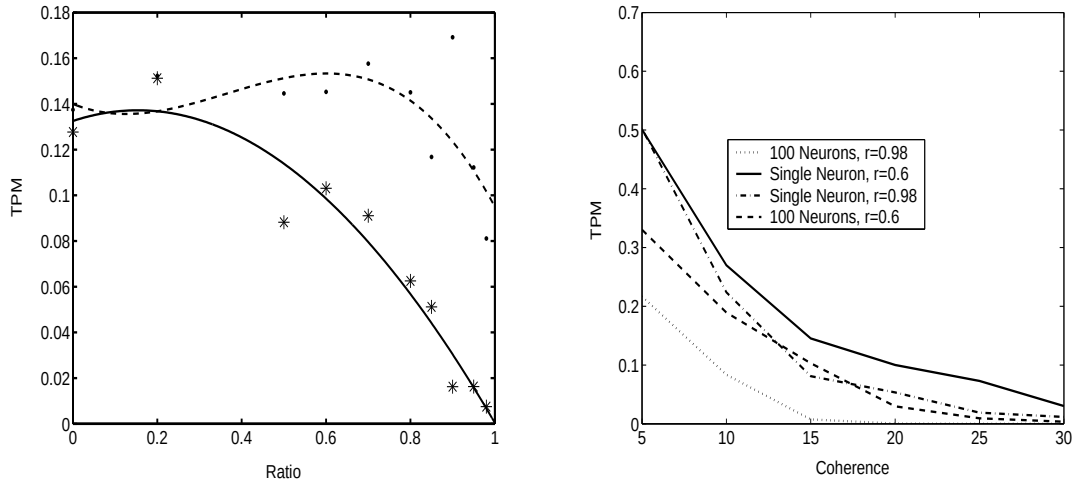


Figure 3.6: Comparison of the TPM of one single neuron and of a population, for various r and coherences, using 100 spikes. Left panel, coherence = 15 %. The time window needed to collect these 100 spikes varies a lot with parameter values, especially it increases dramatically with r.

population approach, we used the first hundred spikes of a hundred neurons to measure the FR, which means that we used on average one spike per neuron, long InterSpike Intervals were unlikely to be produced. In that case, on average, the population will produce a hundred spikes before one neuron can have generated a spike with an InterSpike Interval longer than twice the mean InterSpike Interval. Thus, across all the experiments, we do not obtain exceptionally low FRs, which are generated by exceptionally long ISIs in the corresponding efferent spike train. That means that the distribution of FRs, represented as an histogram, such as in Fig. 3.5 or Fig. 3.3, has its left part cut. Consequently, the histograms overlap less, which ultimately means that the TPM is lower. Graphically, it means that the output FR histograms are better separated than what we see in Fig. 3.5, and consequently the discrimination accuracy of the model is increased.

3.3.2 Time versus accuracy trade-off

In the following, adapted from our paper Gaillard et al. (2006b), we show how input noise can control the trade-off between processing time and accuracy.

Time to a hundred spikes

In this part, we want to compare the time it takes for the model to generate 100 spikes, which is a number that enables us to reliably estimate the FR. This approach enables us to compare the two approaches

with an objective criterion. Fig. 3.7 shows us the main improvement of using a population code: It is much quicker. The population code is very simple. Each neuron of the population is connected to the input in the same way as the others, receiving the same mean and noise values for its post synaptic input, and there are no lateral connections, the neurons of the population are independent from each other.

In this Fig. 3.7, we see a dramatic increase of the time when r tends to one. This is related to the fact that, with perfectly balanced inputs, there is no deterministic synaptic input: the only post synaptic input is created by synaptic noise, which means that reaching the threshold potential to generate a spike is exceptional.

The time to hundred spikes decreases with the coherence of the stimuli: the more dots move coherently, the larger the FR. This is due to the correlated synapses: if more dots move coherently, the positive correlation of the coherent synapses will lead to a generally larger input noise.

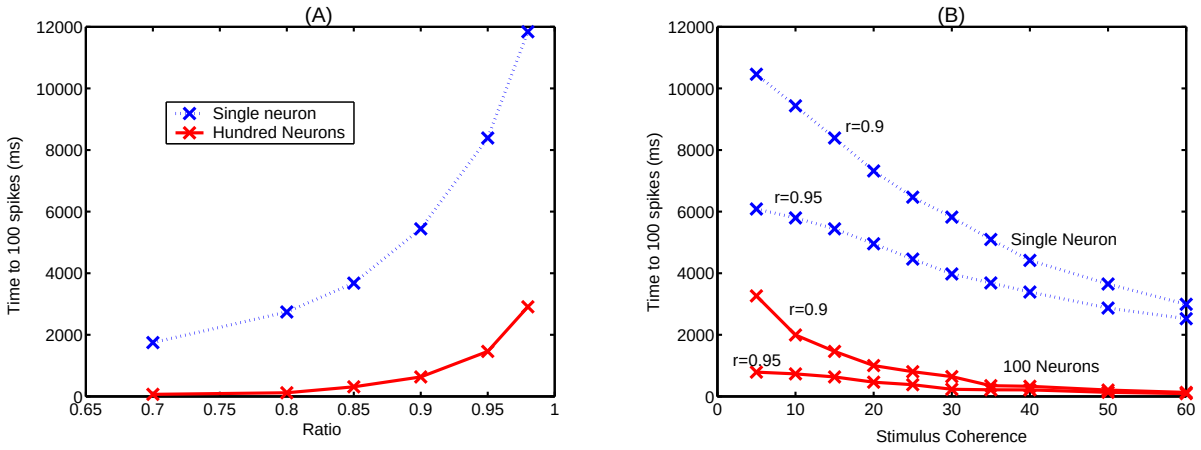


Figure 3.7: Left panel (A): Time to get a hundred spikes versus r, with a population of a hundred neurons and with a single neuron. Coherence=15. We see that this time increases dramatically when r tends to one. As predicted, the population of a hundred neurons generates hundred spikes more quickly, between five and ten times quicker. Right panel (B): Time to get a hundred spikes versus the coherence of the stimulus, with $r = 0.9$ and $r = 0.95$. The dotted lines represent the results for a population of a single neuron, the solid lines, one hundred.

Obviously, the population code is a much faster code: in the case of an almost balanced ratio r, the single

neuron needs around ten seconds to generate a hundred spikes, for coherence values typically between ten percent and twenty percent. Furthermore, the time to one hundred spikes of the population code, with a ratio r close to 1, only goes up to the order of one second with a hundred neurons, and 100 ms with a thousand neurons, without any loss of accuracy. With the increase of the population size, the time to generate a hundred spikes significantly decreases. Using the first hundred spikes of a thousand neurons, we would use 0.1 spike per neuron, on average.

TPM as a function of the size of the time window: a trade-off between accuracy and time

We know that, for most of the biological systems the absolute performance must take into account not only the accuracy at realizing the task, but also the time spent to achieve it (see, for example Ratcliff et al. (1999) or Usher and McClelland (2001)). So, processing time is an important criterion for comparing the various models. Thus, to put the TPM in perspective, we have to measure the evolution of the quantity of errors with the size of the time window during which we collect spikes. We can see, in Fig. 3.8, the performance of the

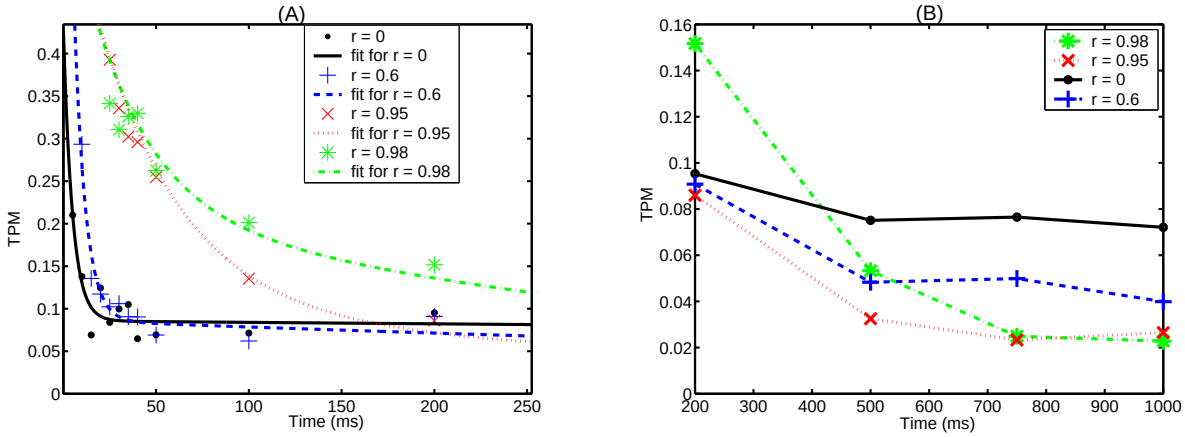


Figure 3.8: Comparison of the evolution of the TPM with the size of the time window, for $r = 0.98$, $r = 0.95$, $r = 0.6$ and $r = 0$. Left panel (A): Zoom on the behaviour for shorter time windows ($t < 250ms$). These curves show that the TPM decreases much faster with the time when r is low. The curves are exponential fits to the data with 4 parameters. The Root Mean Squared Errors of the fits are respectively $RMSE = 0.052$, $RMSE = 0.044$, $RMSE = 0.018$ and $RMSE = 0.022$. Right panel (B): Comparison of the evolution of the TPM for long time windows, reaching to one second, for $r = 0.98$, $r = 0.95$, $r = 0.6$ and $r = 0$. The data points are obtained by measuring the TPM, as explained in the Appendix. We see that for longer time windows, the 'slower' strategy that consists in using $r \approx 1$ outperforms the fast, small ratio approaches.

model for various values of r, as a function of the time, within a one second time window. It shows clearly that we reach much better results with $r \approx 1$ than for $r = 0$, but that we reach acceptable performances much faster with $r \approx 0$. Bearing in mind that significantly increasing the size of the population would allow us to reach such results in a significantly shorter time, these high performances seem relevant to the natural time constraints.

Time to a good discrimination

Because of the Gaussian approximation of the distribution of the efferent FRs over 100 repetitions, the TPM will never be exactly zero. However, we can consider that a good discrimination performance is attained when $TPM = 0.1$, and that an acceptable discrimination performance is reached for $TPM = 0.2$. As seen in the previous subsection, the TPM of the model depends on the size of the time windows during which we collect the spikes in order to accurately evaluate the FR of the population. This is a typical case of speed-accuracy trade-off.

We have measured $TPM_r(t)$, as in Fig. 3.8, for the whole range of ratios, and evaluated t_r such that $TPM_r(t_r) = 0.1$ or 0.2 respectively. We have performed this analysis for a coherence equal to 15%. Then we have fitted these data points with an exponential model with four parameters. The data clearly has an exponential behaviour, and we had to use four parameters to obtain satisfactory Standard Errors between the exponential fit and the data. Thus we obtained numerical evaluations of the time to a good TPM ($TPM = 0.1$) as a function of r ($T_1(r)$), and of the time to an acceptable TPM ($TPM = 0.2$) as a function of r ($T_2(r)$):

$$T_1(r) = a \times e^{(b.r)} + c \times e^{(d.r)}$$

$$T_2(r) = f \times e^{(g.r)} + h \times e^{(k.r)}$$

where $a = 27.94$, $b = 0.6279$, $c = 1.903 \cdot 10^{-8}$, $d = 23.8$, $f = 3.349 \cdot 10^{6}$, $g = 6.936$, $h = -3.349 \cdot 10^{6}$, $k = 6.936$.

We can see an illustration of $T_1(r)$ and $T_2(r)$ in Fig. 3.9 that also shows the data points that were used to evaluate the fitting.

3.3.3 Discussion

The following extract from our paper Gaillard et al. (2006b) discusses the results that we presented in this section.

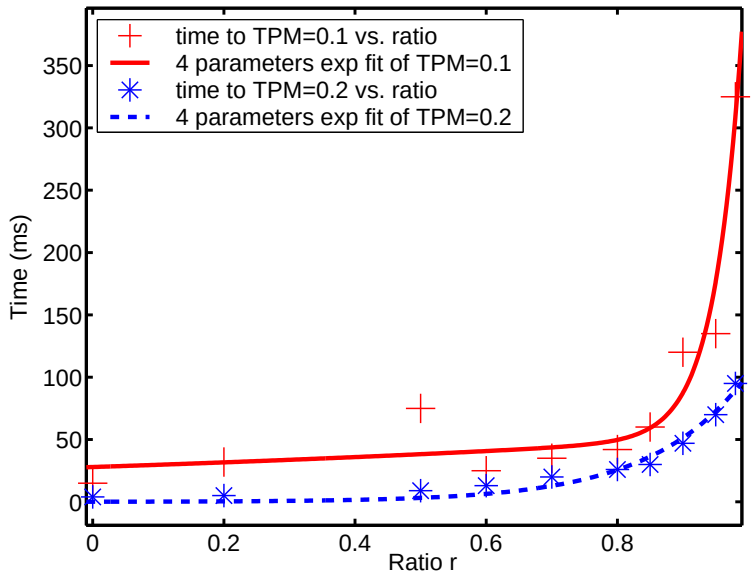


Figure 3.9: Illustration of the numerical estimation of the time to reach an acceptable discrimination performance. The coherence of the input is 15%. The data was obtained using the evaluations of the function representing TPM vs. Time. For each value of r we found t such that $TPM(t) = 0.1$ or $TPM(t) = 0.2$. The curves are exponential fits to the data with 4 parameters. The Root Mean Squared Errors of the fits are $RMSE = 28.74$ $(TPM = 0.1)$ and $RMSE = 6.602$ $(TPM = 0.2)$

Summary of the results

As a first result, we have shown that the discrimination performance increases with the ratio r between inhibitory and excitatory inputs, given that we use the same number of spikes to measure the efferent FR. This is counter-intuitive because, on the other hand, r increases the post-synaptic input noise.

We obtained the surprising result that the population code, without lateral connections is more accurate than the single neuron when we use the same the number of spikes to evaluate the FR. Leaving aside any time considerations, measuring the FR on the first hundred spikes of the population enables much more accurate discriminations than measuring the FR on the first hundred spikes of one neuron. This is due to the fact that we rule out the long InterSpike Intervals.

Although increasing the ratio r increases the performance, it also increases the processing time significantly. The study of the discrimination performance as a function of the time window during which we collect spikes shows that the probability of misclassification decreases much faster for the smaller ratios. We numerically evaluated an expression of the time needed to reach an acceptable TPM as a function of the ratio.

Furthermore, we have seen that only ratios close to one can reach a level of performance unreachable by the FR of a population with exclusively excitatory synapses. These very high performances are reached at the cost of a very long discrimination process.

Experimental evidence for our model predictions

Romo et al. (2003) show that positive correlation is not always harmful to discrimination accuracy, as it can successfully remove the noise from the input. In their article, they show that this happens when pools of sensory neurons have opposite slopes characterising their activity versus intensity of stimulus functions.

Their research experimentally confirms what our model predicts with numerical simulations. In our model, increasing the intensity of the inhibition (i.e. the ratio r) improves the discrimination accuracy when the correlation is positive (here, $c = 0.1$), but is useless if $c = 0$. The activity (FR) of the inhibitory input is modelled to be the same as the activity (FR) of the excitatory input, but with the opposite effect on the synapse: therefore, the post-synaptic effect is opposite. When the post-synaptic input contributions are plotted against the coherence, the inhibitory and excitatory contributions have opposite slopes, and are positively correlated. In both their research and ours, this leads to improved discrimination accuracy. These situations are particular cases of noise subtraction by negative correlation. This phenomenon has been neglected but is now being studied by Durrant and Feng (2006).

Building block of larger decision models

Our model is consistent with larger models that simulate decision processes. For example, Gold and Shadlen (2003) proposes a model for how a decision variable such as the ones we can find in area LIP accumulate motion information that is represented in area MT and MST. Our model gives detailed insights into how the very general motion information represented in MT or MST can be coded by IF neurons into a signal that enables discrimination relevant to the task. We can see in the model of Mazurek et al. (2003), presented in Fig. 2.14, that MT neurons are selective for the direction of the input stimuli, the kinematograms. The output of these selective MT neurons is used as evidence to be used by a further stage of processing (area LIP) where the discussion is taken. In Mazurek et al. (2003), we see three characteristics of this modelling: the FR of the MT neurons is used as evidence, it grows linearly with the coherence, and, as shown in the neural simulation (see again Fig. 2.14), the variability of the ISIs of the output of MT neurons is important. The model presented in this chapter corresponds to this type of models: we extract evidence from the sensory stimulus and express it as FR despite a high variability in the output spike trains.

This model of evidence generation is also used in Wang (2002). He does not model the visual cortex or even area MT but he takes assumptions regarding the evidence used in his Winner-Takes-All model, as

described in section 2.3.2. He assumes that the evidence is generated by area MT, that the output FRs of MT neurons whose receptive fields are centred on one kinematogram direction have a linear relationship with the coherence of the kinematogram, and he acknowledges that this output is highly stochastic. Consequently, he models the evidence for the two opposite alternatives as two overlapping Gaussian distributions and uses them as inputs to his decision model of LIP. As we will see in the next chapter, a simple generalisation of our model of MT neurons renders it perfectly consistent with an output FR proportional to the coherence of the kinematograms. Our model also corresponds to Wang's assumption in so far as we modelled the distribution of output as Gaussians centred at their preferred direction, as explained in 3.2.3.

From a formal probabilistic modelling point of view, we can look at the work of Ratcliff and Rouder (1998). He also posits that decision making is based on accumulation of evidence, whereby a decision variable integrates available evidence over time and triggers the expression of a decision when a threshold is reached. The decision process can be made faster or slower, by modifying the parameters of the network. The first idea found in the literature was to lower the threshold. However Ratcliff proposes that the speed at which the decision variable accumulates evidence, the 'drift rate' as seen in section 2.3.2, varies with the difficulty of the stimuli. In the case of the kinematograms, the higher the coherence, the higher the drift rate. In his modelling work, he showed that this drift rate variation explains the variations of the Reaction Time and accuracy of the decision better than threshold lowering methods. It is straightforward to imagine how the drift rate depends on the information given by our model: It will be lower if the Total Probability of Misclassification is high, because the FRs corresponding to wrong classifications will have a negative effect on the accumulation process. In the next chapter we will see a generalisation of our model of MT neurons into direction detectors, and how their FR is used for accumulating evidence in favour of one or the other alternative. In that case, the drift rate is directly linked to the difference of FR of each detector. If the TPM is high, then the difference of FRs is small, and so is the drift rate. So, we see that the model presented in this chapter fits well with the hypothesis that the 'drift rate' of the accumulation of evidence varies with the difficulty of the stimulus.

