UNIVERSITÉ DE GENÈVE
FACULTÉ DES SCIENCES

LEMANIC NEURO SCIENCE DOCTORAL SCHOOL

DOCTORAT EN NEUROSCIENCES
des Universités de Genève
et de Lausanne

UNIVERSITÉ DE GENÈVE
UNIL | Université de Lausanne

UNIVERSITÉ DE GENÈVE FACULTÉ DES SCIENCES

Dr. Renaud Jolivet, directeur de thèse

TITRE DE LA THÈSE

Energetic Efficiency of Information Transfer at Synapses: Methods of Information Measurement and Effects of Neuromodulation by Serotonin

THÈSE
Présentée à la
Faculté des Sciences

de l'Université de Genève

pour obtenir le grade de
Docteure ès sciences, mention Neurosciences

par

Mireille CONRAD

de Nods (BE)

Thèse N° 282

Genève

Université de Genève

2020

UNIVERSITÉ DE GENÈVE

FACULTÉ DES SCIENCES

DOCTORAT EN NEUROSCIENCES
des Universités de Genève et de Lausanne

Thèse de Madame Mireille CONRAD

Intitulée : ***Energetic Efficiency of Information Transfer at Synapses: Methods of Information Measurement and Effects of Neuromodulation by Serotonin***

Soutenue le : Jeudi 19 Novembre 2020 à 14h00

*La Faculté des sciences, sur préavis du jury de thèse formé par :

Professeur titulaire R. JOLIVET,
Département de physique nucléaire et corpusculaire, directeur

Professeure A.-L. GIRAUD,
Faculté de médecine, Département des neurosciences fondamentales

Professeur A. LITKE,
Institute for Particle Physics, UC Santa Cruz, USA

Docteur L. KOSTAL,
Institute of Physiology, Czech Academy of Sciences, Prague, Czech Republic

autorise l'impression de la présente thèse, sans prétendre par là émettre d'opinion sur les propositions qui y sont énoncées*.

Genève, le 17 décembre 2020

Le Doyen

Thèse No - 282 -

N.B. La thèse imprimée doit porter la déclaration précédente * et remplir les conditions énumérées dans les « Informations aux étudiants relatives aux thèses de doctorat à l'Université de Genève ».

Brain: an apparatus with which we think we think.

-*Ambrose Bierce*

Acknowledgements

During the four years of my PhD thesis, I had the chance to be helped, supported and taught by colleagues, collaborators, friends and family. I would like to take the opportunity to whole-heartedly thank all those people that helped me go through with this thesis.

First of all, I would like to thank Renaud Jolivet for agreeing to be my thesis supervisor, and for his encouragement, trust and advice during these four years. He shared his knowledge and experience with me, offered me the opportunity to travel to present my work in conferences across the world, and made me a better scientist. I am very grateful that he gave me this opportunity.

A big thank you also to Taro Toyoizumi for having me as an intern in his lab for a few months. We had many great talks about Neuroscience, and all I have learned during this internship has helped me tremendously for my thesis. I thank also all the members of his lab in the RIKEN Brain Science Institute, and specifically Roberto Legaspi and Lukasz Kusmierz, for their very warm welcome and their help during my internship.

I also kindly thank the members of my thesis jury, Anne-Lise Giraud, Lubomir Kostal and Alan Litke, for taking the time the read and evaluate my work and for the very interesting questions and discussion that followed my defense.

Un grand merci à David, qui, en plus de souffrir avec moi de la magie noire qui rôde dans le bureau 202, a été un ami précieux et la plus belle rencontre faite durant ma thèse.

Merci infiniment à Guillaume, Marc et Philippe qui ont partagé mes repas de midi, mes pauses café et de nombreuses soirées en dehors de l'université. Ils m'ont écoutée râler quand ça n'allait pas, m'ont soutenue quand j'en avais besoin et ont été une merveilleuse présence quotidienne. Merci aussi à Freda, Loïc, Ralph et tous les autres collègues et amis qui ont partagé des moments avec moi pendant ces quatre années.

Un grand merci aux merveilleux amis que j'ai eu la chance de rencontrer depuis que suis installée à Genève: Anouck, Nadine, Dimitri, Jipé, Alex, Noémi, Eliott, Robin, Aloys, Sam, et les autres membres du Groupe de Nous. Merci pour votre présence, votre soutien et votre bonne humeur.

Je remercie du fond du cœur mes parents, Rémy et Gisèle, ainsi que ma sœur Carine et mon frère Alain. Merci d'être depuis toujours des soutiens à toute épreuve, d'avoir partagé mes bons et mes mauvais moments et de m'avoir menée là où je suis aujourd'hui. Je pense sincèrement que je ne serais pas arrivée au bout de cette thèse sans votre présence. J'ai aussi une pensée toute émue pour ma Granny et mon Grand-Papa, qui nous ont malheureusement quitté avant de pouvoir voire cette thèse, mais qui n'ont jamais cessé de m'encourager à suivre ma voie.

Et finalement, je remercie infiniment Christian d'avoir été mon pilier durant toutes ces années, de m'avoir supportée, confiné avec moi alors que j'écrivais cette thèse, et qui m'a abandonné son bureau sans hésiter pour préserver mon dos. Merci de toujours croire en moi, merci d'être là quoi qu'il arrive, merci pour tout.