Speed-accuracy trade-off

Our model enables better discrimination at the cost of longer processing times. This trade-off between accuracy and speed of processing is a typical pattern of decision making. The time needed to reach an efferent spike train that reflects a reasonable decision goes from a few dozen of milliseconds to hundred of milliseconds. This large discrepancy is consistent with experimental results studied by Shadlen and Gold (2004), where they show that the reaction time of a monkey, in this precise discrimination task, can go up to 800 ms, according to the required reliability and to the difficulty of the task. Of course we only model one neuronal layer, so we cannot precisely compare our results to the Reaction Time, but the magnitude is

compatible. We will see in the next chapters how we integrate this model of MT neurons in a more general decision making model that enable us to compare more precisely our results with real Reaction Times.

Chapter 4

Decision making induced by perturbing a dynamical system

Noise is a fundamental feature of neural activity. More precisely, it has been shown e.g. by Amit and Brunel (1997), or Shadlen and Newsome (1994) that the cortical neurons have a random low level spontaneous activity of around 5 spikes per second and that, due to the sheer number of synapses, this activity is fundamental to the post-synaptic neuron. Initially people considered it detrimental, or that it reflected non relevant patterns of activity. More recently, the influence of noise on the speed of convergence of non-linear systems with various attractors has been analytically studied by for example Berglund and Gentz (2002). So, in the domain of decision making, the background activity of the brain, that was mainly used as a way to explain errors as in Ratcliff and Rouder (1998); Ratcliff et al. (1999), is now getting a closer look as a catalyst for contextually dependent decision making as seen in Salinas (2003). In this chapter, we find out in more detail how noise can influence the dynamics of decision making. We are currently publishing the core idea of this chapter in Gaillard et al. (2006a), but we present here a more precise and developed description.

4.1 Proposed Model

4.1.1 Overview of the model

The model uses eight populations of a hundred Integrate and Fire neurons each to perform a decision based on an imperfect discrimination between highly mixed stimuli, and expresses the decision with a saccadic eye movement. It can be divided into two different parts. First we have what we called global direction detectors, which model direction discrimination in area MT. They are a generalisation of the model described in the previous chapter. The second part models a neural column in area LIP that uses the output of the direction

detectors, receives global inputs from the brain and commands an eye movement that expresses the decision. We can see in Fig. 4.1 an overview of the decision model. Each block is a population of Integrate and Fire neurons, as described in section 3.2.1. On the left, we see a snapshot of the visual stimulus. The direction detectors are detailed in section 4.1.3. Following the direction detectors, we have our model of the LIP column, where the decision is formed. We can see in Fig. 4.1 the various subpopulations that are described in more details in section 4.1.4. As represented on the model, all the neurons in the LIP column are submitted to the background activity of the brain, which is characterised by its mean and variance. These two parameters are the ones that we vary in order to study their influence on the dynamics of decision making. Lastly, we see the eye, whose modelled behaviour is described in section 4.1.5.

We simulate the benchmark decision task based on kinematograms, as described in the introduction of this thesis. The kinematograms are composed of 100 dots. We simulate the decision making from area MT of the visual cortex to the eye movement, via Area LIP. We use two measures: Reaction Time (RT) and Error Rate (ER). The RT is the simulated time needed for the SC activity to reach the threshold necessary to initiate the saccadic eye movement that expresses the decision, and the ER is the number of errors divided by the total number of repetitions of the simulation in the same conditions. An error occurs when an upwards saccade has been generated while the coherent direction of the kinematogram is downwards.

4.1.2 Neural model

We use the LIF model described in the background section. The membrane potential of a neuron is described as in Eq. 2.1.5 by:

$$C\frac{dV}{dt}(t) = -\frac{V(t) - V_{rest}}{R} + I_{syn}(t) \tag{4.1.1}$$

where C is the membrane conductance, R the membrane resistance, $I_{syn}(t)$ is the synaptic input, V_{rest} is the resting potential and V is the membrane potential. When V reaches V_θ, the threshold potential, the neuron emits a spike, V is reset to the resting potential V_{rest} and, after a refractory period τ_{ref}, the membrane dynamics described by Eq. 2.1.5 resume.

Following the work presented in Feng et al. (2006), we assume that the spike trains that constitute the synaptic input to a neuron are renewal processes and are well described by the two first statistical moments of there InterSpike Intervals. Consequently, we use the Usual Approximation Scheme, as described in Eq. 2.2.14, and therefore $I_{syn}(t)$, the synaptic input is described, as in Eq. 2.2.17, by:

$$I_{syn}(t) = \bar{\mu}dt + \bar{\sigma}dB_t \tag{4.1.2}$$

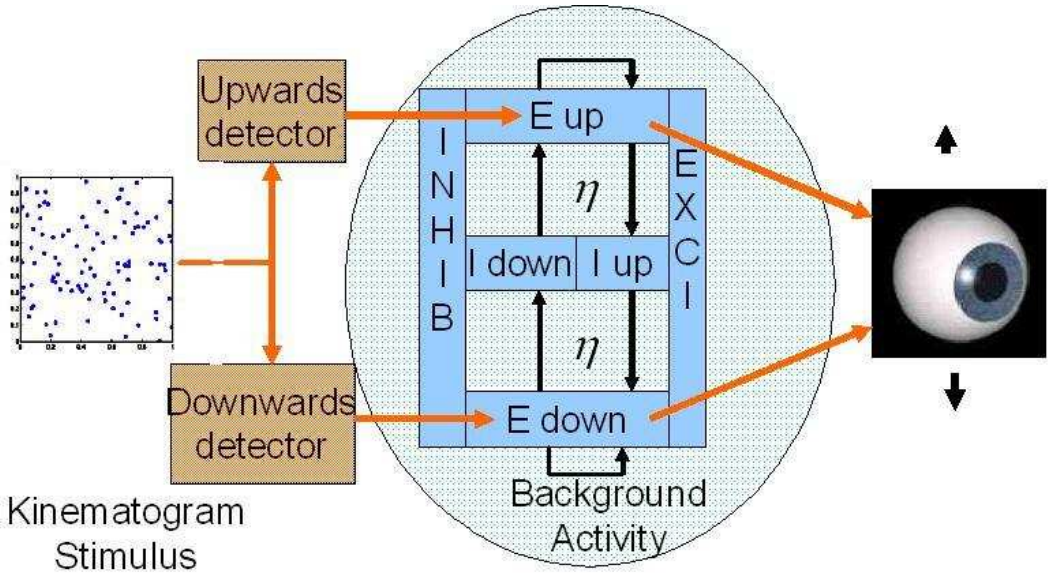


Figure 4.1: The left panel represents the stimulus: this is a snapshot of a kinematogram. In the kinematogram, most of the dots move randomly, but some move coherently in one direction. More dots moving coherently make the direction discrimination task easier, and the decision making quicker. The middle part is an illustration of the neuronal activities in the decision column of area LIP. In the middle part we see the low level background activity of the brain surrounding the column. All the represented subpopulations of neurons in the LIP column are connected to themselves and to each other, but the arrows represent potentiated connections. When the eye on the right has turned to one side, a decision has been taken.

In eq. 4.1.2, $\bar{\mu}$ and $\bar{\sigma}$ are defined as:

$$\bar{\mu} = \sum_{j=1}^{p} w_j \mu_j \tag{4.1.3}$$

$$\bar{\sigma}^2 = \sum_{m,n=1}^{p} w_m w_n \sigma_m \sigma_n c_{m,n} \tag{4.1.4}$$

where p is the number of synapses of the neuron, $c_{m,n}$ is the correlation coefficient between the incoming spike train from presynaptic neuron m and the spike train from presynaptic neuron n, as defined in Eq. 2.2.9, w_m is the synaptic weight from neuron m to the neuron under consideration, as described in Eq. 2.1.11, in Chapter 2. μ_i and σ_i are the mean and standard deviation of the continuous approximation of the spike train coming from the presynaptic neuron i, as described in Eq. 2.2.15.

Similarly to the previous chapter in 3.2.1, we model the spike trains to be Poisson processes. This is a classical and widely used approximation on spike trains as justified in for example Feng (2003a) or Gerstner and Kistler (2002), and it significantly simplifies the modelling. For each incoming spike train from neuron i, we then have: $\mu_i = \nu_i$ and $\sigma_i = \sqrt{\nu_i}$, where ν_i is the rate of the Poisson process. $\bar{\mu}$ and $\bar{\sigma}$ then become:

$$\bar{\mu} = \sum_{i=1}^{p} w_i \nu_i \tag{4.1.5}$$

$$\bar{\sigma}^2 = \sum_{i=1}^{p} w_i^2 \nu_i + \sum_{i,j=1, i\neq j}^{p} c_{i,j} w_i w_j \sqrt{\nu_i \nu_j} \tag{4.1.6}$$

Still following the Moment Network Theory, in the specific case of Poisson processes, we then obtain an expression for the output mean InterSpike Interval ($\langle ISI \rangle$), following Eq. 2.2.20:

$$\langle ISI \rangle = \tau_{ref} + 2\tau \cdot \int_{A(V_{rest})}^{A(V_\theta)} g(x) dx \tag{4.1.7}$$

where τ_{ref} is the refractory period, $\tau = R \cdot C$ is the time constant of the neuron and the limits of the integral are:

$$A(y) = \frac{\frac{y}{\tau} - \bar{\mu}}{\bar{\sigma}\sqrt{1/\tau}} \tag{4.1.8}$$

The function g is defined as:

$$g(x) = e^{x^2} \int_{-\infty}^{x} e^{-u^2} du \tag{4.1.9}$$

Let us rewrite Eq. 4.1.7 in terms of usual functions and output firing rate. We will use the usual function called error function and complementary error function:

$$erf(x) = \frac{2}{\sqrt{\pi}} \int_{0}^{x} e^{-t^2} dt \tag{4.1.10}$$

$$erfc(x) = \frac{2}{\sqrt{\pi}} \int_{x}^{+\infty} e^{-t^2} dt = 1 - erf(x) \tag{4.1.11}$$

we can rewrite g as follows:

$$g(x) = (\frac{\sqrt{\pi}}{2} \cdot e^{x^2}) \cdot \frac{2}{\sqrt{\pi}} \left[\int_{-\infty}^{0} e^{-t^2} dt + \int_{0}^{x} e^{-t^2} dt \right] \tag{4.1.12}$$

which leads to:

$$g(x) = (\frac{\sqrt{\pi}}{2} \cdot e^{x^2}) \cdot (1 + erf(x)) \tag{4.1.13}$$

Therefore, we can express the output firing rate $\nu_{out} = 1/\langle ISI \rangle$, that is also the rate of the Poisson process that models its output spike train, as seen in Eq. 2.2.8:

$$\nu_{out} = \left[\tau_{ref} + \tau \int_{A(V_{rest})}^{A(V_\theta)} \phi(u) du \right]^{-1} \tag{4.1.14}$$

where

$$\phi(u) = \sqrt{\pi} e^{u^2} (1 + erf(u)) \tag{4.1.15}$$

In the previous chapter, the output FR was the inverse of the mean ISI, measured over a hundred spikes. Here, it would require too much computational power to model each spike. Consequently, we use the statistical approximation of the output FR that we presented in Eq. 4.1.14. These equations for the input-output relationships of Integrate and Fire neurons submitted to either Poisson of merely renewal processes have also been directly used in the programming of the models studied in Appendix A which was inspired by the work of Feng et al. (2006). Using the well known error function and optimising the Matlab code (written in collaboration with Jianfeng Feng, himself in collaboration with Nicolas Brunel), we could precisely simulate

networks encompassing a significant number of neurons, as we will see in the simulation results presented further on in this chapter.

4.1.3 Direction detectors

The direction detectors are a generalisation of the MT discriminating neurons described in the previous chapter. In that chapter, the neurons received 100 Poisson inputs whose rates were proportional to the direction of the motion of the corresponding dot: A downward moving dot (direction $-\frac{\pi}{2}$) would generate an input of rate 25 Hz, an upward moving dot (direction $+\frac{\pi}{2}$) would generate an input of rate 75 Hz. It was convenient for discriminating between upwards and downwards movements, but:

- It was not optimal, because it did not use the whole span of rates. In the model, we considered that the maximum incoming firing rate would be 100Hz and the minimum 0Hz. Since we want to discriminate between motion directions, the optimal encoding would associate the most different motions (opposite directions such as upwards/downwards) to the most distant codes (firing rates). The distance between two FRs is the absolute value of their difference. Therefore, optimally, a downwards moving dot would be associated with 0 Hz and an upwards moving dot would be associated with 100 Hz.
- It induced inconsistencies in case of other moving directions: In that scheme, a dot moving to the right but slightly upwards would be associated to a rate of 100 Hz, whilst a dot moving in almost the same direction but slightly downwards would generate a rate of 0 Hz. The difference between their encoding would be 100Hz. That would mean that these two dots would be treated as more different than two dots moving vertically in opposite directions (50Hz encoding difference).

The direction detectors are a more versatile, more powerful and more realistic generalisation of that model. Each direction detector 'detects' a general direction in the kinematogram. It means that each detector has a preferred direction and its activity will be higher in the case of a similar general direction of the kinematogram. This set up overcomes the previously cited drawbacks and enables us to model direction detectors for any direction.

The preferred direction of the detector D_d and the moving dot's direction D_i lie in $[-\pi,\pi]$. First we define the distance between these two real numbers:

$$d_{d,i} = ||D_i - D_d|| \tag{4.1.16}$$

Since the directions of the dot and of the detector lie in $[-\pi,\pi]$, $d_{d,i}$ lies in $[0,2\pi]$, but the maximum distance is π because we are speaking about angles, not real numbers. This is illustrated in Fig. 4.2.

So, if $d_{d,i} > \pi$, then the angular distance is actually $\mathcal{D}_{d,i} = 2\pi - d_{d,i}$. We could express $\mathcal{D}_{d,i}$ with a mathematical notation but it would be too complex, whilst we did not use the expression in the code. In simple terms, we have two cases:

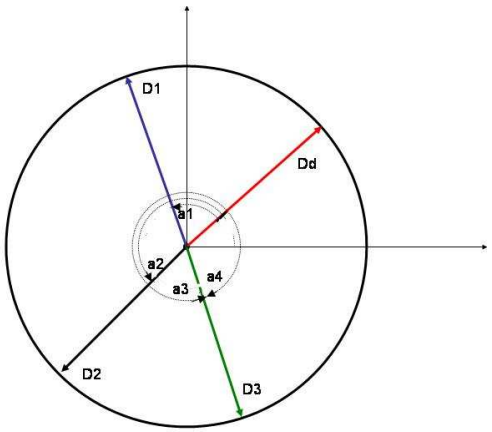


Figure 4.2: Illustration of the angular distance between a detector's preferred direction and various dot's motion's directions. The red arrow represents the preferred direction of the detector (D_d). D_1, D_2 and D_3 represent the directions of the motions of three different dots. a_1, a_2, a_3 and a_4 are the angular differences between the detector's preferred direction and the motion direction of the dot. We see that $a_3 = 2\pi - a_4$ and that the angular distance is $\|a_4\|$. a_2 illustrates the maximal angular distance between a dot's direction and the detector's opposite direction: π.

If $d_{d,i} < \pi$, then $\mathcal{D}_{d,i} = d_{d,i}$ (case 1).

If $d_{d,i} > \pi$, then $\mathcal{D}_{d,i} = 2\pi - d_{d,i}$ (case 2).

We then model the synaptic input to the direction detector of preferred direction D_d from the dot of moving direction D_i as a spike train whose FR is given by:

$$\nu_{d,i} = A(\pi - \mathcal{D}_{d,i}) + B \tag{4.1.17}$$

where $\nu_{d,i}$ is the rate of the Poisson process received by the i^{th} synapse of the d^{th} direction detector. A and B are parameters chosen so that the FRs remain in the range of rates used in our experimental models. The i^{th} synapse receives a spike train corresponding to the motion of the i^{th} dot, similarly to the structure presented in section 3.2. Therefore, in this work, $0 < i \leq 100$ because the kinematograms have 100 dots. In reality, such a direction detector would have many more synapses, but this simplification is justified because we pool all the synapses receiving inputs from the same dot, as explained in 3.2, and we model this combined input from many synapses to be an input with a high FR coming from only one synapse.

We assume as well that the incoming spike trains corresponding to dots having the same direction are correlated. n_c is the number of dots moving coherently and c is the correlation coefficient. The spike trains are modelled as Poisson processes and the correlation coefficient is defined in Eq. 2.2.9.

Fig. 4.3 shows that the FR of the efferent spike train of a direction detector increases linearly with the coherence of the kinematogram, when the kinematogram presents dots that move in the preferred direction of the detector. This result is in accordance with the literature, as for example in Wang (2002), where he assumes a linear dependency on the kinematogram's coherence: *"It is known that the response of an MT neuron increases with the input coherence in its preferred direction and decreases with it for stimulus in the direction* 180^o *opposite... I implemented this linear dependence by assuming linear relations"*. This is also consistent with the idea that the drift rate of the accumulation of the evidence variable depends on the difficulty of the task, as seen in Ratcliff and Rouder (1998), and discussed in section 3.3.3.

Furthermore, by sampling the output FRs of the direction detectors over many simulations, we could evaluate the distribution of these FRs. This distribution fits the assumption of a Gaussian distribution of FRs very well. We used this approximation to measure the TPM in section 3.2 and in our paper Gaillard et al. (2006b).