Résumé

Le cerveau est l'une des structures les plus complexes de notre organisme. Chez l'homme, même s'il s'agit d'un relativement petit organe (environ 2% de la masse totale du corps), le cerveau consomme par lui-même environ un cinquième de toute l'énergie produite par le corps. La quantité totale d'énergie disponible pour le cerveau est cependant limitée. Ainsi, la consommation d'énergie est un paramètre essentiel du fonctionnement du cerveau, et de nombreuses études ont démontré qu'une altération des ressources énergétiques de cet organe peut être en lien avec diverses maladies et troubles neurologiques. Ces contraintes énergétiques ont aussi façonné la manière dont le cerveau fonctionne.

Les synapses, les structures cellulaires entre les neurones responsables de transmettre les signaux d'un neurone à l'autre, sont le lieu où la majorité de l'énergie du cerveau est consommée. Étonnamment, la transmission de signaux par les synapses n'est pas nécessairement fiable. En effet, dans le système nerveux central, la probabilité qu'un signal électrique soit transmis entre deux neurones est estimée entre 25 et 50%. Cela signifie que les neurones ne transmettent pas toute l'information qu'ils reçoivent, probablement car cela coûterait trop d'énergie à l'organisme. A la place, de précédentes études ont démontré que certains neurones maximisent le rapport entre l'information qu'ils transmettent et l'énergie utilisée pour transmettre cette information. Ce phénomène est appelé l'efficacité énergétique de la transmission d'information.

Pour pouvoir étudier l'efficacité énergétique, il est important d'être capable d'estimer de manière fiable l'information transmise par les neurones. Bien que développée dans le but d'être utilisée en sciences de la communication, la théorie de l'information est aussi largement utilisée en neurosciences et est très utile dans ce cas de figure. Cependant, il est parfois difficile d'utiliser cette théorie sur des données biologiques. En effet, la quantité de données collectées lors d'expériences est souvent limitée (que cela soit une expéri-

ence avec un animal ou avec des cellules *in vitro*) alors que la théorie de l'information nécessite souvent une quantité de donnée élevée pour éviter toute forme de biais. Il peut aussi parfois être compliqué de binariser ces mêmes données biologiques.

Cette thèse de doctorat a pour but de mieux comprendre l'efficacité énergétique de la transmission d'information, en se focalisant sur les neurones du système visuel. L'approche choisie est une approche computationelle, où des données d'entrée sont générées selon une certaine statistique et des modèles de transmission synaptiques sont utilisés pour générer les données de sortie. Trois études sont incluses dans cette thèse. La première se concentre sur la consommation d'énergie dans le cerveau et les découvertes expérimentales liées à l'efficacité énergétique. La deuxième étude compare la performance de deux mesures développées en théorie de l'information et adaptées aux neurosciences pour mesurer l'information transmise par les neurones: l'information mutuelle et l'entropie de transfert. Ces deux mesures sont appliquées avec des méthodes de correction visant à réduire les biais et elles sont comparées selon la taille des données générées et la complexité du modèle utilisé. La deuxième étude a aussi pour but de reproduire certaines des découvertes expérimentales concernant l'efficacité énergétique dans le noyau géniculé latéral (dans le thalamus), à l'aide d'un modèle de type Hodgkin-Huxley pour simuler la transmission synaptique. La troisième étude ajoute à ce modèle un moyen de modéliser la dépression synaptique, afin d'examiner les effets, sur l'efficacité énergétique, de la neuromodulation induite par la sérotonine aux synapses des cellules thalamiques relai.

La comparaison entre l'information mutuelle et l'entropie de transfert montre que leurs performances respectives dépendent du type de données simulées ainsi que des propriétés de transmission synaptiques considérées entre les données d'entrée et les données de sortie. Pour les cas simples, généralement peu représentatifs du fonctionnement d'un neurone, l'étude montre que l'entropie de transfert souffre moins des biais induit par les petits jeux de donnés que l'information mutuelle, même en utilisant des méthodes de correction. Une simulation plus complexe, basée sur un modèle de type Hodgkin-Huxley, montre que l'entropie de transfert est plus sensible aux propriétés de transmission, la rendant ainsi plus difficile à utiliser, alors que l'information mutuelle (avec les méthodes d'estimation utilisées) est robuste face à ces problèmes. L'étude montre aussi que dans les cas complexes, l'information mutuelle et l'entropie de transferts ne mesurent plus nécessairement la même quantité.

L'étude de l'efficacité énergétique avec notre modèle montre que le neurone modélisé (basé sur les propriétés des cellules thalamiques relai) maximise bien le ratio entre information et énergie correspondante, en accord avec les études expérimentales. L'étude montre aussi que, quand la neuromodulation induite par la sérotonine a lieu, l'information et l'énergie sont réduites, mais

que le ratio entre les deux est toujours maximisé par le neurone.

Les résultats obtenus durant cette thèse sont d'une grande aide pour mieux comprendre la transfert d'information à travers les synapses. La comparaison entre les deux mesures de théorie de l'information étudiées (information mutuelle est entropie de transfert) offre de très utiles aperçus des limitations de ces métriques et il est permit d'espérer que ces résultats seront utiles pour aider les neuroscientifiques à concevoir leurs expériences. De plus, les différents résultats obtenus lors de l'étude de l'efficacité énergétique du transfert d'information indiquent que ce concept pourrait être un principe général dans le cerveau, même si d'autres études sont encore nécessaires pour comprendre comment les neurones atteignent cet état. Dans ce sens, le modèle développé pendant cette thèse sera un excellent outil pour étudier plus en détail l'efficacité énergétique et la transmission synaptique.