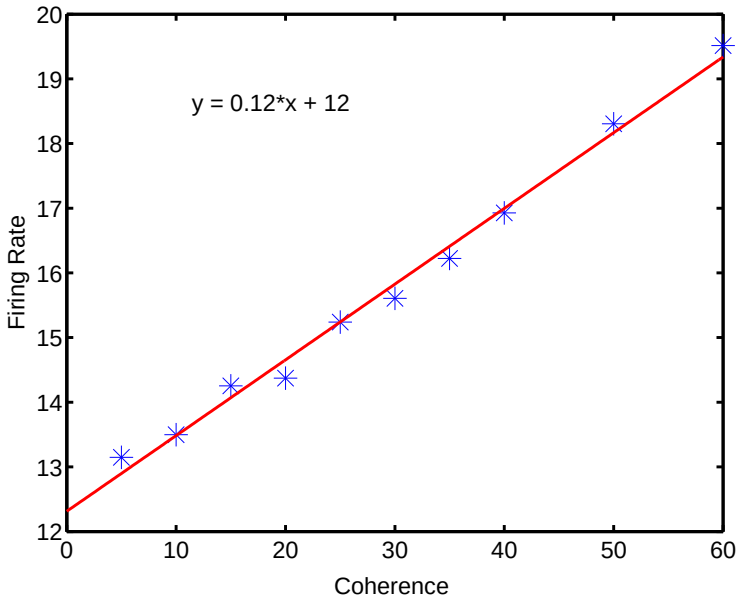


Figure 4.3: The output FR of a motion detector is a linear function of the coherence of the stimulus. The global direction of the stimulus is the preferred one of the detector.

4.1.4 Column organisation: Two competing subpopulations

The decision is performed by a competition between groups of LIF neurons that model a visual column in the Lateral IntraParietal area (area LIP) of the brain. In this column, groups of neurons inhibit each other and recursively excite themselves, which generates the competition. In fact, this competition is not solely stimulus driven but depends on the global neuronal activity of the brain because, besides the stimuli and the recurrent interactions, all the neurons of our model receive a low level excitation from the rest of the brain. This structure has several stable states: one of them occurs when all activities are weak, and other ones when one local group of neurons has a higher activity and dominates the others through lateral inhibition. The convergence to one of these stable states models decision making: The group of neurons that is sensitive to the specific stimulus has a comparative advantage over the others during the competition.

We have two specific inputs from the upwards direction detectors and from the downwards direction detector. We modelled each direction detector to be formed of a homogeneous population of 100 neurons. We modelled the LIP column to be constituted of the following 6 subpopulations of 100 neurons, as illustrated in Fig. 4.1:

- one subpopulation accumulating evidence for an upwards general direction of the kinematogram: $\mathcal{E}_{up}$
- one subpopulation accumulating evidence for an downwards general direction of the kinematogram: $\mathcal{E}_{down}$.
- one subpopulation of non specific excitatory neurons: $\mathcal{E}$.
- one subpopulation of non specific inhibitory neurons: I.
- one subpopulation of inhibitory neurons, with stronger connections from the upwards subpopulation and to the downwards subpopulation: I_{down}.
- one subpopulation of inhibitory neurons, with stronger connections from the downwards subpopulation and to the upwards subpopulation: I_{up}.

Potentiated connections

We have seen in 2.2.2 that one fundamental rule for learning and adaptivity is the Hebbian rule. The Hebbian rule states that the learning takes place at the synapse. It explains that if the postsynaptic neuron fires consistently soon after the presynaptic one, then the strength of the synapse is increased: it is potentiated. Conversely, if the firing occurs in the wrong order then the synapse is depressed. In the case of the LIF model that we are studying, we express this rule with the synaptic weights $w_{i,j}$ that we introduced in Eq. 2.1.11. In the current model, we assume that initially all the neurons had the same connection strength, for all i, j, k, n, $w_{i,j}^{init} = w_{k,n}^{init}$. Then through some learning (that we have not modelled in this dissertation), certain connections between neurons have been potentiated with a coefficient η:

$$w_{i,j} = \eta \cdot w_{k,n}^{init} \tag{4.1.18}$$

with $\eta > 1$ whilst others have been depressed with a coefficient η^-:

$$w_{i,j} = \eta^- \cdot w_{k,n}^{init} \tag{4.1.19}$$

In this model we defined η^- as a function of η, the number of potentiated synapses and the number of depressed synapses:

$$\eta^- = \frac{n^+ + n^- - n^+\eta}{n^-} \tag{4.1.20}$$

where n^+ is the number of potentiated synapses and n^- the number of depressed synapses. This way the total synaptic strength of the network remained constant.

Column connectivity

The neurons in the modelled LIP column receive inputs from neurons of all the subpopulations described above, but not from every neuron of the column. Furthermore, all the neurons of the LIP column receive background low level activity. However only two subpopulations receive synaptic inputs from the direction detectors: the upwards direction detector sends spikes only to the subpopulation $\mathcal{E}_{up}$, whilst the downwards direction detector only sends spikes to the subpopulation $\mathcal{E}_{down}$. The direction detectors do not either receive feedback from the LIP column. We can thus describe the synaptic input to one neuron of the column by two categories: *Internal input*, the input coming from the neurons of the column and *external input* the input coming from neurons outside the column, the direction detectors and the low level background activity.

As explained in Eq. 4.1.18 to 4.1.20, some synapses are potentiated with a coefficient η, while the rest of the synapses are depressed, in order to keep a normalized synaptic strength. The potentiated synapses are shown in Fig. 4.1.

We model that each of the 6 subpopulations of a hundred neurons is homogeneous. This means that two neurons from the same subpopulation will have the same connectivity (their synaptic weights are the same as well as the number of their synapses receiving inputs from specific subpopulations) and that the parameters of their LIF model will be the same. Consequently, the neurons will behave similarly and we can describe their behaviour with the equation that models one neuron.

The low level background activity is a neural activity spread in the whole cortex. However, the neurons of the LIF column that we model are not connected to all the neurons of the cortex. Similarly, the direction detectors are constituted by a population of neurons that do not all send their output to the LIP column. In order to specify how many synapses of one neuron of the modelled LIP column receive external input, we introduced a ratio $x = \frac{n_{in}}{n_{ext}+n_{in}}$ where n_{in} is the number of synapses receiving input from neurons of the LIP column (internal input), and n_{ext} is the number of synapses receiving external input, from the direction detectors or from the background activity. We also introduced the ratio between the number of external synapses receiving inputs from the low level background activity, and the number of external synapses receiving inputs from the direction detectors $f_b = \frac{n_b}{n_{dir}}$. We then define $f_1 = f_2 = \frac{1-f_b}{2}$, the fraction of external synapses receiving input from the two direction detectors input in comparison to the total number of external synapses. In the model, inhibitory synapses are slightly stronger than excitatory synapse: $w_{i,j_1} = -Cw_{i,j_2}$ where neuron j_1 is inhibitory and neuron j_2 is excitatory.

We mentioned earlier that we modelled that some synapses were potentiated with a coefficient η and some other depressed with a coefficient η^-. In the case of inhibitory neuron, we have introduced yet another parameter for the depression of the synapses from inhibitory neurons: $C^- = C \cdot \eta^-$.

A neuron does not have the same number of synapses connecting it to a subpopulation or to another. In fact, in the model, only a few synapses convey to a neuron spike trains from the same subpopulation. In order to describe this feature, we introduced another parameter, $f = \frac{n_{spe}}{n_{int}}$, where n_{spe} is the number of synapses conveying to the neuron the efferent spike trains of subpopulation $\mathcal{E}_{up}$ and subpopulation $\mathcal{E}_{down}$ and n_{int} is the number of synapses conveying internal input to the neuron.

The following parameters describe the connectivity of the network in more details, giving the values that we used for the parameters described above.

- The ratio of synapses conveying internal input is $x = 0.8$.
- The inhibitory synaptic strength is slightly stronger than excitatory. The multiplying coefficient is $C = 1.01$.
- The synaptic potentiation coefficient is $\eta = 9$.
- The ratio between the number of synapses from subpopulation $\mathcal{E}_{up}$ or $\mathcal{E}_{down}$ and the number of internal synapses, for *excitatory* neurons is $f(= 0.01)$.
- The ratio between the number of synapses from subpopulation $\mathcal{E}_{up}$ or $\mathcal{E}_{down}$ and the number of internal synapses, for *inhibitory* neurons is $f_i(= 0.05)$.
- The synaptic depression coefficient is $\eta^- = max(0, \frac{1-\eta \times f}{1-f})$.

- The inhibitory synaptic depression coefficient is $C^- = max(0, \frac{C*(1-\eta\times f_i)}{1-f_i})$.
- The ratio of external synapses conveying background activity is $f_b(=2/3)$.
- The ratio of external synapses conveying direction detectors output to the LIP neurons is $f_1 = f_2 = \frac{1-f_b}{2}$.

So, we can eventually describe the connectivity of the column by a matrix of 6 lines and 9 columns. We call this matrix W. For $j < 7$, $W_{i,j}$ is the relative strength of the synaptic input from subpopulation j to one neuron in subpopulation i: the fractions of synapses multiplied by corresponding potentiation, inhibitory, or depression coefficients. The subpopulations are numbered as follows:

- $\mathcal{E}_{up}$ is the subpopulation 1
- $\mathcal{E}_{down}$ the subpopulation 2
- I_{up} is the subpopulation 3
- I_{down} is the subpopulation 4
- I is the subpopulation 5
- $\mathcal{E}$ is the subpopulations 6

$$W = \begin{pmatrix} xf\eta & xf\eta^- & -xf_iC^- & -xf_i\eta C & -x(1-2f_i)C & x(1-2f)\eta^- & (1-x)f_b & (1-x)f_1 & 0 \\ xf\eta^- & xf\eta & -xf_i\eta C & -xf_iC^- & -x(1-2f_i)C & x(1-2f)\eta^- & (1-x)f_b & 0 & (1-x)f_2 \\ xf\eta & xf\eta^- & -xf_iC & -xf_iC & -x(1-2f_i)C & x(1-2f)\eta^- & (1-x)f_b & 0 & 0 \\ xf\eta^- & xf\eta & -xf_iC & -xf_iC & -x(1-2f_i)C & x(1-2f)\eta^- & (1-x)f_b & 0 & 0 \\ xf & xf & -xf_iC & -xf_iC & -x(1-2f_i)C & x(1-2f) & (1-x)f_b & 0 & 0 \\ xf & xf & -xf_iC & -xf_iC & -x(1-2f_i)C & x(1-2f) & (1-x)f_b & 0 & 0 \end{pmatrix} \quad (4.1.21)$$

Then in order to fully describe the synapses to a neuron we multiply this matrix by the Excitatory Post Synaptic Potential a and by the total number of synapses to a neuron N. The multiplying coefficient thus obtained is $J = N \cdot a = 11$. In this simulation project, we chose J arbitrarily, and since a and N do not explicitly enter the simulation equations we have not defined their value.

As seen in biology and in section 3.2.5 the spike trains emitted by neurons belonging to the same groups tend to be correlated. Therefore, we introduced correlation on our model. The correlation is described by the matrix c, where $c_{i,j}$ is the correlation coefficient between spike trains emitted from a neuron of population i and a neuron of population j. The correlation coefficient between spike trains is defined in section 2.2.9. We

suppose that the correlation between neurons that belong to the same subpopulation ($i = j$) is $c_{i,i} = 0.1$. The correlation between neurons of two different subpopulations is $c_{i,j} = 0.01$, $i, j < 7$. The correlation between inhibitory and excitatory neurons is null. So we obtain the following correlation matrix c (different from C, the synaptic inhibitory strength):

$$c = \begin{pmatrix} 0.1 & 0.01 & 0 & 0 & 0 & 0.01 & 0.01 & 0.01 & 0.01 \\ 0.01 & 0.1 & 0 & 0 & 0 & 0.01 & 0.01 & 0.01 & 0.01 \\ 0 & 0 & 0.1 & 0.01 & 0.01 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.01 & 0.1 & 0.01 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.01 & 0.01 & 0.1 & 0 & 0 & 0 & 0 \\ 0.01 & 0.01 & 0 & 0 & 0 & 0.1 & 0.01 & 0.01 & 0.01 \\ 0.01 & 0.01 & 0 & 0 & 0 & 0.01 & 0.1 & 0.01 & 0.01 \\ 0.01 & 0.01 & 0 & 0 & 0 & 0.01 & 0.01 & 0.1 & 0.01 \\ 0.01 & 0.01 & 0 & 0 & 0 & 0.01 & 0.01 & 0.01 & 0.1 \end{pmatrix} \tag{4.1.22}$$

For $i, j = 7, 8, 9$, $c_{i,j}$ represents the correlation coefficient for the inputs from respectively the background activity, the upwards detector and the downwards detector. One can see that the outputs of the detectors are correlated with a small correlation coefficient, for example $c_{8,9} = 0.01$, despite the fact that they detect motions with opposite directions. This is because the neurons involved are close to each other and because the firing of each detector will be marginally stimulated by the same stimulus, as long as it involves moving dots.

The background activity

The spike trains are all modelled as Poisson processes in order for us to be able to compute the neural equations presented above. However we want to have control over the two first moments of the low level background activity, the mean interspike interval and the variance of the interspike interval. In order to solve this apparent contradiction (in the Poisson process, the variance and the mean are equal) we use the following strategy. As we will see in the next subsection, the evolution of the system is studied by recalculating its variables along successive timestep, a standard method of numerical simulation. At each timestep t_n we assume that the spike trains that constitute the background activity are Poissonian. However, we redefine their rate each time t. At each timestep and for each neuron that receives background activity, we randomly draw a variable $F(t_n$ from the normal distribution $\mathcal{N}(0, \sigma)$, where σ, the standard deviation, is a parameter that we change in order to study its influence over the system (see the simulations section). Then it is added to the mean FR μ of the background activity, itself a varying parameter. The rate of the background activity Poisson process incoming to one neuron at one timestep t_n is then $\nu_{global}(t_n) = F(t_n) + \mu$. As described before, for example in section 3.2.1 we simplified many processes

coming to many synapses into one process of higher rate coming to one synapse whose weight is defined in matrix W, in Eq. 4.1.21.

The background activity of the brain can control the dynamics of the competition in many ways, by modifying the stability of the homogeneous and heterogeneous states of the system. We will see, in section 4.2, simulation results that highlight some features of this control.

System Equations

We use a discrete formalism. At each time step of one millisecond, we re-evaluate the variables of the system (the rates of the Poisson processes that model the spike trains generated by the neurons), as a function of their value at the previous time step, and of the value of the contextual variables. The contextual variables are the global activity of the brain and the direction detectors' FR. Using the LIF model presented in section 3.2.1, we obtain the following update rule for the FR of the various populations of the network.

$$\nu_{n+1}^{int} = F(\nu_n^{int}, \nu_n^{ext}) \tag{4.1.23}$$

where

$$F(\nu_n^{int}, \nu_n^{ext}) = \left[\tau_{ref} + \tau \int_{\frac{V_{rest}-\mu}{\sigma}}^{\frac{V_\theta-\mu}{\sigma}} \phi(u)du \right]^{-1}$$

where

- $\phi(u) = \sqrt{\pi} \cdot e^{u^2} \cdot (1 + erf(u))$
- $$\mu = \begin{pmatrix} \mu_{up} \\ \mu_{down} \\ \mu_{i,up} \\ \mu_{i,down} \\ \mu_i \\ \mu_0 \end{pmatrix} = J \cdot W \cdot \begin{pmatrix} \nu_{up} \\ \nu_{down} \\ \nu_{i,up} \\ \nu_{i,down} \\ \nu_i \\ \nu_0 \\ \nu_{global} \\ \nu_{ext,up} \\ \nu_{ext,down} \end{pmatrix} \tag{4.1.24}$$
- $$\sigma_n^2 = \frac{1}{2} J^2 \left(\sum_{i=1}^{9} W_{n,i}^2 \nu_i + \sum_{i,j=1, i \neq j}^{9} c_{i,j} W_{n,i} W_{n,j} \sqrt{\nu_i \nu_j} \right) \tag{4.1.25}$$
- $\tau_{ref} = 0.0005$: refractory period in seconds.

- $\tau = 0.05$: neuronal time constant in seconds.
- ν^{int} is the vector representing the output rate of the neurons in the column:

$$\nu^{int} = \begin{pmatrix} \nu_{up} \\ \nu_{down} \\ \nu_{i,up} \\ \nu_{i,down} \\ \nu_i \\ \nu_0 \end{pmatrix}$$

- ν^{ext} is the vector representing the output rates of the contextual neurons: background activity and MT neurons:

$$\nu^{ext} = \begin{pmatrix} \nu_{global} \\ \nu_{ext,up} \\ \nu_{ext,down} \end{pmatrix}$$

- W is the matrix of synaptic weights.
- $c_{i,j}$ is defined in the matrix presented in Eq. 4.1.22.

This structure induces the competition that leads to a decision. All the neurons receive synaptic inputs from the low-level activity of the brain, which we modelled with two parameters: FR and standard deviation. The effect of these two parameters on the dynamics of the whole decision model is the object of this study.

Recursive dynamics of the LIP column: the generation of a decision

As we can see in the matrix W of synaptic weights, the two subpopulations $\mathcal{E}_{up}$ and $\mathcal{E}_{down}$ that receive evidence from the direction detectors inhibit each other via specific inhibitory neurons, and excite themselves. If we only look at the mean postsynaptic input (μ as defined in Eq. 4.1.24), we see that

$$\begin{aligned} \mu_{up} = J \cdot (& xf\eta\mu_{up} + xf\eta^{-}\mu_{down} \\ & - xf_i C^{-}\nu_{i,up} - xf_i\eta C\nu_{i,down} - x(1-2f_i)C\nu_i \\ & + x(1-2f)\eta^{-}\nu_0 + (1-x)f_b\nu_{global} + (1-x)f_1\nu_{ext,up}) \end{aligned} \tag{4.1.26}$$

We see in this equation a few inputs that come from potentiated synaptic weights: recursive input from the same subpopulation $\mathcal{E}_{up}$, and from the inhibitory subpopulation I_{down}, the inhibitory neurons that receive stronger inputs from the subpopulation $\mathcal{E}_{down}$. If we write the same equation for $\nu_{i,down}$, we see that it

receives inputs from $\mathcal{E}_{down}$ with potentiated synapses. Consequently, by applying Eq. 4.1.24 repeatedly, we see that that subpopulation $\mathcal{E}_{up}$ receives strong inhibition from subpopulation I_{down} which itself receives strong excitation from $\mathcal{E}_{down}$. The exact symmetric reasoning can be held to show how $\mathcal{E}_{up}$ inhibits $\mathcal{E}_{down}$ via I_{up}.