Summary

The brain is one of the most complex structure in our organism. In humans, even though it is quite a small organ (it represents around 2% of the whole body mass), it is responsible by itself for around one fifth of the whole-body energy consumption. Still, the total amount of energy available to the brain is limited. Energy consumption is thus an essential feature of the brain's function, and studies have shown that alterations to the brain's energy supplies are linked to many diseases and neurological disorders. These energy constraints have also shaped how the brain works.

Synapses, the part of the neurons where signals are transmitted to other neurons, are where the majority of the energy of the brain is consumed. Surprisingly, signal transmission across synapses is not necessarily reliable. In the central nervous system, the probability of signal transmission between two neurons is estimated to be around 25 to 50%. This means that those neurons do not convey all the information they receive, probably because this would cost too much energy to the organism. Instead, previous studies have shown that some neurons work at an energetically efficient level for information transmission, maximizing the ratio between information conveyed and energy used doing so.

In order to study energetic efficiency, it is important to be able to reliably estimate the information transmitted by neurons. Although it was developed to be used in communication sciences, information theory is also widely used in Neuroscience. It can however be difficult to use this theory on biological data. The amount of data collected during experiments is usually limited (the experiment being conducted on an animal or directly on *in vitro* cells) and limited data can lead to biases when information theory is applied. It can also be complicated to binarize those same biological data.

This PhD thesis aims to better understand energetic efficiency of information transfer, focusing on neurons of the visual pathway. The approach used is a computational approach, where input data are generated with a

given statistic and models of synaptic transmission are used to generate output data. Three studies are included in this thesis. The first study focuses on energy consumption in the brain and experimental findings of energetic efficiency. The second study compares the performances of two metrics developed in information theory and adapted in Neuroscience to measure information conveyed by neurons: mutual information and transfer entropy. Those two metrics are used with correction methods designed to reduce the biases and they are compared according to the size of the dataset generated, as well as the complexity of the model used. The second study also aims at reproducing experimental findings about energetic efficiency in the lateral geniculate nucleus (in the thalamus), with the help of a Hodgkin-Huxley-type model to simulate synaptic transmission. The third study adds to this model a way of modelling modulation of synaptic depression in order to investigate the effects of neuromodulation induced by serotonin at thalamic relay synapses on energetic efficiency.

The comparison between mutual information and transfer entropy shows that their respective performance depends on the type of data simulated, as well as the transmission properties considered between inputs and outputs. For simple cases, usually not representative of how a neuron operates, the study shows that transfer entropy suffers less from biases induced by small datasets than mutual information, even with the correction methods used. A more complex simulation, based on a Hodgkin-Huxley-type model, shows that transfer entropy is more sensitive to transmission properties, making it more difficult to apply and that mutual information (at least with the estimation method used) is robust to this kind of issues. The study also shows than in complex cases, mutual information and transfer entropy do not necessarily measure the same quantity.

The study of energetic efficiency with our model shows that the modelled neuron (based on thalamic relay cells properties) indeed maximize the ratio between information and concomitant energy consumption, in accordance with experimental studies. It also shows than, when neuromodulation induced by serotonin occurs, the information transmitted by the neuron, as well as the energy used doing so decrease, but that the ratio between the two is still maximized by the neuron.

The results obtained during this thesis are of great help to better understand information transfer across synapses. The comparison of the two metrics of information theory studied (mutual information and transfer entropy) offers very useful insights on the limitations of those metrics and will hopefully be helpful for neuroscientists when designing experiments. Moreover, the different results obtained when studying energetic efficiency of information transfer at synapses indicate that this concept could be a generic principle in the brain, even though further studies will be necessary in order to understand how this is achieved by neurons. In this sense, the

models and tools build during this thesis will be of great use to study in details energetic efficiency in neural networks.

Acronyms

5-CT	5-carboxytryptamine
5-HT	5-hydroxytryptamine
AMPA	α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid
ATP	adenosine triphosphate
CNS	central nervous system
EPSC	excitatory postsynaptic current
GABA	γ-aminobutyric acid
ISI	interspike interval
LGN	lateral geniculate nucleus
NMDA	N-methyl-D-aspartate

Contents

CHAPTER 1

General Introduction

1.1 Aims of the study

Brain function has been shaped throughout evolution by energetic constraints. The brain is a relatively small and lightweight organ, but it consumes a big portion of the energy produced by the body. Importantly, the energy the body can allocate to the brain is limited and this has led to notable features in the functioning of the brain, one of them being the fact that some synapses do not maximize the information they convey, but maximize instead the ratio between the information they convey and the energy they use doing so: the so-called energetic efficiency of information transfer. A brief discussion of these issues has been published as Conrad *et al.* in the proceedings of IEEE IEDM 2017 (published in 2018; see Paper 1 in Chapter 2).

The present thesis aims to better understand how energetic efficiency in the brain works and is modulated. For this purpose, the first point addressed is to better characterize how the information conveyed by synapses can be measured, by comparing the advantages and trade-offs of two information theory metrics: the mutual information and the transfer entropy. Those two metrics were compared on different types of modelled spike trains, from very simple synthetic spike trains to more biologically realistic spike trains. These metrics were then used to calculate the energetic efficiency of information transfer in relay neurons in the lateral geniculate nucleus (LGN). This work is currently under review. A preprint can be found at `https://www.biorxiv.org/content/10.1101/2020.06.01.127399v1` (see Paper

2 in Chapter 2).