$\mathcal{E}_{up}$ receives excitation from the upwards detection detector, whilst $\mathcal{E}_{down}$ receives excitation from the downwards detection detector. When there is no explicit motion direction in the kinematogram, the two subpopulations receive exactly symmetrical inputs, so the system is balanced in a homogeneous state: the activities of all neurons remain stable at a low level. (This level of activity can be tuned with the parameters of the network, results not shown here). In contrast, if one of the two subpopulations receives stronger evidence because the kinematogram shows a significant number of dots moving in one direction, then the equilibrium is broken. This subpopulation will apply a stronger inhibition to the opposite one, which will thus reach a lower activity. Consequently, the latter will apply a weaker inhibition to the former, which will thus be able to sustain a higher activity. This recursive pattern will lead the column in a heterogeneous state in which one subpopulation will sustain a high activity, dominating the other one into near zero activity. Once the difference of activities is important enough, removing the stimulus will not bring the system back to the balanced homogeneous state: the model is committed to a decision and can remember it.

In Fig. 4.4 we illustrate the time course of a decision making. We show the competing FRs of the two evidence gathering subpopulations, and the integration of the difference of their activity over time. This integrating process is defined later in section 4.1.5. It represent the activity of neurons in the Superior Colliculus (SC), preparing for the performance of an eye saccade. We see that the recurrent loops enable the model to keep in mind the decision even when the stimulus disappears and we artificially reset the integrative SC activity to 0. In fact this decision is not expressed by the model until the difference of activity between the two populations has been consistent enough for the eye to have accomplished its saccadic movement after the SC activity reached the threshold.

4.1.5 Expressing the decision with a Saccadic Eye Movement

In the benchmark experiments presented in the introduction chapter, we have seen that the monkeys and the humans who perform the discrimination experiment are asked to express their decisions with a Saccadic Eye Movement (SEM). This experimental set up has been used to evaluate the decisions of monkeys during the same discrimination task, for example in Shadlen and Newsome (1996). Despite the fact that modelling the biophysics of the generation of a saccade turned out to be beyond the scope of this thesis, in order to have an embodied measure of the Reaction Time (RT), we needed to attempt at modelling at least the neural activity leading to the generation of a saccade. Eye saccades can be seen as stereotypical patterns of motor

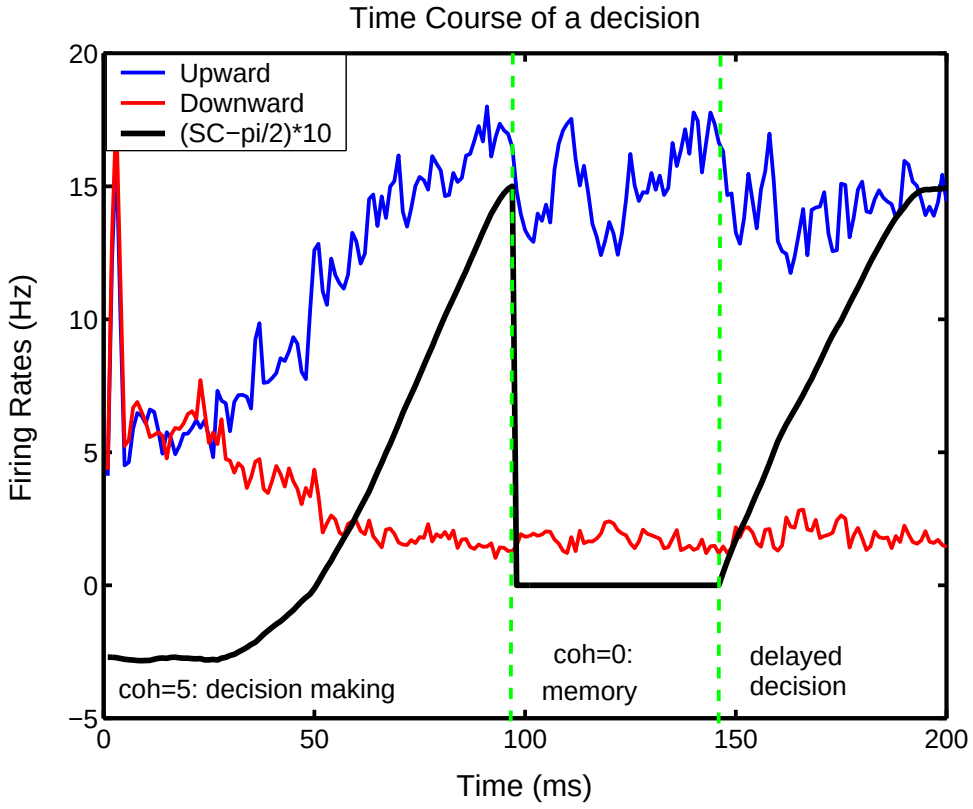


Figure 4.4: Illustration of a decision making and working memory, under the influence of random background activity. The stimulus is presented during the first 100 ms, with a coherence of 5 percent ($coh = 5\%$). We represented the activity of the Superior Colliculus (SC), defined in section 4.1.5, but normalised so that its curve would compare to the FRs of the populations accumulating evidence. In this figure, we represented $FR(SC)' = 10(FR(SC) - \frac{\pi}{2})$, where FR(SC) is the firing rate of SC neurons. Then, the activity of the Superior Colliculus (SC), defined in section 4.1.5. is reset to the middle position and the stimulus disappears (100 ms to 150 ms). However, we see that the decision is maintained when we free the SC activity and the system expresses a delayed decision, even in the absence of stimulus.

control activity and eye movement. In the literature, we find evidence that this saccade is generated when the activities of neurons in the Superior Colliculus reach a threshold. Therefore, we modelled the activity of SC neurons and assumed that a saccade was generated upon the modelled SC activity reaching a threshold value. The following equation describes the SC activity:

$$\frac{dx}{dt} = k \cdot R - l \cdot tan(x - \frac{\pi}{2}) \tag{4.1.27}$$

- x represents the SC activity, varying from 0 to π.
- k and l are parameters. In the simulations of this chapter, $k = l = 3$.
- R is the input that comes from the LIP column to the SC model. $R = \nu_{up} - \nu_{down}$.
- $x(0) = x_0$ is chosen randomly around $\frac{\pi}{2}$

The system defined by this equation can be seen as an integrator that is similar to the diffusion model of decision making. As seen in Usher and McClelland (2001), instead of considering two populations competing to reach a threshold, we used one variable representing the difference of activities. The term $tan(x - \frac{\pi}{2})$ models a leak that tends to bring the SC activity in its central position if no specific decision is generated in SC, and that also prevents the SC activity from diverging. In fact, when x tends to π or 0, the leak modelled by $tan(x - \frac{\pi}{2})$ tends to ∞, thus the activity is limited. We consider that a saccade is generated when $x = \frac{\pi}{2} - 0.95\frac{\pi}{2}$ or $x = \frac{\pi}{2} + 0.95\frac{\pi}{2}$. This is an arbitrary value, as the parameters k and l are. It would be possible to fit these values to biophysical data, but the qualitative results would not be changed. With this model we can develop a measure of Reaction Time: In this study, this is the time between stimulus onset and initiation of a saccade. To compare these results to experiments, we would need to add to our RT the typical time necessary to perform a saccade, plus the preprocessing time of the early visual cortex, that is not modelled in the simulations of this chapter.

4.1.6 Methods

The model was considered to have taken a decision when an SEM was generated. However, if after a certain time limit (usually, 0.3 seconds) no decision was taken, then we interrupted the experiment. This saved computational time and took into account that the model might never reach a decision. The model was presented with the same kinematograms as in the previous chapter. Every ten milliseconds, the direction of each dot in the kinematograms were re-generated and the output FRs of the direction detectors were re-evaluated. These FR were used to numerically simulate Eq. 4.1.23, with a time step for integration of one millisecond. At each time step and for each population, a value of the background noise was randomly generated from the Gaussian distribution of the specified mean and variance.

For each simulation, we obtained a Reaction Time and a decision that was either correct or erroneous. After 100 simulations with the same parameters we obtained a mean RT and an Error Rate (ER) (Number of errors divided by the total number of answers). All simulations were conducted with stimuli that had 5 percent coherence.

We repeated these simulations for various values of mean and variance of the background activity, in order to highlight the influence of these two parameters on the dynamics and quality of decision making. We covered the following parameter space : The mean of the global background activity varied from 3 to 10 Hz, and the standard deviation varied from 0 to 5 Hz. The results are presented in the next subsection.

4.2 Influence of low level background activity

4.2.1 Mean of background activity controls the speed/accuracy trade-off

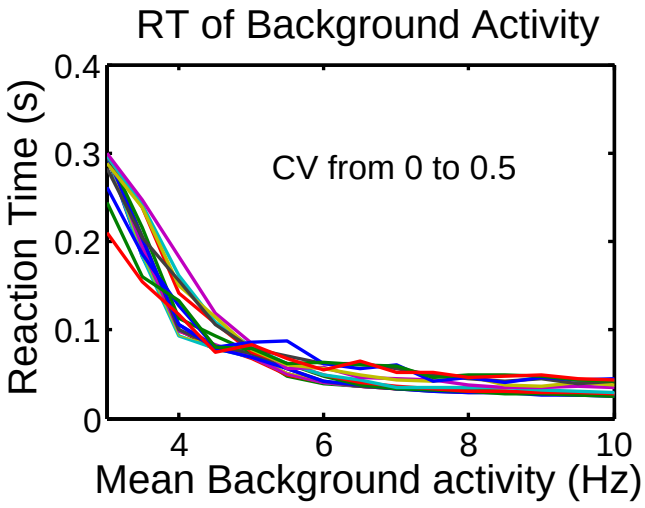


Figure 4.5: Reaction Time as a function of the mean intensity of the background activity of the brain. The Decision processes were stopped after 0.3 seconds, in order to reduce computational time. In fact, if no decision is taken, when the background activity is too weak, RT tends to infinity.

In Fig. 4.5, we clearly see that increasing the intensity of the background activity reduces the Reaction Time. For a mean background activity inferior to 3Hz, the RT is very long, quickly tending to infinity when the activity reduces. The decision cannot be taken because none of the two competing subpopulation manages

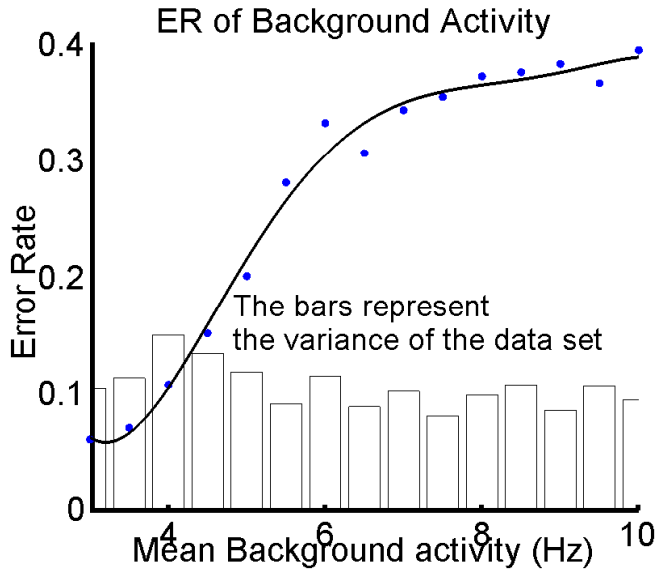


Figure 4.6: Error Rate as a function of the mean intensity of the background activity of the brain. This is an average over all the values of the Standard Deviation, between 0 and 5.

to take over, and both keep a low level activity. This is in accordance with the 'switch' effect described by Salinas (2003), because within a few Hertz of background activity the system is stable at homogeneous activity or tends to develop blobs of higher activity that inhibit each other. This is however richer than a switch effect because the change is continuous, the background activity controlling how fast a dominant population emerges. If it emerges infinitely slowly as happens under 3 Hz, then no decision can be taken. In parallel to the Reaction Time, the Error Rate increases with the intensity of the background activity following a sigmoid function. In Fig. 4.6, we can see a 4 parameters exponential approximation of this increase.

4.2.2 Variance of the background activity

Fig. 4.7 shows that the variance of the background activity also speeds up decision making. In this simulation, the mean of the low level background activity is constant. Its standard deviation is the only parameter that varies. The standard deviation of the background activity is what characterises its noisy essence. This is a new result, purely based on the second order moment of the brain stochastic neural activity. Fig. 4.8 shows a significant influence of noise on the Error Rate.

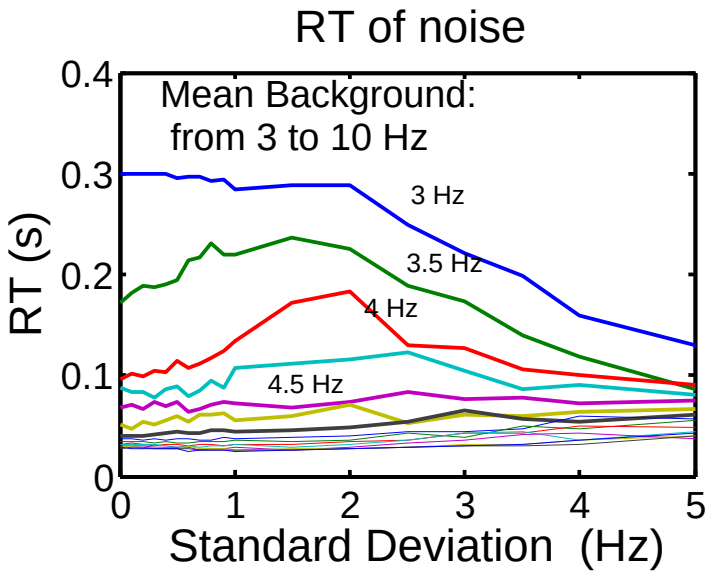


Figure 4.7: Reaction Time as a function of the Standard Deviation of the Background Activity. The mean activity varies between 3Hz and 10Hz.

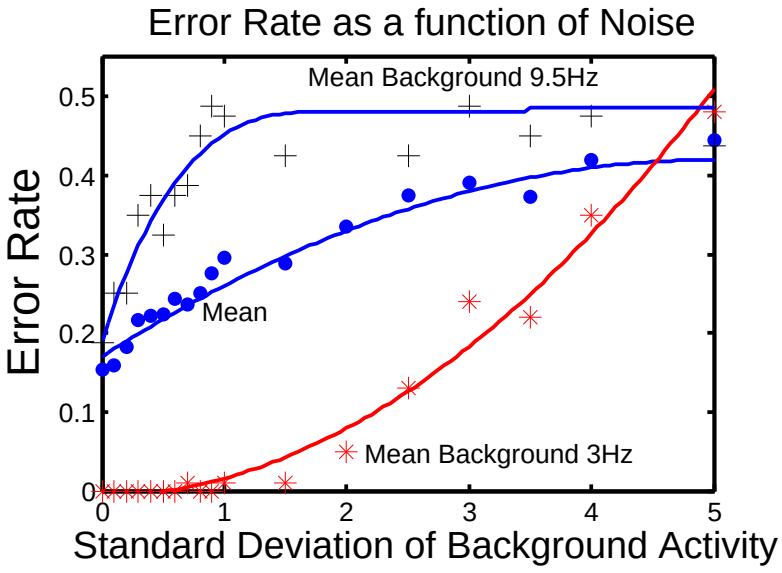


Figure 4.8: Error Rate as a function of the Standard Deviation of the background activity of the brain. It shows three different behaviours for three different mean intensities of the background activity.

4.2.3 Summary and prediction

Increasing the mean intensity of the background activity of the brain reduces the Reaction Time whilst increasing the Error Rate. Increasing the variance also increases the Error Rate, and can decrease Reaction Time, depending on the mean background activity. So, we have shown in this chapter that this background noise is not necessarily only a switch to a decision making attitude of the brain, as shown by Salinas (2003), but that it can precisely tune the speed of decision making and even reduce the Error Rate. The assumptions made in this paper fit well with the current literature on decision making as seen in reviews such as Glimcher (2003b), or in other models such as the one described in by Wang (2002).

We used two measures (RT and ER) that are not defined by neural activity but by external movements of the tested animal or human. We have evaluated the dependence of these two measure on what we call background noise, the statistical signature of the background low level activity. More precisely, if μ is the mean FR of the background activity, we have measured the functional dependencies $ER(\mu)$ and $RT(\mu)$. From these measures, we can then evaluate the claim that μ is the neural translation of urgency. In that case RT must be a decreasing function of μ. This is the case of the results that we have presented. For a level of noise μ_0, the corresponding Reaction Time will be RT_0 and the corresponding Error Rate will be ER_0. RT_0 is associated to one unique ER_0 thus we were able to define and evaluate the function $ER(RT)$ with our simulations results.

This function is a prediction produced by our model, and since it only involves external variables such as the reaction time and the error rate, it could be compared with data obtained through non-pervasive biophysical experiments, in order to test our hypothesis. Experimental research is beyond the scope of this thesis, but we have written the qualitative prediction that our simulations suggest. Our model predicts that if the mean intensity of the background activity controls the trade-off between RT and ER, then the experimental results should be qualitatively compatible with the following equation:

$$ER = 0.055 + 0.6 \cdot e^{-18RT}$$

This prediction is described in Fig. 4.9, where we see that it corresponds to the intuitive result that the faster the decision has to be, the less accurate it will be. This characterises the trade-off between time and accuracy.