This work also aims to better understand the effect neuromodulation can have on synapses by evaluating its effect on the energetic efficiency of information transfer. Specifically, we are interested in better understanding what is the effect of serotonergic modulation of paired-pulse depression on energetic efficiency of information transfer in the LGN, using the same metrics of information theory. Preliminary findings on this work will be published in the proceedings of ICANN (international Conference on Artificial Neural Networks) 2020 (see Paper 3 in Chapter 2).

The next section introduces the literature relevant to the subjects studied in this thesis. We will start by discussing energy consumption in the brain, how energy is used (and by which processes) and stored, and what is the interplay between presynaptic release probability, postsynaptic gain, energy consumption and energetic efficiency of information transfer at synapses. We will then introduce information theory and the metrics studied in this work and their application in Neuroscience. Lastly, we will overview the role and effects of neurotransmitters, focusing on serotonin and synaptic depression.

1.2 Introduction

1.2.1 Energy consumption in the brain

The brain is a very complex structure, and the human brain has evolved to become what is probably the most sophisticated computational device we know. This organ is also one of the main consumers of the energy produced by the body. The human brain, for example, represents approximatively 2% of the whole body mass, but it consumes by itself around 20% of the oxygen needed by the resting body in adults [1]. The vertebrate brain is also the one that consumes the most oxygen relative to the total body needs [1].

In humans, the average power consumption of an adult is about 100 W [2], which translates to the average power consumption of the brain (20%) being around 20 W. We will discuss later in this introduction how the brain is not necessarily extremely reliable when it comes down to information transmission between neurons, but let us first try to better understand what those 20 W represent.

In March 2016, AlphaGo, a computer program using artificial neural networks, beat Lee Sedol (9th dan) 4:1 at Go, a Chinese strategic board game, which is considered to be the most computationally intensive game ever created. It took no less than 1920 CPUs and 280 GPUs [3] to beat one of the best players of Go in the world. It was estimated than the total power consumption of AlphaGo was around 1 MW [4], 50'000 times the consumption of the human brain. It is important, however, to note that the only thing AlphaGo can do is to play Go, or as Lee Sedol put it after the

game [5] *"[...] robots will never understand the beauty of the game the same way that we humans do".*

The brain is a very powerful computational device, and this was achieved through evolution. As humans evolved from primates, their cranial capacity increased approximatively 4-fold, which also caused an increase in energy use [1, 6]. The question of what allowed humans to have larger brain is still debated. One of the main hypothesis is that cooking was what allowed human encephalization [6, 7], as cooked food is easier to masticate and to digest than raw food, thus potentially delivering more energy to the body and the brain from the same amount of foodstuff extracted from the environment. Another hypothesis is the use of bi-pedal locomotion by humans, as it consumes less energy when displacing than quadrupedal locomotion [8]. Nonetheless, a lot of energy is needed to fuel the modern human brain, but the amount of energy the body can supply is limited. This most likely shapes information processing in the brain.

During adulthood, brain energy use is mainly stable, but this is not the case during development [9]. Before adolescence, the energy use of the brain increases, in relation with the thickening of the cortex (as many synaptic connections are made). From adolescence onwards, energy use constantly decreases in correlation with cortical thinning (as connections are pruned) until the brain reaches its mature state, between 20 and 30 years of age [10–12].

Energy consumption at synapses

As mentioned above, the brain itself is responsible for about 20% of the whole-body energy consumption, but what exactly is responsible for this high energy consumption in the brain?

Adenosine triphosphate (ATP) is the main store of chemical energy in the brain (and in living matter in general) [13]. In 2001, Atwell and Laughlin estimated the ATP consumption for the main subcellular processes for signalling in rat cerebral cortex [14] (their main results can be seen in Figure 1.1). In 2003, Lennie performed a similar study for the human cortex [15]. In both these studies, estimates are based on a bottom-up detailed biophysical budget, considering the rate of ATP consumption of every process consuming energy in the brain. Interestingly, a study by Jolivet *et al.* published in 2009, developing instead a top-down budget approach to calculate the brain's energy budget, where the total average glucose and oxygen utilization of the brain was used to estimate the average ATP production of processes, reached similar conclusions with respect to the relative contribution of various ATP-consuming processes [16]. It is nonetheless important to note that those different budgets are based on many hypotheses and the results should be considered with caution.

Those papers showed that the main processes that use ATP in the brain are action potentials, the resting potentials of neurons and glia and synaptic transmission (see Figure 1.1). The processes mediating synaptic transmission are, by themselves, responsible for about 60% of the total energy use [17] (highlighted in red in Figure 1.1). Figure 1.1 also shows that the great majority of those 60% is used to reverse ion movements generating postsynaptic responses through cellular membranes, even though a lot of other processes are necessary for synaptic transmission. Housekeeping energy (*i.e.* all the biochemical processes that maintain the infrastructure of the cell), is not taken into account in this budget.

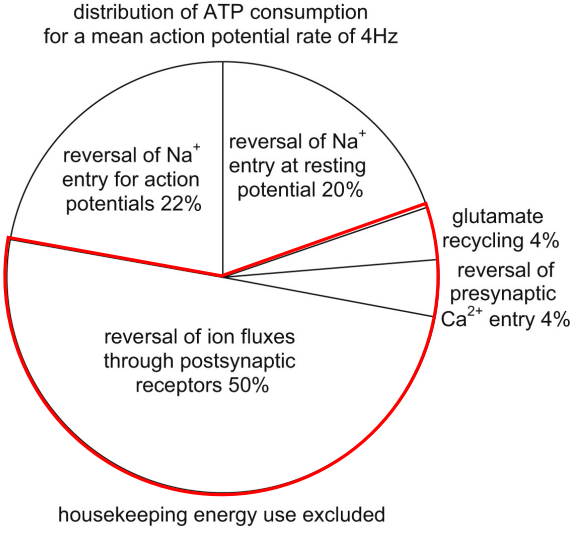

Figure 1.1: Percentage of adenosine triphosphate (ATP) predicted to be used on subcellular mechanisms underlying signalling in the rat cortex [14] . The three main processes consuming ATP are reversal of ions entry at resting potential (20%), reversal of ions entry for action potentials (22%) and synaptic transmission (58%, highlighted in red). Adapted from reference [17].

Figure 1.2 summarizes all the processes underlying signalling occurring at synapses. Let us focus on the processes consuming ATP. Presynaptically, those processes are the Na^+/K^+-ATPase (restore the sodium gradient and drives the Na^+/Ca^{2+} exchanger responsible of calcium removal), the Ca^{2+}-ATPase (lowers internal calcium concentration), the vacuolar H^+-ATPase (energizes transmitters uptake by vesicles), the endocytosis of empty vesicles and motor proteins (move mitochondria and vesicles inside the cell). Postsynaptically, those processes are the Na^+/K^+-ATPase, the Ca^{2+}-ATPase,

returning Ca^{2+} to intracellular stores and motor proteins. In astrocytes, those processes are restoration of the sodium gradient through the Na^+/K^+-ATPase and the conversion of glutamate into glutamine.

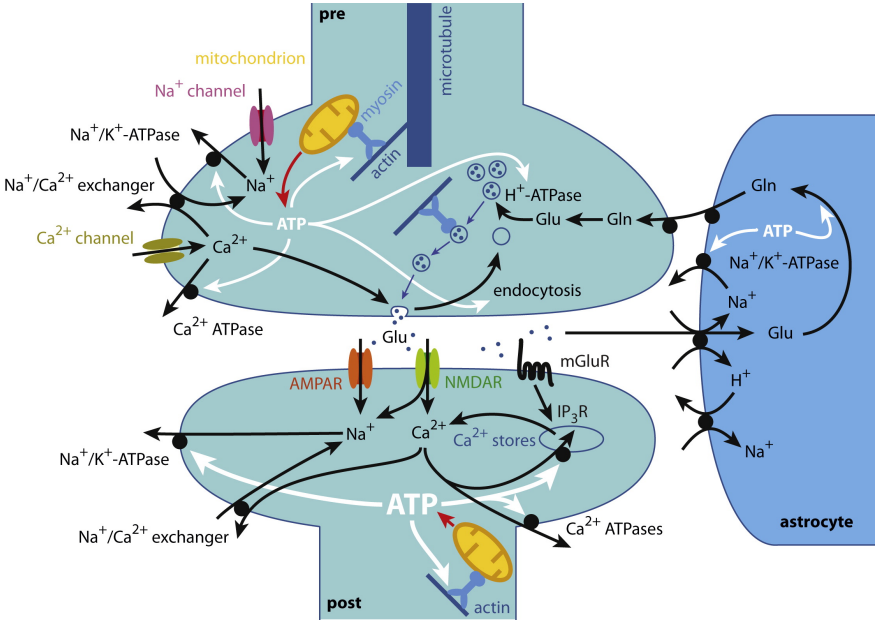

Figure 1.2: Schematic representation of mechanisms underlying signalling at synapses. Post- (bottom) and presynaptic (up) sides of the synapse are shown, as well as an astrocyte (right). ATP consumption is shown with white arrows (the thicker the arrow, the larger is ATP use). Presynaptically, ATP is used by the Na^+/K^+-ATPase, the Ca^{2+}-ATPase, the vacuolar H^+-ATPase, endocytosis of vesicles and motor proteins. Postsynaptically, ATP is used by the Na^+/K^+-ATPase, the Ca^{2+}-ATPase, returning Ca^{2+} to intracellular stores and motor proteins. In astrocytes, ATP is used by the Na^+/K^+-ATPase and conversion of glutamate into glutamine. Some processes that do not consume ATP are also shown. Adapted from reference [17].

ATP and mitochondria

The ATP used by the adult brain is almost entirely generated by the glycolysis, followed by complete oxidation of glucose in oxidative metabolism. Consequently, mitochondria provide about 93% of the ATP used in the brain [18]. The fact that the majority of the energy is used by synapses can also be verified via the physical localization of mitochondria, that matches loci of high energy consumption [19, 20]. On average, there is 1 mitochondria on either side of most synapses (see [21] for the presynaptic terminals and [22] for postsynaptic terminals). ATP-producing mitochondria are drawn in yellow

in Figure 1.2.

An increase of neuronal activity will result in an increase of energy usage, and most of the surplus ATP necessary in this case is generated by mitochondria [23, 24]. In order to sustain neuronal activity, it is essential that ATP is made available to the cells when necessary. This means that correct location of functioning mitochondria is essential, both for the development of the nervous system, and for its plasticity and health. Mitochondria are formed at the soma. It is thus necessary that they can travel rapidly and efficiently to the place where they are needed. They are transported long-distance along microtobules, driven by kinesin and dynein motors with a speed of between 0.3 and 1 $\mu m/s$ [25, 26]. For short distances, they travel along actin filaments, driven by myosin motors [27]. However, movement of mitochondria costs ATP, meaning that there are energetic limitations to their movement. This might explain why mitochondria are stationary 80% of the time [17].