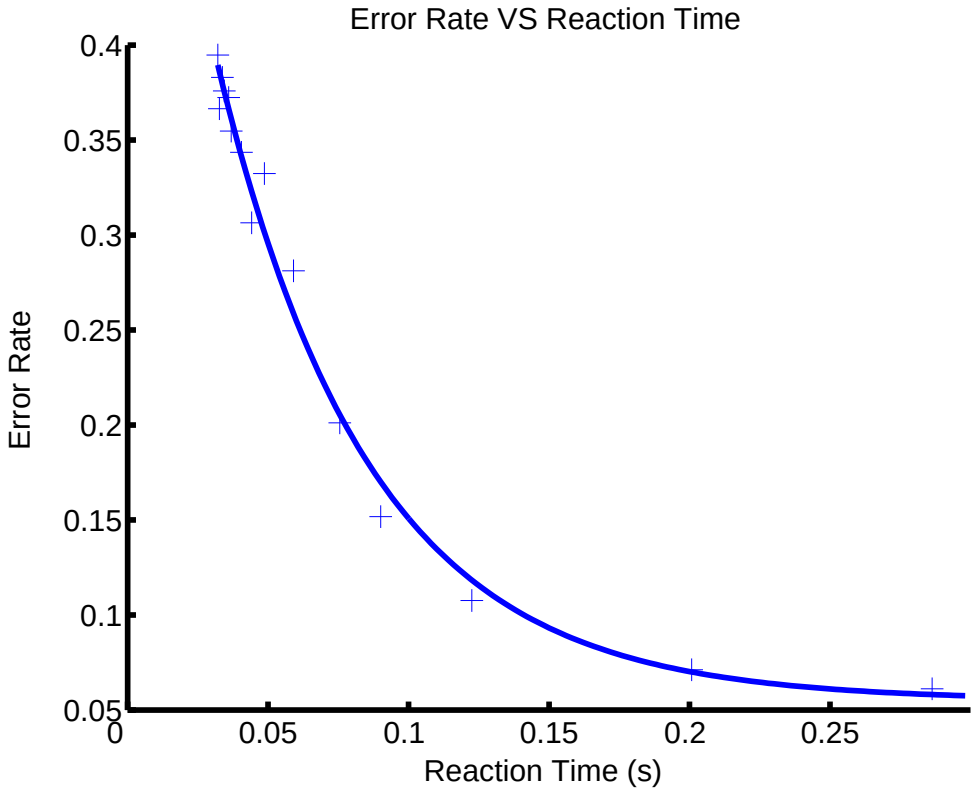


Figure 4.9: Error Rate as a function of the Reaction Time and the corresponding exponential fit. Our method to obtain this curve is a change of variables from the mean background activity to the Reaction Time. This works, because we have seen previously that the RT is a monotonic function of the background Intensity. Here, we have taken the mean over all the values of the standard deviation of the background activity. The curve is an exponential fit of the difference between the ER and its minimum, 0.055*, with two parameters:* $a = 0.6$*,* $b = -18$*. The standard error of the regression (Root Mean Squared Errors of the fit) is* $RMSE = 0.014$*.*

4.3 Conclusion

The statistical signature of the neural background activity influences the dynamics of decision making. We provided in this chapter an evaluation of the graded control exerted by random neural activity over a biologically inspired model of decision making. This goes further than the model of Wang (2002), because the role of the low level background activity goes beyond forcing a decision in case of balanced or absence of stimuli, and it goes beyond the investigation of Salinas and Sejnowski (2001) who only saw it as a switch between dynamical states.

Our simulations showed that the RT of a decision clearly decreases with the mean background activity, whilst the ER increases. The ER also clearly increases with the variance of the background activity but the decrease of the RT with the variance is less clear. This investigation also suggests a method for asserting the validity of the hypothesis according to which the background activity is the neural expression of the feeling of urgency of the subject. However, the results of the model as it is can not be compared with the results of biophysical experiments. We would need a more detailed modelling of the visual cortex and of the saccade generation. Furthermore, the parameters of the models should be chosen more akin to biophysical parameters. So, this chapter shows that our hypothesis is possible, but that it needs to be more grounded in biology to be able to be tested.

Some modelling questions remain. Since we have assumed that the neurons communicate with Poisson processes, we have neglected second order statistics that might have an important role in the convergence to the stable states that reflect a decision. We have not modelled either how the background activity arises and varies. As we have seen in the end of Chapter 2, the background activity should probably be modelled in an embedded manner, taking into account, if not the whole organism at least the whole cortex and its interactions with other brain structures.

So, in the next and last chapter, we develop these few questions and directions. First we see how far we can compare our model to experimental results, then we suggest how we could develop the model to include second order statistics and how we could model the low level background activity, before concluding by evaluating what this investigation tells us about the mind and the body.

Chapter 5

Discussion

5.1 Summary of contributions

In this thesis, we have shown that randomness improves perception tasks modelled by populations of Leaky Integrate and Fire neurons. We have focused on two types of randomness in neural activity. First, we have studied what we called input noise, the variance of the modelled post-synaptic input. We have shown that increasing the ratio between inhibitory and excitatory inputs increases the input noise to a Leaky Integrate and Fire model of a neuron. We have applied our Leaky Integrate and Fire model to the simple task of discriminating the general direction of kinematograms. This application mimicked experiments that have been fundamental in showing that neurons in area MT perform this task. Thus, we applied to our discriminating neurons an input that models the output of the early visual system, after using simple modelling attempts to justify assumptions regarding this pre-processing. Simulations of this model of discrimination in MT showed that the accuracy improves with the input noise. We also showed that population coding increases the speed and accuracy of the discrimination task.

Secondly, we have seen that neurons in LIP are central to visual decision making based on kinematogram evidence and expressed by an eye saccade. It is established in the literature that these neurons accumulate sensory evidence until a decision is taken. Furthermore, it has been shown that the enhanced activity of neurons in favour of one decision hinders the activity of neurons in favour of the opposite decision. This is typical of a competition system. Chapter 4 has provided a model that accounts for a mixed process of neural integration and neural competition, and has done so solely in terms of a simple and well established neural model, the Leaky Integrate and Fire neuron and its synaptic connections.

In Chapter 4, we have also shown that random neural activity speeds up the accumulation of evidence occurring in the area LIP. In that chapter, what is referred to as 'random neural activity', or 'noise' is the background neural activity surrounding the competing neural populations. We modelled this background activity with its mean and variance. Furthermore, we used our model to evaluate the hypothesis that noise actually controls the dynamics of decision making, creating a trade-off between Reaction Time and Error Rate. Since we have established, on one hand, a monotonic functional relation between the mean background activity of the brain and the Reaction Time, and, on the other hand, between the same mean background activity and the Error Rate, we were able to predict an analytical relation between Reaction Time and Error Rate. This prediction will, of course, not be exact, due to the various approximations such as the simple model of the early visual cortex or the saccade generation. In order to have a satisfying comparison with psychophysical data, we can add a time constant representing non-modelled processing: In Palmer et al. (2005), they compare experimental data to a simple diffusion model. They add a constant, called mean residual time, to the model's Reaction Time, and fit this constant to match data without losing the meaning of their model's predictions.

5.2 Comparison

5.2.1 The direction detectors

As argued in Wang (2002), in Britten et al. (1993) and in Palmer et al. (2005), the response rate of MT neurons is a linear relation of the stimulus strength. This thesis has been focusing on a benchmark discrimination task based on kinematograms. In this case, the measure of stimulus strength is the coherence of the kinematogram. Fig. 4.3 shows the response of our modelled direction detectors in area MT. It is clearly a linear relation. Thus, the evidence upon which our LIP model bases its decision is coherent with experimental data.

5.2.2 Commitment to a decision

The classical modelling of decision making is now based on accumulation of evidence via neural integrators as in Huk and Shadlen (2005), or more abstractly through diffusion models and their extensions as seen in Palmer et al. (2005). These models involve a very low level of commitment: Until the threshold of activity is reached the decision toward which the model was drifting can be changed in the case of input reversal. The only commitment lies in the fact that if the neural activity that represents accumulation of evidence is close to the decision threshold then there still is a chance that through random perturbation it will cross this

threshold after stimulus reversal. However, the drift itself changes sign when the stimulus reverses. On the other hand, in Wang (2002) and in our model, there is a high level of commitment to a decision, because of the mutual inhibition and the recurrent auto-excitation. When one population inhibits the other with its self-sustained higher activity, this dynamic interaction becomes more prominent than the actual stimulus, and the trend toward a decision is very unlikely to be altered by stimulus reversal.

In Fig. 1.6, we see a recording of neural activity in area LIP during the decision task that we focus on. As pointed out by the authors, we can see that after stimulus onset, the Firing Rate's increase rate grows with the stimulus strength. However it cannot be argued that this ramp of activity is the results of the integration of evidence constituted as a neural activity linearly related to the coherence of the stimulus. We can distinguish two phases in the accumulation process: a steep initial increase followed by a slower ramp. This recording is evidence in favour of a non-linear process happening at the same time as the summation of evidence, and thus supports our model with recurrent inhibition and excitation as opposed to temporal linear integration.

5.2.3 Effect of urgency

We are looking for data that describes the speed accuracy trade-off under urgency instructions. Recent and authoritative research about decision making, Reaction Time (RT) and Error Rate (ER) such as Palmer et al. (2005) refer to psychological studies made in Ruthruff (1996), dedicated to compare models to experimental data concerning the speed accuracy trade-off. They only have two instructions: stress on accuracy and stress on speed. This data is not enough for us to compare with our continuous prediction. In fact most of the studies of the effect of urgency on decision making are only concerned with RT, as in Rao (2004), Reddi and Carpenter (2000) or Reddi et al. (2003). The joint study of RT and ER is more developed in terms of the influence of stimulus strength such as in Palmer et al. (2005). In Ratcliff and Rouder (1998), they also manipulate the speed-accuracy threshold instructions with only two levels of urgency. Furthermore, we cannot really compare our model with their experiments, as their task is about brightness and not about kinematograms. Again, in Huk and Shadlen (2005) we find discussion about the speed-accuracy trade-off, but the influence of urgency is only considered in so far as they discourage the monkey to produce fast erroneous answers. They refer to Palmer et al. (2005) and Reddi and Carpenter (2000) for control over the speed-accuracy trade-off.

Here we attempt to compare our results with these of Palmer et al. (2005). They give data relative to three instructions for the urgency: fast normal and slow. This comparison is interesting because they simultaneously consider the variations of RT and ER using the same benchmark decision task as the one that we studied in this thesis. They mainly study the effect of varying stimulus strength (kinematogram coherence) on these two performance measures, but in one experiment they present measures of performance variations

resulting from external urgency requirements. In order to confirm our hypothesis that this urgency is mediated via low level background activity, our results should accord with theirs. From their paper we extracted the data that we needed and traced the plot of ER as a function of RT in the case of urgency constraints. This is shown in Fig. 5.1.

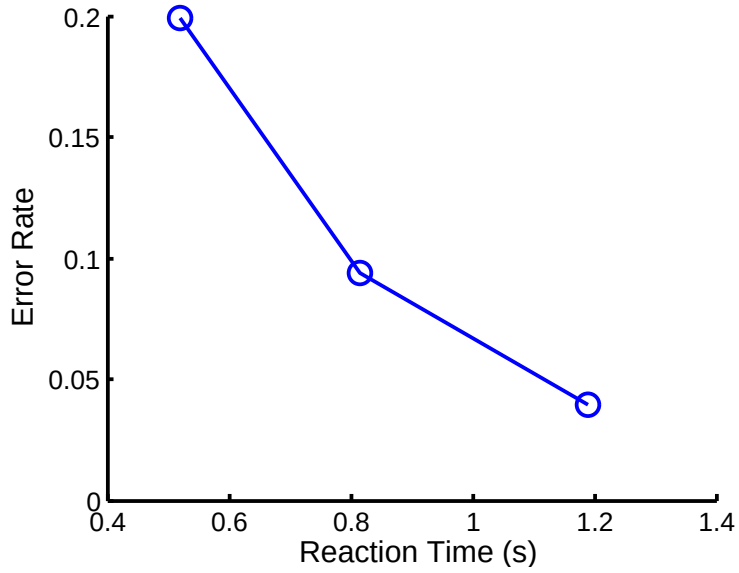


Figure 5.1: Error Rate as a function of the Reaction Time, measured on humans. The values have been extracted from Palmer et al. (2005), experiment 2, with a coherence of 5 percent. In the paper they fit their data with a diffusion model that gives them a function of RT and ER vs. stimulus strength (kinematogram's coherence), for three levels of urgency. Their data is precisely fitted. To obtain the values presented here, we computed RT and ER as a function of the coherence (5%) with the parameter of the diffusion models that best fitted their data.

Comparing Fig. 4.9 to Fig. 5.1, we see that our model is consistent with their results in so far as we have a non-linear Error Rate decrease with the increase of Reaction Time. However, we do not have enough biological data to precisely compare their results to our model. We should point out that, since we did not simulate the actual saccade or the early visual processing in details, our modelled reaction is delayed from psychophysical measures. A quantitative comparison would require adding processing time to the Reaction Time defined in this thesis.

5.3 Limitations and developments

5.3.1 Limitations

The Poisson assumption is simple but, as seen in Feng et al. (2006), it does not hold very well in the case of the Integrate and Fire model. Furthermore, we are studying the effects of random neural activity on the dynamics of decision making. So, it is important to keep the richness of behaviour generated by second order statistics. It can only be achieved if the mean and variance of the ISIs vary independently in the whole model, not only in the background activity as we have done so far. In 5.3.2, we present a more complex model that allows us to study the evolution over time of the two first statistical moments of spike trains.

So far, we have assumed that the neurons we modelled were embedded in a global low level background activity coming from other neurons surrounding them or possibly distributed across the brain. We have artificially simulated this low level background activity and described it directly with its two first statistical moments. The existence of this background activity is well known amongst neuroscientists. The question remains to know whether it is possible to model such a stable low level activity in a network of neurons. More precisely, we wanted to know what features are necessary for a network of LIF neurons to exhibit a stable low level activity, and what parameters of the network are influential to its statistical signature.

A network exhibiting such a stable a low activity has been modelled with statistical approximations in Amit and Brunel (1997), and with spiking neurons in Izhikevich et al. (2004) and Izhikevich (2006). However, the neuron model used by Izhikevich is not an IF model. Moreover in many such models the authors assume yet another external input that stabilises the network. So, the last series of simulations that we present in this thesis, in section 5.3.3, is an attempt at answering these two questions: How can stable low level activity emerge from the properties of a network of LIF neurons whose every spike is modelled, without external input, and what parameters in the network might be susceptible to control the statistical signature of this activity.

5.3.2 Overcoming the Poisson assumption with moment mapping

The Moment Neuronal Networks (MNN) for Integrate and Fire neurons have been developed by Feng et al. (2006) for feedforward network. We described in details the ideas behind them in 2.2.1. In order to use that theory in this development of our decision making model, we had to adapt this feedforward approach to a recurrent situation. The detailed simulations and results are presented in Appendix A, but the next paragraph gives a summary of the partial results we obtained.

Having rigorously modelled the propagation of second order statistics in the network of LIF neurons that simulates decision making, we have shown that the variance of the background activity can control the speed of decision making independently of the value of the mean background activity. This is a result that could not be shown properly with the Poisson approximation. For this model, we used a slightly different structure and different parameters, so we cannot compare the two models. This emphasizes that there are many parameters that we can alter in such networks, and that exploring the whole parameter space with simulations is not a realistic strategy. In order to understand such complex systems, we would need to use a new formalism based on mathematical descriptions issued from non-linear dynamical systems theory, in concordance with detailed statistical analysis. In this research we have mainly used simulations, but we have hinted at such analytical pioneering research, for example in Berglund and Gentz (2002).

5.3.3 Can low level background activity emerge from network properties?

Topology

In the work presented so far in this dissertation, we have neglected to look at the distance between neurons, and the effects that this distance has on neural coding. We only took advantage of the natural distance between neurons when, in the model of the early visual system 3.1.1, we tried to implement the winner takes all strategy, whereby the neuron whose receptive field was the closest, in terms of physical distance, form the input would fire first and prevent its neighbours from firing through inhibitory lateral connection. This was an interesting but limited approach to the use of distances, because we did not use any feedback and because we looked partially at the effects of the distances between some neurons.

How far two neurons are influences how strongly connected they are and the time needed for a spike to propagate from one neuron to another, modifying the speed of their communication, but these effects do not follow a single simple rule. Some distant neurons can be strongly connected, and some axons carry spike faster than others. The existence in biological neurons of connection delays has been outlined previously in this thesis, and has recently received renewed interest by for example Izhikevich (2006). Furthermore, Defining a distance between neurons is an interesting problem. Since the cortex is a folded layer of neurons surrounding what can hardly be described as a sphere, one has to wonder whether classical three dimensional Euclidian distance is the most relevant one.

So, one issue that remained after the work presented so far was to know whether, without external simulation a brain could generate stable random low-level activity. We tried to answer this question by modelling a network of Integrate and Fire neurons organised and connected according to a simple topology. This

topology enabled us to take into account that spikes need time to propagate, and that neurons are often less strongly connected when they are distant from each other.

Initially, we wanted to get enough control on this model to try to reproduce the decision making dynamics studied in the previous chapter, but it proved to be beyond the scope of this thesis. So the detailed model and simulation results are presented in Appendix B.

Results

With that simple model, we attempted to stabilise a network of Leaky Integrate and Fire neurons without the use of an external stabilising stochastic activity. Izhikevich and Edelman, in Izhikevich et al. (2004) and Izhikevich (2006) did not do it, they used random thalamic input. We wanted to see the network stabilise at a low level activity: the network should be stable in state where it is not completely silent, but where the neurons are far from all firing at their maximum firing rate, defined in the model as the inverse of the refractory period. In this work, this stable low level activity is called Submaximal Stable Activity (SSA).

We also searched for a network parameter that could control the value of the SSA of the network. We have managed to stabilise the network at a submaximal activity without external input, but this SSA is much higher than that observed in biological neural networks, and we do not have control over its level. However, by adding thalamic input to the system, as is often done in the literature, we can perfectly control the SSA.

By looking at the raster plots presented in Appendix B, one can see regular patterns of firing, but we think that this is because the population is too small. With larger neural populations, one would get more intricate patterns, and thus local firing would be more irregular. Beyond a sufficient level of complexity, this irregularity may be what is currently considered as intrinsic noise in the brain. Furthermore we suggest that larger networks would enable us to control the level of SSA with the radius of the sphere. This hypothesis could not be tested because larger network of neurons required too much computational power, but it is a very interesting line of future investigation.

5.4 Evaluation and directions

5.4.1 Strength and weaknesses of the models

In this thesis, after reviewing the neurology of decision making, we proposed a model that produces evidence in the case of a benchmark decision making case. This decision making task involving kine-

matograms has been extensively experimented upon, and a wealth of neurological and biophysical data has been produced. Following the literature, we situated our evidence producing model in area MT, or MST. Then we linked it to a larger model of decision making. Again, we followed the literature by situating our model in area LIP and investigating the integration of evidence by neural variables.

This model, connecting two parts, is relevant because it is consistent with the most up to date and widely accepted models and neurological data about this decision making task. Together, the two parts described in Chapter 3 and 4 reproduce a range of observed phenomena such as working memory, increased errors and longer decisions in case of mixed stimuli, speed/accuracy trade off, fluctuating evidence as opposed to robust decisions.