Changes in energy use and diseases

As mentioned above, the brain use in energy is mainly stable during adulthood. Nonetheless, synaptic plasticity can alter energy expenditure because it can change synaptic strength, as can be seen in Figure 1.3A. For example, long-term potentiation can increase the number of AMPA (α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid) receptors in postsynaptic membranes, which increases the strength of synapses (green synapse in Figure 1.3A) but also their energy consumption [28]. The inverse is also possible, when a negative feedback mechanism adapts to low levels of cellular energy supply by stopping the maintenance of the synapse [29] (red synapse in Figure 1.3A). Synaptic plasticity can also help synaptic transmission save energy. An example of this is the case of silent synapses. These appear when long-term depression leads to synapses producing zero postsynaptic current, as observed by Isope and Barbour in cerebellar parallel fibers connecting to Purkinje cells [30]. These silent synapses were predicted theoretically in 2004 by Brunel *et al.* [31] for optimal storage of information. A few years later, in 2009, Howarth *et al.* [32] showed that they massively reduce the amount of energy used synaptically.

For stable energy use in the brain, it is also important that depleted energy levels are restored when necessary, or possible. There are a lot of reasons to think that sleep helps to restore energy in the brain [33]. Physiological evidence of this was found by Dworak *et al.* [34]. They reported differences between energy consumption during awake and sleep states. If cells use a lot of ATP during awake periods, for electrical and chemical signalling at synapses, then it is only during sleep that ATP can be allocated to other tasks in order to fuel restorative biosynthetic processes.

Sleep is probably also the time when the brain adjusts synaptic strengths, and therefore ultimately also the energy expenditure of synapses, with the help of homeostatic plasticity [35], as can be seen in Figure 1.3B. During the awake state, synaptic plasticity leads to strengthening of synaptic connections (Figure 1.3B, left), and consequently, to an increase in energy usage. The signal-to-noise ratio is also reduced as synaptic weights can become saturated. The homeostasis hypothesis [36] proposes that sleep allows for the renormalization of synaptic strengths (Figure 1.3B, right), leading to a desaturation of synapses, and thus to an increase of signal-to-noise ratio. According to this hypothesis, sleep also allows restoration of energy supplies, and consolidation and integration of memories.

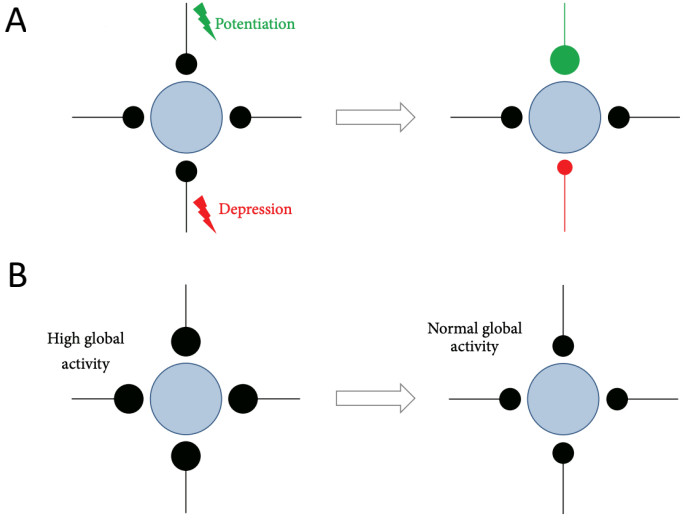

Figure 1.3: Schematic representation of synaptic plasticity and homeostatic plasticity effects on synaptic strength. Blue circles are postsynaptic cells and black lines and circles are synaptic inputs. Greater synaptic strength is represented with larger synaptic inputs. A: Synaptic plasticity induced by particular inputs to the neuron can lead to strengthening (potentiation; green) or weakening (depression; red) of synaptic connections. B: Homeostatic plasticity leads to proportional changes of the strengths of all synapses to a neuron. For instance, strong synapses are renormalized with homeostatic plasticity so that less energy is used and the signal-to-noise ratio is increased. Adapted from reference [37].

Synapses account for the majority of energy use in the brain. This means that disorders of mitochondria trafficking or function will affect synapses, and *vice-versa*. Mitochondria dysfunctions or abnormalities have been shown to play a role in several diseases including Parkinson's disease [38], Huntington's disease [39], Alzheimer's disease [40, 41], the motor neuron disease called familiar amyotrophic lateral sclerosis [42], and cerebral ischemia [43],

highlighting the importance of brain energy metabolism in these pathologies

An other interesting way to approach the link between brain's dysfunctions and energy is the free energy principle developed by Karl Friston [44, 45]. This principle, quite complex, is based on the minimization of the free energy, defined as the difference between the state a system expects to be in and the state its sensors tell it that it is in. In other words, this principle aims to reduce any form of surprise. The free energy principle can in theory be applied to any system, from a monocellular organism to a very complex system like the brain. According to this principle, a change in functioning and energy usage in the brain (either too much or too little) would be associated with a change of the state the brain's sensors can evaluate, potentially leading to neurological conditions implying hallucinations and delusions. For example, the free energy principle has been used to explain features of Parkinson's disease [46], addiction [47] and schizophrenia [48]. Nonetheless, it is important to note that the free energy considered in this principle is a concept wider than the concrete concept of energy considered in Biology in general.