It is an original work because it bridges approaches. Firstly it links the detailed low level approach of the Integrate and Fire neuron to the higher level cognitive processes such as decision making, attention, urgency and error rates that belong more to the field of psychology. Secondly, contrary to the general intuition we did not consider variance as a detrimental addition to the signal, to be edited out for better overall performance, but we investigate the beneficial effects of synaptic noise in Chapter 3 and of random low level background activity in Chapter 4. Finally, it puts decision making into perspective, and looks for its place in the global picture of the mind, the brain and the body. We have formulated and simulated a hypothesis about how this precise cognitive phenomenon is controlled by the global organism, through low level background activity.

The main drawback of our model is that our simulation results can not be compared with experimental results, for two reasons. Firstly, the model lacks biologically relevant modelling of the low level sensory and motor systems, which would have been very interesting, especially because we claim to take an embodied approach. We would need to provide a model of the treatment of the stimulus in the visual cortex, from the retina to area MT. We would also need to realistically model the generation of saccade. At the moment the picture we provide starts at the area MT receiving 100 directions of moving dots coded in Poissonian spike trains and stops at the FR of the superior colliculus reaching a threshold. Secondly, we have not consistently defined the parameters of the model according to biological values. We wanted to illustrate some neural dynamics and in many cases we chose ad hoc parameters. We believe that one could build similar models that respect measured values for FRs, synaptic weights, number of neurons, synapses and so on, even if this was beyond the scope of our work.

We identified that modelling neural communication with Poissonian spike trains when we wanted to study the effect of the second order statistics on the system was a simplification detrimental to the results of the

simulations. We investigated a third model that used Moment Neural Networks (MNN) and kept track of both the first and the second order statistics of the spike trains. We can see in Appendix A that we obtained results that are not consistent with those of Chapter 4. These result do not challenge that there is a graded influence of background activity on the decision making modelled as a competitive system with lateral inhibition, but the graded effect is different. This is an issue that is not solved yet in this research, and that prevents us from making claims on the nature of the control exerted by low level background activity on the dynamics of the competition.

All in all this research proposes a hypothesis that could help understand the embodiment of rationality and motivation and investigates its various aspects. We investigated if the notion of motives from Merleau Ponty, and the notion of somatic states from Damasio could not be implemented in the statistical signature of low level background activity. This if verified, would be an answer to challenging questions such as the neural basis of attention. In order to evaluate the plausibility of this proposal, we investigate the role of the background activity on a specific cognitive task. Our work shows that the idea is appealing and can work in models that are consistent both with detailed neural biology and higher level psychological features. One now needs to test it against biophysical and neurological evidence. For that one would need to build a model that would be more biologically relevant, more developed and more precisely tuned to specific experiments.

5.4.2 Towards an integrated model

A global system

The simulations presented in the previous section are fundamental because they tackle the basis of a general theory of the cortex as a substrate for psychology. This theory sees specific cortical activity as local divergences from a baseline stochastic low level global activity, sustained by the intrinsic characteristic of the neural network, as seen in Appendix B. We have seen that, in the ideal case of an homogeneous cortex, this activity would be stable. Specific divergences would be made possible by local parameter differentiations, such as potentiated synaptic weights. With this structure, local groups of neurons would, upon specific situations or task requirement, diverge from the global pattern of activity into a local, specific and transient pattern. For example, in our thesis, we have shown a divergence of groups of neurons of LIP from low level activity towards a higher or lower Firing Rate. The development of these local patterns of activity would be controlled by the statistical signature of the background activity, to which the local neurons are submitted, as exposed in Chapter 4 for the example of decision making. This hypothesis of background activity as a switch has also been formally studied by Salinas (2003). Within the scope of our research we have dissociated the attractor system we looked at from the rest of the brain activity, but it is all part of the same global system, because the local attractors influence the global statistical signature. Such ideas are becoming widespread, even in the domain of this specific decision

making task, as we can see in Sugrue et al. (2005). In this work, we have mainly focused on the FR and the variance of the ISIs to describe the state of the system. Other powerful characterisations of the state of the system can be used, such as the correlation between spike trains, synchronicity or even polychronicity.

Learning

These local parameters that control the attractor landscape of the brain as a dynamical system could, for instance, be the potentiation coefficient of synapses. We have seen in Chapter 2 that one of the most famous biologically relevant learning rule for neural systems is the Hebbian rule, which potentiates or depresses the synaptic weights according to the correlation of activity between the pre-and post-synaptic neurons. The invaluable interest of this learning rule is that it is a local rule that can have global effects. In this thesis, we have not simulated any kind of learning, but we have defined specifically potentiated synaptic connections that could potentially be developed using the Hebbian rule. This would mean that the competing populations that we have been focusing on would not be fixed but constantly evolving, strengthening their connections and expanding each time their neurons are involved in the specific decision task. By expanding, we mean that nearby neurons would tend to fire in a correlated manner with the group and thus would be drawn into the pattern of activity of the local cluster via Hebbian learning. That would explain the effects of training and specialisation of individuals that has been experimentally shown, whereby more neurons get involved in performing a task that is extensively trained. These ideas are consistent with the theory of neural Darwinism, by Edelman (1987), and their practical implementations have started to be investigated for example by Izhikevich (2006).

5.5 Conclusion

In this thesis we have developed a biologically inspired model of decision making based on evidence gathered from visual stimuli. We have modelled the system from early visual processing to the neural preparation of the eye saccade that expresses the decision. Nevertheless, the modelled decision making process is not fully determined by the stimulus, as it exhibits dynamics of its own that are determined by the internal connectivity of the network and by stimulus independent surrounding neural activity. Consequently, we have a model that shows decision making as neither fully dependent on stimuli, nor totally abstracted from the sensory inputs: it answers the questions raised in section 1.3.5 that were paradoxical in a dualist approach to mind and body.

In addition, we formulated a model inspired by the notion of motives introduced by Merleau-Ponty to explain perception in terms that solve the problems posed by dualism. It is, of course, a very simplified

model. How could we hope to have a realistic model of intentionality that relies on two parameters (the mean and variance of the low level background activity of the brain)? This work suggests that even our simplest decisions are not only dependent on stimulus based reasoning (modelled here as the MT-LIP-SC chain) but also, for a large part, are controlled by our inner mental life, depicted here as the low level background activity, an approximation for the motive. Even if it is only a small first step towards a global understanding of the physiological basis of our psychology, it already shows that it is possible to control the speed-accuracy trade-off with a widely distributed neural activity that is dependent on global parameters such as the topology of the system, as seen in the previous subsection. This essential part of decision making will never be totally explicable, because it is informed in each individual by an infinity of factors such as the shape of their body shape, their current situation, their education and culture.

In the case of ambiguous stimuli, when arguments in favour of one decision are balanced with opposite arguments, the decision outcome mainly depends on the low level background activity. If this background activity is considered as noise, than the decision is deemed arbitrary. On the contrary, if we see it as the implementation of motives, or inner mental life, the decision becomes an expression of our free will, of our humanity: What differentiates us from machines is that our history and personality is inscribed in our body and our brain as specific and rich patterns of global neural activity. In the western scientific world, this part of our humanity has long been stifled under the dogma of logic, rationality and scientism. This piece of research is an attempt at scientifically investigating this side of our humanity that is neither mechanistic nor religious.

Bibliography

Abbot, L. and Van Vreeswijk, C. (1993). Asynchronous states in networks of pulsed-coupled oscillators. *Physical Rev. E*, 48:1483–1490.

Abeles, M. (1991). *Corticonics: Neural Circuits of the Cerebral Cortex*. Cambridge University Press, Cambridge, UK.

Abeles, M. (2002). Synfire chains. In Arbib, M., editor, *The Handbook of Brain Theory and Neural Networks*, pages 1143–1146. MIT Press.

Amit, D. J. and Brunel, N. (1997). Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cereb. cortex*, 7:237–252.

Barone, P., Batardiere, A., Knoblauch, K., and Kennedy, H. (2000). Laminar distribution of neurons in extrastriate areas projecting to visual areas V1 and V4 correlates with the hierarchical rank and indicates the operation of a distance rule. *J Neurosci*, 20:3263–81.

Benzi, R., Sutera, A., and Vulpiani, A. (1981). The mechanism of stochastic resonance. *J. Phys. A: Math Gen.*, 14:L453–L457.

Berglund, N. and Gentz, B. (2002). Pathwise description of dynamic pitchfork bifurcations with additive noise. *Probab. Theory Related Fields*, 122(3):341–388.

Braitenberg, V. and Schutz, A. (1991). *Anatomy of the cortex: statistics and geometry*. Springer Verlag, Berlin.

Britten, K. H., Newsome, W. T., Shadlen, M. N., Celebrini, S., and Movshon, J. A. (1996). A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Visual Neurosci.*, 13:87–100.

Britten, K. H., Shadlen, M. N., Newsome, W. T., and Movshon, J. A. (1992). The analysis of visual motion: A comparison of neuronal and psychophysical performance. *J. Neurosci.*, 12:2331–2355.

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

Brooks, R. (1991). Intelligence without representations. *Artificial Intelligence*, 47:139–159.

Carpenter, R. H. S. (1988). *Movements of the Eyes. (2nd ed.)*. Pion Limited, London.

Carpenter, R. H. S. and Williams, M. L. L. (1995). Neural computation of log likelihood in the control of saccadic eye movements. *Nature*, 377:59–62.

Chance, F., Abbott, L., and Reyes, A. (2002). Gain modulation from background synaptic input. *Neuron*, 35:773782.

Churchland, P. S. and Sejnowski, T. (1992). *The computational brain*. MIT Press.

Cox, D. R. (1962). *Renewal Theory*. Chapman and Hall, London, UK.

Cox, D. R. and Isham, V. (1980). *Point Processes*. Chapman and Hall.

Cronin, J. (1987). *Mathematical Aspects of Hodgkin-Huxley Neural Theory*. Cambridge University Press, Cambridge, UK.

Damasio, A. R. (1994). *Descartes Error, Emotion, Reson and the Human Brain*. A. Grosset/Putnam Books, New York.

De Schutter, E. and Bower, J. (1994). An active membrane model of the cerebellar purkinje cell, i. simulation of current clamps in slice. *Journal of Neurophysiology*, 71:375–400.

Deleuze, G. and Guattari, F. (1980). *Mille Plateaux - Capitalisme et schizophrénie 2*. Les éditions de minuit, Paris.

Deng, Y., Williams, P., Liu, F., and Feng, J. (2003). Discriminating between different input signals via single neuron activity. *Journal of Physics A: Mathematical and General*, 36(50):12379–12398.

Descartes, R. (1664). *L'Homme*. Charles Angot, Paris.

Di Paolo, E. (2002). Spike timing plasticity for robot control. *Adaptive Behaviour*, 10 (3-4):243–263.

Durrant, S. and Feng, J. (2006). Negative correlated firing: The functional meaning of lateral inhibition within cortical columns. *Biological Cybernetics*, 95 (5):431–453.

Edelman, G. (1987). *Neural Darwinism: the theory of neuronal group selection*. Basic Books, New York.

Elman, J. (1990). Finding structure in time. *Cognitive Science*, 14:179–211.

Feng, J. (2001). Is the integrate-and-fire model good enough? - a review. *Neural Networks*, 14:955–975.

Feng, J. (2003a). *Computational Neuroscience-A Comprehensive Approach*. Chapman-Hall/CRC Press, Boca Raton.

Feng, J. (2003b). Effects of correlated and synchronized stochastic inputs to leaky integrator neuronal models. *Journal of Theoretical Neurobiology*, 222:151162.

Feng, J. and Brown, D. (2000). Impact of correlated inputs on the output of the integrate-and-fire models. *Neural Computation*, 12:671–692.

Feng, J., Deng, Y., and Rossoni, E. (2006). Theory of moment neuronal networks. *Phys. Rev. E*. (accepted).

FitzHugh, R. (1961). Impulses and physiological states in theoretical models of nerve membrane. *Biophys. J.*, 1:455–466.

FitzHugh, R. (1969). Mathematical models for excitation and propagation in nerve. In Schwann, H., editor, *Biological Engineering*, pages 1–85. MacGraw Hill.

Fitzpatrick, D., Usrey, W., Schofield, B., and Einstein, G. (1994). The sublaminar organization of corticogeniculate neurons in layer 6 of macaque striate cortex. *Vis Neurosci*, 11:307–15.

Foucault, M. (1970). *The Order of Things: An Archaeology of the Human Sciences*. Tavistock, London.

Gaillard, B. (2002). A neuronal model for visual discrimination task. Master's thesis, University of Sussex. http://www.cogs.susx.ac.uk/users/bg22/monkey.pdf.

Gaillard, B. (2003). Winner takes all strategy with spike timing: an application to motion perception. Technical report, University of Sussex. http://www.cogs.susx.ac.uk/users/bg22/report.pdf.

Gaillard, B. and Feng, J. (2005). Modelling a visual discrimination task. *Neurocomputing*, 65-66:203–209.

Gaillard, B., Feng, J., and Buxton, H. (2006a). Neuronal model of decision making. In Feng, J., Qian, M., and Jost, J., editors, *Networks: From Biology to Theory*. Springer Verlag. In Press.

Gaillard, B., Feng, J., and Buxton, H. (2006b). Population approach to a neural discrimination task. *Biological cybernetics*, 94(3):180–191.

Gerstner, W. and Kistler, W. (2002). *Spiking Neuron Models, Single Neurons, Populations, Plasticity*. Cambridge University Press.

Glimcher, P. (2003a). *Decisions, Uncertainty and the brain: the Science of Neuroeconomics*. MIT Press, Cambridge, MA.

Glimcher, P. W. (2003b). The neurobiology of visual-saccadic decision making. *Annu. Rev. Neurosci.*, 26:133–179.

Gold, J. I. and Shadlen, M. N. (2001). Neural computation that underlie decisions about sensory stimuli. *Trends in Cognitive Science*, 5:10–16.

Gold, J. I. and Shadlen, M. N. (2003). The influence of behavioral context on the representation of a perceptual decision in developping oculomotor commands. *Journal of Neuroscience*, 23:632 – 651.

Gulyás, A., Miles, R., Sik, A., Toth, K., Tamamaki, N., and Freund, T. (1993). Hippocampal pyramidal cells excite inhibitory neurons through a single release site. *Nature*, 366:683–687.

Harris, C. S. (1965). Perceptual adaptation to inverted, reversed and deplaced vision. *Psychological Review*, 72 (6):419–444.

Haykin, S. (1999). *Neural Networks, A comprehensive Foundation*. Prentice Hall.

Hebb, D. (1949). *The Organization of Behavior: A Neurophysiological Theory*. John Wiley, New York.

Heeger, D., Huk, A., Geisler, W., and Albrecht, D. (2000). Spikes versus bold: what does neuroimaging tell us about neuronal activity? *Nature Neuroscience*, 3:631–633.

Heeger, D. J., Simoncelli, E. P., Carandini, M., and Movshon, J. A. (1995). Computational models of cortical visual processing. In *Proc National Academy of Science*, volume 93, pages 623–627.

Heekeren, H. R., Marrett, S., Bandettini, P. A., and Ungerleider, L. G. (2004). A general mechanism for perceptual decision-making in the human brain. *Nature*, 431:859–862.

Hille, B. (1992). *Ionic channels of excitable membranes*. Sinauer Associates, Sunderlan, Mass.

Hodgkin, M. and Huxley, A. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol. (London)*, 117:500–544.

Hopfield, J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci USA*, 79:2554–2558.

Hopfield, J. (1984). Neurons with graded response have collective computational properties like those of two-state neurons. *Proc. Natl. Acad. Sci USA*, 81:3088–3092.

Hopfield, J. (1995). Pattern recognition computing using action potential timing for stimulus representation. *Nature*, 376:33–36.

Horwitz, G. d., Batista, A. P., and Newsome, W. T. (2004). Representation of abstract perceptual decision in macaque superior colliculus. *Journal of Neurophysiology*, 91:2281–2296.

Hubel, D. H. and Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol. (Lond.)*, 148:574–591.

Huk, A. C. and Shadlen, M. N. (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.

Izhikevich, E. M. (2003). Simple model of spiking neurons. *IEEE Transactions on Neural Networks*, 14 (6):1569–1572.

Izhikevich, E. M. (2004). Which model to use for cortical spiking neurons. *IEEE Transactions on Neural Networks*, 15 (5):1063–1070.

Izhikevich, E. M. (2006). Polychronization: Computation with spikes. *Neural Computation*, 18:245–282.

Izhikevich, E. M., A., G. J., and Edelman, G. M. (2004). Spike-timing dynamics of neuronal groups. *Cerebral Cortex*, 14:933–944.

Jack, J., Noble, D., and Tsien, R. (1975). *Electric Current Flow in excitable Cells*. Oxford University Press, Oxford, UK.

Jensen, D. (2000). *A Language Older than Words*. Context Books, New York.

Katz, B. (1969). *The Release of Neural Transmitter Substances*. Liverpool University Press, Liverpool, United Kingdom.

Kelso, S., Ganong, A., and Brown, T. (1986). Hebbian synapses in hippocampus. *Proc. natl. Acad. Sci. USA*, 83:5326–5330.

Kim, J. N. and Shadlen, M. N. (1999). Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nature Neurosci.*, 2:176–185.

Koch, C. (1999). *Biophysics of Computation: Information Processing in Single Neurons*. Oxford University Press.

Kohonen, T. (1995). *Self-organising maps*. Springer, Berlin.

Kreps, D. (1990). *A course in microeconomics*. Princeton University Press, Princeton, NJ.

Laberge, D. (1962). A recruitment theory of simple behavior. *Psychometrika*, 27:375–387.

Lapicque, L. (1907). Recherches quantitatives sur l'excitation électrique des nerfs traitée comme une polarisation. *J. Physiol. Paris*, 9:620–635.

Lapicque, L. (1926). *L'excitabilité en fonction du temps*. Presses Universitaires de France, Paris.

Liu, J. and Newsome, W. T. (2003). Functional organization of speed tuned neurons in visual area MT. *J. Neurophysiology*, 89:246–256.

Lund, J., Lund, R., Hendrickson, A., Bunt, A., and Fuchs, A. (1975). The origin of efferent pathways from the primary visual cortex, area 17, of the macaque monkey as shown by retrograde transport of horseradish peroxidase. *J Comp Neurol*, 164:287–303.

Maunsell, J. and Van Essen, D. (1983). The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *J Neurosci*, 3:2563–86.

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

McCulloch, W. and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5:115–133.

Merleau-Ponty, M. (1945). *Phénoménologie de la perception*. Gallimard, Paris.