Synaptic release probability and energetic efficiency

As mentioned above, synapses are responsible for a large fraction of the energy usage of the brain, but the total energy available to the brain is limited. As studies in Anthropology and Neuroscience have showed, energy constraints are probably one of the key factors that have shaped the evolution of the human brain [6], but also the evolution of the brains of other species [49]. It is also likely that energy constraints have shaped how the brain processes information.

Information is conveyed from one neuron to another when an action potential reaches the synapse, activating the release of vesicles of glutamate (or other neurotransmitters) in the synaptic cleft, as can be seen in Figure 1.2. Glutamate then docks to the corresponding receptors of the postsynaptic neuron, triggering an excitatory postsynaptic current (EPSC) that can lead to the generation of an action potential [50]. One could expect that a vesicle is released every time an action potential reaches the presynaptic terminal, but, surprisingly, in the central nervous system (CNS), the typical vesicle release probability at synapses is around 0.25-0.5 [14]. This low release probability offers several advantages to synapses: allowing efficient computation for time-varying signals, increasing information transmission from correlated inputs and maximizing the storage of information [51–53]. More in context with the present work, the low release probability can also have energetic advantages.

Firstly, synaptic failure is essential to reduce energy wastage [54]. Let us consider a dendritic tree receiving signals from several synapses, for

example a cortical neuron receiving inputs from around 8000 presynaptic terminals [55]. In this case, the rate at which information arrives from all synapses is greater than the rate at which the output axon can convey information. Energy would then be wasted if information that cannot be passed on by the postsynaptic cell was transmitted. A low release probability is thus essential to avoid this waste of energy.

Moreover, most axons create more than one release sites to a postsynaptic cell (for example in the cortex, it is usually between 4 and 6) [56], each with a given probability of releasing a vesicle. Having more than one release site increases the reliability of signal transmission and the fraction of the input information that is received [56–58]. Several release sites in parallel (often on different spines) also allow stable network connectivity even in the case of spine turnover [59], because there is a response in the postsynaptic cell when it receives at least one synaptic current.

This is not the only advantage of several release sites. Indeed, in the case of several release sites, the ratio of the information transmitted per ATP molecules needed is maximized for low release probability [14]. This means that the energetic design of synapses is optimized for low neurotransmitters release probability. As suggested by Levy and Baxter [60], the nervous system does not maximize the rate of information transfer, but the ratio of transmitted information to energy consumed. This phenomenon is called, or referred to as, energetic efficiency of information transfer [61, 62].

Levy and Baxter were actually the first to theorize this principle, when they studied synaptic firing rates [60]. As we will develop in Section 1.2.2, information theory [63] is very helpful in Neuroscience and is used abundantly to characterize information flow in neural systems. By using information theory, Levy and Baxter have argued that the firing rate that maximizes the rate at which information arrives in an input train is around 200 Hz [60]. However, the mean firing rate of *in vivo* neurons is around 4 Hz, or much lower [14, 64]. This low firing rate *in vivo* means that the nervous system could convey more information (by firing more action potentials), but this would have a higher energy cost

Figure 1.4 shows a theoretical example based on reference [17] to help better understand how the information transmitted and the ratio between information and energy vary with the number of release sites N and the release probability p. At first, let us focus on the black curves. In this example, the input firing rate is 4 Hz, which corresponds to a probability $s = 0.01$ of an action potential arriving in any given interval $\Delta t = 2.5$ ms. Figure 1.4A shows the fraction of information transmitted as a function of the release probability p. The information transmitted increases sigmoidally with p (once all the information has been transmitted, it is impossible to transmit more, even if p increases). In this example, the energy needed to convey this information is calculated as $N \cdot s \cdot p$, the number of vesicles released

per Δt, assuming that energy use is proportional to the number of vesicles released. Figure 1.4B shows then the ratio between transmitted information and energy use. As it can be seen, for $N > 1$ the maximum of the ratio is at $p = 0$, which is not very realistic.

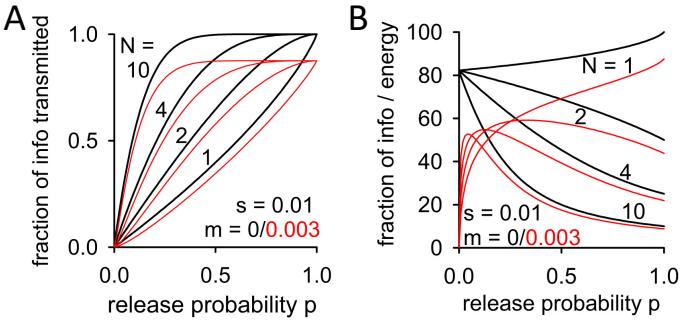

Figure 1.4: Theoretical evaluation of the energetic efficiency as a function of the release probability. A: Fraction of information arriving from an axon and transmitted across N release sites, each releasing one vesicle with a probability p, in the absence (black) or presence (red) of spontaneous release. B: Ratio of fraction of information transmitted (from A) to postsynaptic energy used, taken as $N \cdot s \cdot p$ (the number of vesicles released by action potentials per Δt), in the absence (black) or presence (red) of spontaneous release. In both panels, s is the probability of an action potential arriving in any given interval Δt and spontaneous release occurs in the whole cell with a probability m per Δt. $\Delta t = 2.5$ ms. Adapted from reference [17].