Merleau-Ponty, M. (1968). *Résumés de cours au Collège de France*. Gallimard, Paris.

Minsky, M. and Papert, S. (1969). *Perceptrons*. MIT Press, Cambridge, MA.

Mountcastle, V. B. (1957). Modality and topographic properties of single neurons of cat's somatosensory cortex. *J. Neurophysiol.*, 20.

Nagumo, J., Arimoto, S., and Yoshizawa, S. (1999). An active pulse transmition line simulating nerve axon. *Proc. IRE*, 50:2061–2070.

Newsome, W. T., Britten, K. H., and Movshon, J. A. (1989). Neural correlates of a perceptual decision. *Nature*, 341:52–54.

O'Regan, J. K. and Noe, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24(5):939–1031.

Palmer, J., Huk, A., and Shadlen, M. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. *Journal of Vision*, 5:376–404.

Platt, M. L. and Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex. *Nature*, 400:233–238.

Polyak, S. (1957). *The vertebrate visual system*. Univ. of Chicago Press, Chicago.

Puskorius, G., LA, F., and LI, D. (1996). Dynamic neural network methods applied to on-vehicule idle speed control. *Proceedings of IEEE*, 84:1407–1420.

Ramón y Cajal, S. (1909). *Histologie du système nerveux de l'homme et des vertébrés*. A. Maloine, Paris.

Rao, R. and Sejnowski, T. J. (2001). Spike-timing hebbian plasticity as temporal difference learning. *Neural Computation*, 13(10):22212237.

Rao, R. P. N. (2004). Bayesian computation in recurrent neural circuits. *Neural Computation*, 16(1). To appear.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85:59–108.

Ratcliff, R. and Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, 9:347–356.

Ratcliff, R., Zandt, T. V., and McKoon, G. (1999). Connectionist and diffusion models of reaction time. *Psychological Review*, 106:261–300.

Reddi, B. A., Asress, K. N., and Carpenter, R. H. S. (2003). Accuracy, information and response time in a saccadic decision task. *J. Neurophysiol*, 90:3538–3546.

Reddi, B. A. and Carpenter, R. H. (2000). The influence of urgency on decision time. *Nature Neuroscience*, 3(8):827–830.

Reid, R., Soodak, R., and RM, S. (1987). Linear mechanisms of directional selectivity in simple cells of cat striate cortex. In *Proc Natl Acad Sci USA*, volume 84, pages 8740—-8744.

Rieke, F., Warland, D., De Ruter Van Steveninck, R. R., and Bialek, W. (1997). *Spikes, exploring the neural code*. MIT Press, Cambridge, MA.

Robinson, D. (1964). The mechanics of human saccadic eye movements. *Journal of Physiology*, 174:245–264.

Rockland, K., Saleem, K., and Tanaka, K. (1994). Divergent feedback connections from areas V4 and TEO in the macaque. *Vis Neurosci*, 11:579–600.

Roitman, J. D. and Shadlen, M. N. (2002). Response of neurons in the lateral intraparietal area (LIP) during a combined visual discrimination reaction time task. *J. Neurosci.*, 22:9475–9489.

Romo, R., Hernandez, A., Zainos, A., and Salinas, E. (2003). Correlated neuronal discharges that increase coding efficiency during perceptual discrimination. *Neuron*, 38:649–657.

Rorie, A. E. and Newsome, W. T. (2005). A general mechanism for decision-making in the human brain? *Trends in Cognitive Sciences*, 9. Advanced Online Publication.

Rosenbaltt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65:386–408.

Rumelhart, D. and McClelland, J. (1986). *Parallel Distributed Processing: Exploration In the Microstructure of Cognition*. MIT Press, Cambridge, MA.

Ruthruff, E. (1996). A test of the deadline model of speed-accuracy tradeoffs. *Perception and Psychophysics*, 58:56–64.

Salinas, E. (2003). Background synaptic activity as a switch between dynamical states in a network. *Neural Computation*, 15:1439–1475.

Salinas, E. and Sejnowski, T. (2001). Correlated neuronal activity and the flow of neural information. *Nature Reviews Neuroscience*, 2:539–550.

Salzman, C. D., Murasugi, C. M., Britten, K. H., and Newsome, W. T. (1992). Microstimulation in visual area MT: Effects on direction discrimination performance. *Journal of neuroscience*, 12:2331–2355.

Sejnowski, T. (1977). Statistical constraints on synaptic plasticity. *J. theor. Biology*, 69:385–389.

Shadlen, M. and Newsome, W. (1994). Noise, neural codes and cortical organization. *Curr Opin Neurobiol*, 4:569.

Shadlen, M. N. and Gold, J. I. (2004). The neurophysiology of decision making as a window on cognition. In Gazzaniga, editor, *The Cognitive Neurosciences, 3rd edition*. MIT Press.

Shadlen, M. N. and Newsome, W. (1996). Motion perception : seeing and deciding. *PNAS*, 93:628–633.

Shadlen, M. N. and Newsome, W. T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J. Neurophysiol.*, 86:1916–1935.

Sherringdon, C. (1906). *The Integrative Action of the Nervous System*. Scribner, New York.

Sheth, B. R., Sharma, J., Rao, S. C., and Sur, M. (1996). Orientation maps of subjective contours in visual cortex. *Science*, 274:2110–2115.

Simoncelli, E. P. and Heeger, D. J. (1998). A model of neuronal responses in visual area MT. *Vision Research*, 38:743–761.

Simons, M. and Pellionisz, A. (2006). Genomics, morphogenesis and biophysics: triangulation of purkinje cell development. *The Cerebellum*, 5(1):27–35.

Smith, P. L. and Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neuroscience*, 27(3):162–168.

Softky, W. and Koch, C. (1993). The highly irregular firing of cortical cells is inconsistent with temporal integration of random epsps. *J. Neurosci.*, 13:334–350.

Sparks, D. and Le, M. (2004). Movement fields of saccade-related burst neurons in the monkey superior colliculus. *Brain Res.*, 190 (1):39–50.

Stein, R. (1967). Some models of neuronal variability. *Biophys. J.*, 7:37–68.

Stratton, G. M. (1897). Vision without inversion of the retinal image. *Psychological Review*, 4:341–360 and 463–481.

Sugrue, L. P., Corrado, G. S., and Newsome, W. T. (2005). Choosing the greater of two goods: Neural currencies for valuation and decision making. *Nature Reviews Neuroscience*, Advanced Online Publication.

Suzuki, W., Saleem, K., and K, T. (2000). Divergent backward projections from the anterior part of the inferotemporal cortex (area TEO) in the macaque. *J Comp Neurol*, 422:206–28.

Swadlow, H. A. (1994). Efferent neurons and suspected interneurons in motor cortex of the awake rabbit: axonal properties sensory receptive fields, and subthreshold synaptic inputs. *Journal of Neurophysiology*, 71:437–453.

Thomas, N. J. T. (1999). Are theories of imagery theories of imagination? an active perception approach to conscious mental content. *Cognitive Science*, 23:207–245.

Tovée, M. J. (1996). *An Introduction to the Visual System*. Cambridge University Press.

Townsend, J. T. and Ashby, F. G. A. (1983). *The stochastic modelling of elementary psychological processes*. Cambridge University Press, Cambridge.

Tuckwell, H. C. (1988). *Introduction to Theoretical Neurobiology*, volume 2. Cambridge University Press, Cambridge.

Ungerleider, L. and Desimone, R. (1986a). Cortical connections of visual area mt in the macaque. *J Comp Neurol*, 248:190–222.

Ungerleider, L. and Desimone, R. (1986b). Projections to the superior temporal sulcus from the central and peripheral field representations of v1 and v2. *J Comp Neurol*, 248:147–63.

Usher, M. and Feingold, M. (2000). Stochastic resonance in the speed of memory retrieval. *Biological Cybernetics*, 83:L11–L16.

Usher, M. and McClelland, J. (2001). On the time course of perceptual choice: The leaky competing accumulator model. *Psychological Review*, 108:550–592.

Van Essen, D., Anderson, C., and Felleman, D. (1992). Information processing in the primate visual system: an integrated systems perspective. *Science*, 255:419–423.

Vanrullen, R. (2002). *Une première vague de potentiels d'action, une première vague idée de la scène visuelle*. PhD thesis, Centre de Recherche Cerveau et Cognition, Faculté de Médecine de Rangueil, Université Paul Sabatier, Toulouse.

VanRullen, R. and Koch, C. (2003). Is perception discrete or continuous? *Trends in Cognitive Science*, 7(5):207–213.

Wang, X. J. (2002). Probabilistic decision making by slow reverberation in cortical circuits. *Neuron*, 36:955–968.

Waxman, S. G. and Bennett, M. (1972). Relative conduction velocities of small myelinated and non-myelinated fibres in the central nervous system. *Nature New Biol.*, 238:217–219.

Wehemeier, U., Don, D., Koch, C., and Van Essen, D. (1989). Modeling the mamalian visual system. In Koch, C. and Segev, I., editors, *Methods in neuronal modeling*. MIT Press.

Wolfart, J., Debay, D., Masson, G. L., Destexhe, A., and Bal, T. (2005). Synaptic background activity controls spike transfer from thalamus to cortex. *Nature Neuroscience*, 8:1760–1767.

Wurtz, R., Sommer, M., M, P., and Ferraina, S. (2001). Signal transformations from cerebral cortex to superior colliculus for the generation of saccades. *Vision Research*, 41(25-26):3399–3412.

Wurtz, R. H. and Goldberg, M. E. (1989). *The Neurobiology of Saccadic Eye Movements. Reviews of Oculomotor Research Volume 3.* Elsevier, New York.

Zohary, E., Shadlen, M. N., and Newsome, W. T. (1994). Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature*, 370:140–143.

Appendix A

Moment Neural Networks

Convergence to stable low-level activity in a recurrent network

We simulated the convergence to a low level activity equilibrium with the values expressed in Feng et al. (2006), and also with other parameter values. They studied the propagation of activity in a feed forward network. We adapted their idea and connected the neurons to themselves, creating a homogeneous recurrent neural network. Such a network can be described as a dynamical system, the firing rate being a variable of the system. We showed that the state of homogeneous low level activity is an attractor of this dynamical system, and in Fig. A.1, we can see that this attractor has a limited basin of attraction.

Non homogeneous case: a model for decision making

As before, we modelled decision making in a LIP column, where the accumulation of evidence and the competition were generated via recursive excitation and inhibition. The structure of this network is simpler than that of the model presented in Chapter 4 because we used negative recursive connections instead of inhibitory inter-neurons. Similarly to the Poisson case, we used standard values for the LIF neuron parameters, and we only studied the effect of variation of certain parameters: the relative synaptic strength between the various subpopulations and the statistical signature (two first moments) of the background activity. The parameters of the network are described below:

- $200 \leq p \leq 400$ number of excitatory incoming synapses from one subpopulation.
- $a = 0.5$ mV EPSP.
- $J = ap$ global input strength.

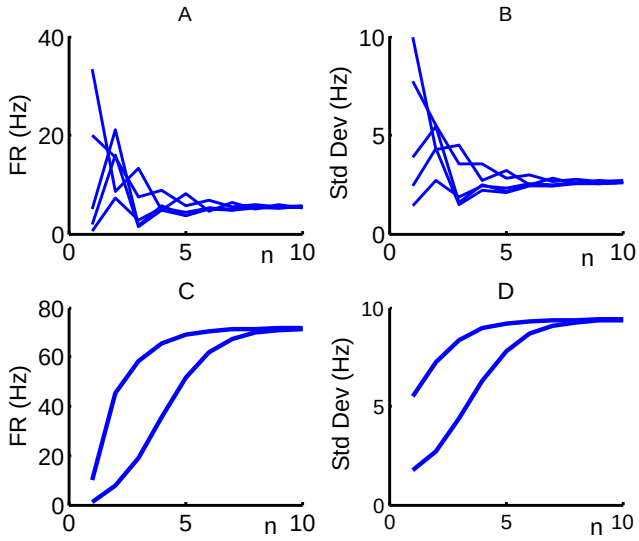


Figure A.1: Illustration of the convergence of the homogeneous system to a low level activity state, with 300 excitatory neurons. Panel A: FR as a function of the number of recursive steps, for a ratio between inhibitory and excitatory inputs $r = 3$ and for various initial states. Panel B: Standard deviation of efferent ISIs as a function of the number of recursive steps, for $r = 3$ and for the same various initial states. Panel C: FR as a function of the number of recursive steps, for $r = 1$ and for a low and a high initial activity. Panel D: Standard deviation as a function of the number of recursive steps, for $r = 1$ and for a low and a high initial activity.

- $x = 0.6$ relative strength of recursive connections.
- $xi = 0.2$ relative strength of background noise synapses.
- $f = 1 - x - xi$ relative strength of stimulus synapses.
- $V_\theta = 20 * 0.001$ threshold potential in Volt.
- $V_r = 0$ resting potential in Volt.
- $L = 1/(20 * 0.001)$ membrane decay rate in 1/s.
- $r = 1.1$ ratio between inhibitory and excitatory input.

The subpopulations are almost the same as in Chapter 4. The differences of this model lie in the fact that we do not use inhibitory interneurons or non specific excitatory or inhibitory neurons. On the other hand, we have divided the external input (form direction detectors and from background activity) into excitatory and inhibitory inputs. This leads to the following description of the homogeneous subpopulations of neurons:

- subpopulation 1 is $\mathcal{E}_{up}$, specifically sensitive to the input from the upwards detectors
- subpopulation 2 is $\mathcal{E}_{down}$, specifically sensitive to the input from the downwards detectors

- subpopulation 3 is the subpopulation that generates excitatory background activity
- subpopulation 4 is the subpopulation that generates inhibitory background activity
- subpopulation 5 is the upwards direction detector
- subpopulation 6 is the downwards direction detector

The synaptic matrix *Syn* describes concisely the synaptic weights. $Syn_{i,j}$ is the relative strength of the synaptic input from subpopulation j to one neuron in subpopulation i, as was $W_{i,j}$ in matrix W in Chap 4:

$$Syn = J \begin{pmatrix} x & -x\cdot r & xi & -r\cdot xi & f & -f\cdot r \\ -x\cdot r & x & xi & -r\cdot xi & -f\cdot r & f \end{pmatrix} \tag{A.0.1}$$

The matrix *Syn* is not a square matrix because we do not model the synaptic interactions that generate the external inputs (background activity and direction detectors) which are modelled independently from this model. Consequently the neural activity recursively generated by the model presented here is only the activity of population 1 and 2, that is why *Syn* only has two lines.

In order to compute the two first output moments of the neurons of the system, we need to compute their global post-synaptic input and then apply the formula presented in Eq. 2.2.20 and Eq. 2.2.23. The global post-synaptic inputs are defined by Eq. 2.2.18 and Eq. 2.2.19:

$$\bar{\mu} = Syn \times \vec{\mu_{in}} \tag{A.0.2}$$

In this equation $\bar{\mu}$ is a vector of two dimensions, and $\vec{\mu_{in}}$ is the six dimensional input vector that includes the background activity and the signals from the direction detectors.

For the vector $\bar{\sigma}$, the relation is not linear, so we cannot straightforwardly collapse all the incoming synaptic weights in the elements of matrix *Syn*. As in Chapter 4, we described the correlation between the subpopulations by a correlation matrix c, where $c_{m,n}$ is the correlation coefficient between a spike train emitted by a neuron from subpopulation n and one from subpopulation m. I we look at matrix *Syn* in more details, we see that, for example, $syn_{1,2} = -a\cdot p\cdot x\cdot r$. We can decompose this quantity into $p\cdot r$, the number of synapses, and $w_{1,2} = -a\cdot x$, the synaptic weight from one neuron of subpopulation 2 ($\mathcal{E}_{down}$) to one neuron of subpopulation 1 ($\mathcal{E}_{up}$). We then reach the following expression for the standard deviation of the postsynaptic input of neurons in $\mathcal{E}_{up}$ if $i = 1$ or $\mathcal{E}_{down}$ if $i = 2$:

$$\bar{\sigma_i}^2 = \sum_{m,n=1}^{p_{tot}} |w_{i,n}||w_{i,m}|\sigma_n\sigma_n c_{m,n} \tag{A.0.3}$$

where p_{tot} is the total number of incoming synapses to the neuron.

Simulations of decision making

We simulated the model previously described with kinematogram stimuli. The Superior Colliculus (SC) activity was, however, only influenced by the FRs of the neurons that gather evidence in the LIP column. As before, the time needed for the SC activity to reach a specific threshold was measured in order to determine a Reaction Time.

We reproduced the competitive accumulation of evidence that is experimentally observed and that we simulated for the case of Poissonian spike trains. This is illustrated in Fig. A.2.

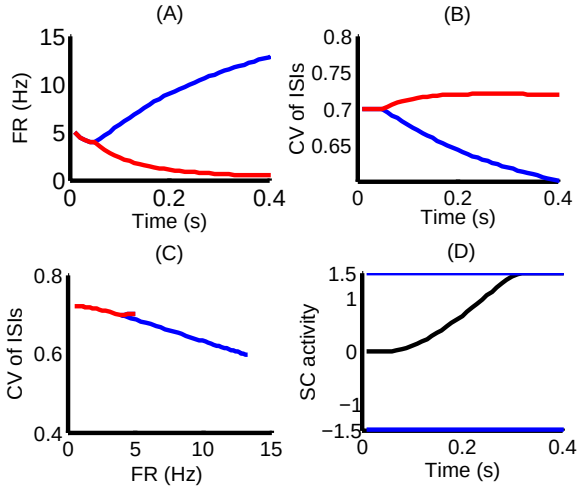


Figure A.2: Illustration of the time course of a decision making with the MNN. The blue lines represent the activity of the neural group gathering evidence for the upwards movement, the red lines for the downwards movement. The activity of the Superior Colliculus in panel (C) is represented without unit, because it is normalised to $\pm\frac{\pi}{2}$*. During the first 50 ms not stimuli were given to the system so that it could stabilise in a homogeneous low activity before the decision making. The background activity had a FR of 5 Hz and a CV of ISIs of 1. The simulation was run for 300 neurons, the ratio between inhibitory and excitatory inputs:* $r = 1.1$*, and* $k = l = 1$ *in the SC equation. We simplified the evidence and modelled it by a FR (50 Hz and 48 Hz respectively) and a CV of ISIs of 0.5.*

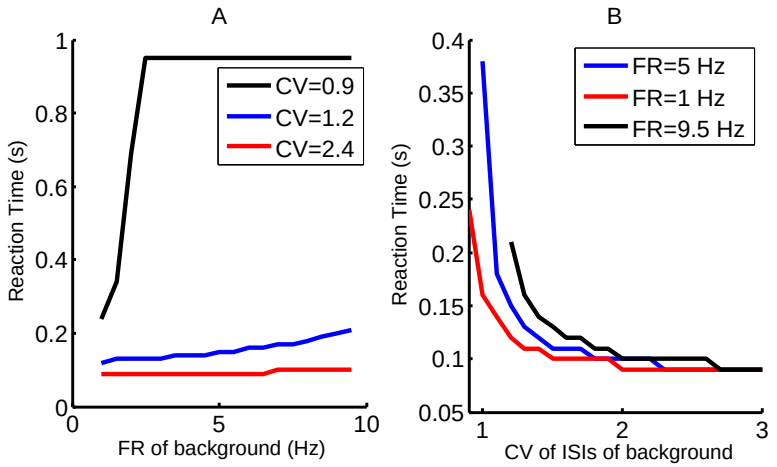


Figure A.3: Reaction Time of the MNN for decision making as a function of the FR of the background activity (panel A), and as a function of the coefficient of variation (panel B). The parameters were the same as in Fig. A.2. We have interrupted the decision making process after 1 s, hence the plateau in panel (A). Panel (A) contradicts previous results of Chapter 4 because, in the current model, the background activity is both inhibitory and excitatory.

We measured the dependence of RT on the statistical signature of the background activity. The results are shown in Fig. A.3. Contrary to the results of the model with Poisson spike trains, in Chapter 4, the RT increases with the mean background activity. This happens because the two models have different structures. In the Poisson model, the only inhibitory connections come from the inhibitory neurons that connect the two competing subpopulations, so the background activity is purely excitatory as we can see in the connectivity matrix W in Eq. 4.1.21. On the contrary, in the MNN model for decision making, the background activity is inhibitory and excitatory, with the same ratio r as the ratio between the strength of the mutual inhibition and the internal recurrent excitation, as it can be seen in the connectivity matrix Syn in Eq. A.0.1. Consequently, increasing the mean background activity with a ratio r greater than one inhibits all neurons and thus prevents local neural activities from diverging from a low activity. The two models are therefore not comparable with the current parameters.

The interest of the MNN model for decision making lies precisely in the fact that, by having a ratio between inhibitory and excitatory input slightly greater than one (in the present simulations, $r = 1.1$), we highlight the role of second order statistics. This actually is the reason why we designed this model. We can see in Fig. A.3 the effect of the coefficient of variation of the ISIs on the decision dynamics: the RT is monotonically decreasing with the coefficient of variation. Moreover we see that when the variance is past a certain threshold above which the decision becomes possible, the actual mean intensity of the background activity has little effect, and the second statistical moment controls the speed of decision making.

Appendix B

Topology and stable states of a network

In this appendix, we present our attempts to simulate the emergence of low level background activity of large networks of the same LIF neurons as the ones used in the rest of this dissertation. We modelled the generation of each spike. We also simulated the propagation of spikes in the network. The aim was to create a population whose neural activity would converge to a FR of 5 Hz without external input.

B.1 Topology

We decided to spread the neurons evenly over a sphere. We assumed that the connections between neurons were straight lines cutting across the sphere. Therefore, if the position of neuron i is defined by (θ_i, ϕ_i), then the distance between neuron i and neuron j is:

$$d_{i,j} = \sqrt{(2 \cdot rad \cdot sin(\theta_i - \theta_j))^2 + (2 \cdot rad \cdot sin(\phi_i - \phi_j))^2} \tag{B.1.1}$$

where *rad* is the radius of the sphere. This topology has been used by Izhikevich et al. (2004). In Fig. B.1, we have an illustration with special subpopulations of neurons in a different colour. The red and green coloured neurons represent respectively neurons from $\mathcal{E}_{up}$ and $\mathcal{E}_{down}$ as described in Chapter 4 and Appendix A. Their positions on the sphere highlight the fact that they are competing and the presence of long range inhibition.

The inhibitory connections were made slower than the excitatory connections. We defined two more parameters, V_{unmyel}, the velocity of a spike coming from an inhibitory neuron, and V_{myel}, the velocity of a spike coming from an excitatory neuron. The time for a spike to reach a synapse is $t = \frac{d}{v}$.

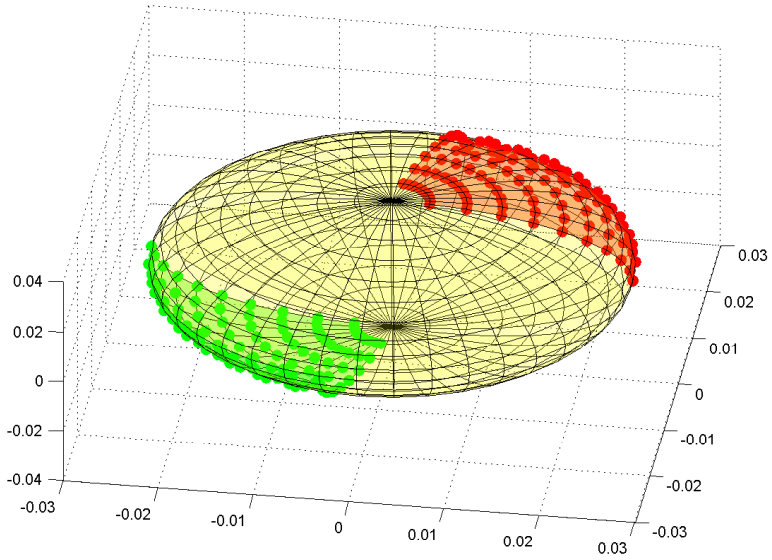


Figure B.1: Topology of the decision making network of spiking neurons. We attempted to model decision making, but could not obtain satisfying results. This is an illustration of the network that we attempted to use to model decision making, where we can see the two populations of neurons that were meant to accumulate evidence over the course of the decision, in red $\mathcal{E}_{up}$ and in green $\mathcal{E}_{down}$.

B.2 Parameters and initial conditions

B.2.1 Synaptic connections

Each neuron has a probability P_{inh} to be excitatory. We note $w_{i,j}$ the strength of the depolarisation of the membrane of neuron j after receiving a spike from neuron i. It is referred to as 'synaptic weight' or 'Post Synaptic Potential' (PSP), that is either inhibitory or excitatory.

If neuron i is excitatory,

$$w_{i,j} = A \cdot e^{-\frac{d_{i,j}}{k}} \tag{B.2.1}$$

where k is a parameter that we vary. A is the upper limit of the strength of the EPSP.

If neuron i is inhibitory,

$$w_{i,j} = B \cdot e^{-\frac{d_{i,j}}{k}} \tag{B.2.2}$$

where $B = r_{inh} \cdot A$ and $r_{inh} = \frac{IPSP}{EPSP}$.

B.2.2 Initial conditions

The network of neurons can remain indefinitely in the state of zero activity. So, in order to make it converge to a low non-zero stable activity, we had to kick-start the dynamics of spike generation. In order to do so more realistically than forcing it in an initial activity level, we 'poked' it: We assumed that, initially, all neurons of the network received a Poissonian synaptic input added to their recurrent connections. That put the network into a state of low level activity. After a short time, we released this drive in order to see the dynamics of activity of the network by itself. The characteristics of the poking input were as follows:

- H, the rate of the Poisson process (FR).
- $n_{syn} = 75$ is the number of synapses that did work in the moment mapping network presented in Feng et al. (2006).
- $r = 1.5$ is the ratio between inhibitory and excitatory inputs.
- $a = 0.5$ EPSP in mV.
- $c = 0.1$ correlation between inputs.

B.2.3 Parameters

During the simulations, we varied the following parameters:

- rad, the radius of the sphere.
- A the EPSP value.
- $r_{inh} = \frac{IPSP}{EPSP}$.
- k, in Eq. B.2.2, controls how sharply the synaptic weight decreases with distance.
- p_e, the proportion of excitatory neurons.
- $N_{neurons}$, the number of neurons.
- H, the rate of the Poisson process used for the poking.
- V_{myel}, the velocity of an excitatory spike.
- V_{unmyel}, the velocity of an inhibitory spike.
- τ_{ref}, the refractory period.

B.3 Simulation results

B.3.1 Stability of the network without a refractory period

We present our results in a form of raster plots. Each dot on a raster plot represents a spike. On the abscissa, we have time and on the ordinate, we have the neural number. On the raster plot, we do not see the topology of the network.

With a random external stimulation, it was very easy to generate low level random activity. However, when we stopped the external activation, the network either fell into a completely quiet state or its activity diverged to its maximum possible activity. We illustrate this phenomenon in Fig. B.2. In these two figures, the same network in one case diverges to a very high activity and in the other case converges to a quiet activity, for two different instances of artificial poking (with the same parameters). That suggests that this network is not stable at a low or even reasonable activity. We would not claim that the network only has two stable states (quiet and maximal activity), because we have not formally studied its dynamics, we only have performed simulations. However, we have simulated the network's behaviour many times and have explored the parameter range quite extensively, without finding a stable state other than these two. It would be interesting to attempt to prove the hypothesis, suggested by these simulations that the network only has two stable states in this case.

We have reached this conclusion using a dichotomy search method over one parameter: If the parameter is too large, the network diverges, but if the parameter is too small the network's activity converges to zero spike per second (or vice-versa). When we get close to the middle value of the parameter, the network probabilistically diverges or dies out, but it does not maintain an intermediate activity as we had hoped would happen. By contrast, we can see in Fig. B.3 that, in presence of external Poisson perturbation, the network is stable.

B.3.2 Stability of the network with delays and refractory period

By including a refractory period in the model, we managed to stabilise the network. With delays that exactly corresponded to the distance divided by the spike's velocity, we reached synchrony. However, if the delays were randomised, we obtained repetitive asynchronous patterns. We illustrate this result in Fig. B.4, where we clearly see with two different sets of parameters that the network's FR converged to a level smaller than the maximal FR that it could reach ($FR_{max} = \frac{1}{\tau_{ref}}$), but clearly non-zero. This is what we call a Submaximal Stable Activity (SSA).

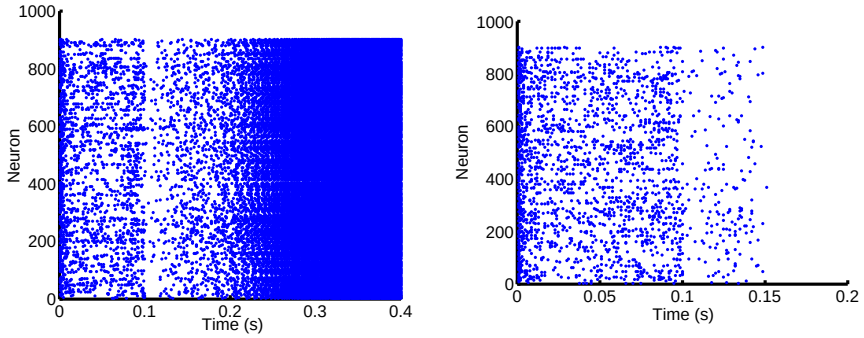


Figure B.2: Raster plots of neural activity. The parameters are: rad $= 38$ *mm, corresponding to* 2.5 *neurons per* mm^2, $A = 0.5$ *mV,* $r_{inh} = 5$, $k = 1$ *for inhibitory synapse,* $k = 1/12$ *for excitatory synapses,* $p_e = 0.5$, $N_{neurons} = 900$, $V_{myel} = 1$ *m/s,* $V_{unmyel} = 0.5$ *m/s, and there is no refractory period. We simulated* 0.4 *seconds, during which the network was subjected for* 0.2 *seconds to a Poisson external input of rate* $H = 2.5$, $r = 2.5$, *and then left alone. We see that two similar instantiations of poking lead the system to two radically different behaviours, as soon as the external input is removed. This proves the instability of the system. In the right panel, the raster plot stops after the last spike of the network.*

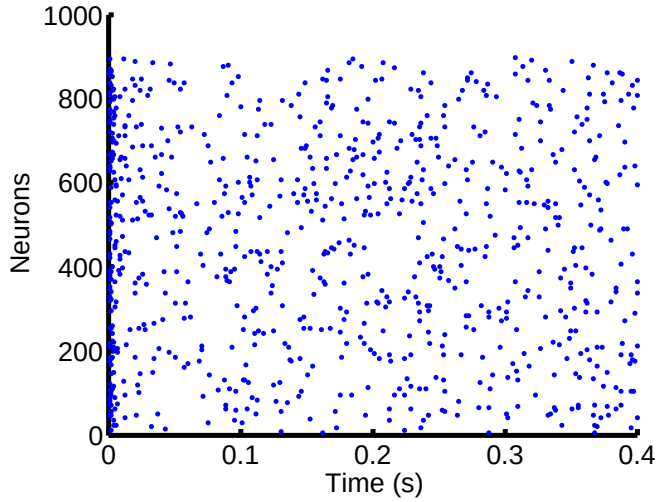


Figure B.3: Raster plots of neural activities. The parameters are: rad $= 75$ *mm, corresponding to* 1.25 *neurons per* mm^2, $A = 0.5$ *mV,* $r_{inh} = 5$, $k = 10$ *for inhibitory synapses,* $k = 1$ *for excitatory synapses,* $p_e = 0.5$, $N_{neurons} = 900$, $V_{myel} = 1$ *m/s,* $V_{unmyel} = 0.5$ *m/s, and there is no refractory period. We simulated* 0.4 *seconds, during which the network was subjected to a Poisson external input of rate* $H = 2.5$, $r = 2.5$.

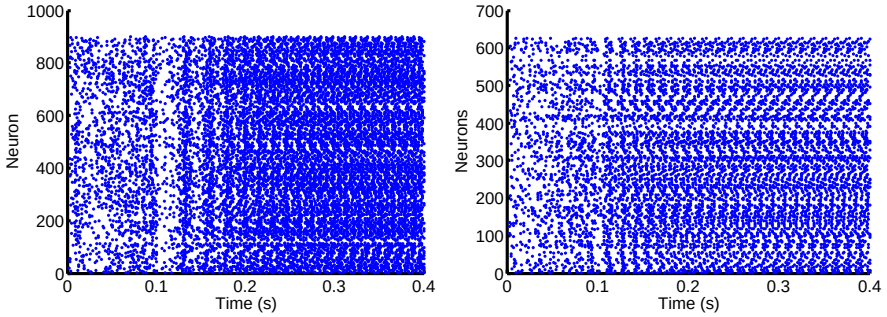


Figure B.4: Convergence to SSA. . We simulated 0.4 *seconds. During* 0.1 *seconds, the network was submitted to a Poisson external input of rate* $H = 5.2$, $r = 1.5$. *Left panel:* $rad = 37.8$ *mm,* $A = 0.5$ *mV,* $r_{inh} = 5$, $k = 1$ *(inhibitory synapses),* $k = 1/15$ *(excitatory),* $p_e = 0.3$, $N_{neurons} = 900$, $V_{myel} = 1$ *m/s,* $V_{unmyel} = 0.15$ *m/s,* $\tau_{ref} = 7$ *ms. The efferent FR was* $FR = 43$ *Hz. Right panel:* $rad = 13.9$ *mm,* $A = 0.5$ *mV,* $r_{inh} = 5$, $k = 1$ *(inhibitory synapses),* $k = 1/15$ *(excitatory),* $p_e = 0.5$, $N_{neurons} = 625$, $V_{myel} = 1$ *m/s,* $V_{unmyel} = 0.15$ *m/s,* $\tau_{ref} = 7$ *ms. Efferent* $FR = 37$ *Hz.*

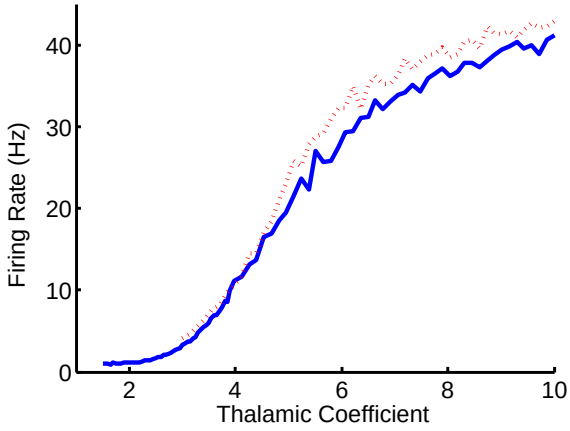


Figure B.5: Stable network FR as a function of thalamic FR. The network was subjected for 1 *s. to a Poisson external input of rate* H *that varies from 1.5 to 10 HZ along the* x *axis.* $r = 3$. *The membrane potential time constant was adjusted to* $\tau = 0.01s$. *The dotted line presents results of a second simulation with a new instantiation of the network (same parameters). It shows that the results are consistent. The parameters are:* $rad = 18.9$ *mm,* $A = 0.5$ *mV,* $r_{inh} = 5$, $k = 1$ *for inhibitory synapses,* $k = 1/15$ *for excitatory synapses,* $p_e = 0.5$, $N_{neurons} = 900$, $V_{myel} = 1$ *m/s,* $V_{unmyel} = 0.15$ *m/s,* $\tau_{ref} = 7$ *ms.*

B.3.3 Using thalamic input to control global activity

As seen in the previous paragraph, we managed to stabilise the network. However, the FR to which the network converges is in the range of a few tens of Hz. This is much larger than the mentioned 5 Hz that is characteristic of the low level activity across the brain. Modifying the radius of the sphere gave us a small amount of control, and we observed that the Stable Submaximal Activity decreased when the radius increased. However, very quickly, we observed the same phenomenon of instability as the one that proved that we cannot stabilise the network without refractory period: adjusting the radius to its maximum value leads to a point where, for two experiments with the same parameters, the network can converge to zero activity or to the illustrated Submaximal Stable Activity (SSA). We hope that, if we significantly increase the number of neurons we will be able, by increasing the radius, to reduce this SSA to the experimental 5-10 Hz.

As seen in Izhikevich (2006), extensive models of stable spiking networks are often given an additional input that is supposed to come from the thalamus. We have already used this idea to stir the network from its initial zero activity, and also to stabilise a network with no refractory period. We were also able to use it to control the value of the SSA. As seen in Fig. B.5, the amount of external input can control the SSA in a monotonic way.