This example is not very representative of neurons as it is lacking one important features: spontaneous release. Synapses can spontaneously releases a vesicle, even without an incoming action potential. This produces a postsynaptic effect indistinguishable from the one triggered by an action potential [17]. The red curves in Figure 1.4 show the case with spontaneous release at a rate of 1.2 Hz. This is equivalent to a probability $m = 0.003$ that the release occurs per $\Delta t = 2.5$ ms. Spontaneous release from all input synapses reduces the maximum amount of information transmittable (see Figure 1.4A), because some postsynaptic currents will not be caused by presynaptic action potentials, thus creating noise. In the case where the frequency of action potential-evoked release probability is close to the one of spontaneous release, the ratio of information transmitted per energy used decreases (less information is transmitted but more energy is needed). A typical physiological spike rate of around 4 Hz combined with a spontaneous release rate of 1.2 Hz (as measured in cortical pyramidal cells by Dani *et al.* [65]) would give a release probability that maximizes the ratio information over energy of around 0.05-0.25 depending on the number of release sites, as shown in Figure 1.4B.

As can be seen in Figure 1.4B, the maximum of the ratio between information and energy emerges for $p \neq 0$ only in the case where spontaneous release is taken into account, but also only for more than one release site ($N > 1$). Figure 1.4B also shows that the optimal release probability decreases if the number of release sites increases. Interestingly, similar results have been obtained in experimental studies, comparing the release probability with the number of release sites. A correlation between the number of release sites and the release probability was found in cortical synapses [66] (see Figure 1.5A) and in hippocampal synapses [67] (see the closed circles in Figure 1.5B). Figure 1.5B also shows that, for these synapses, the release probability adapts only to the number of synapses made by the presynaptic cell on the same dendrite, while no correlation is observed when the synapses are made on different dendrites (open circles).

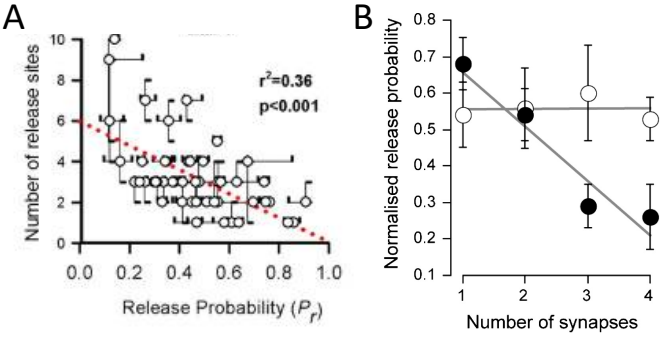

Figure 1.5: Experimental correlation between the number of release sites and the release probability. A: Correlation of the number of release sites with the release probability for pyramidal neurons in layer 2/3 of the rat visual cortex. Adapted from reference [66]. B: Correlation of the release probability with the number of release sites for rat hippocampal neurons. The closed circles represent the number of synapses that one axon makes on the same dendritic branch (these are the release sites discussed in the Introduction). The open circles represent the number of synapses that one axon makes on different different dendrites (in this case, no correlation is observed with the release probability). Adapted from reference [67].

There also have been experimental studies on the energetic efficiency level, and they have shown that some neurons do apply this (or a similar) principle of energetic efficiency with respect to how much information is conveyed. Figure 1.6 shows the results obtained by Harris and colleagues at synapses of thalamic relay neuron in the lateral geniculate nucleus (LGN; a nucleus in the thalamus) in rat brain slices [68]. Figure 1.6A shows the circuitry of those relay neurons. Thalamic relay cells are part of the visual pathway. They receive signals from the retina, with a one-to-one connection, and then send their signals to the cortex. This circuit is thus essentially feed-forward. As shown by the grey arrows in Figure 1.6A, modulation by

descending inputs from the cortex and the local inhibitory network in the thalamus were removed or inactivated (see reference [68] for details).

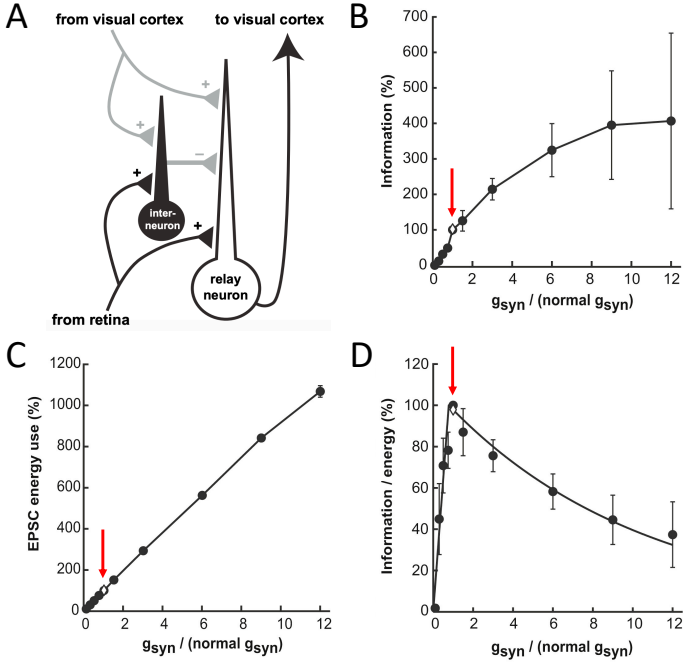

Figure 1.6: Measure of energetic efficiency of information transfer when experimentally modulating the postsynaptic gain of synapses of relay neurons in the lateral geniculate nucleus (LGN) in rat brain slices. A: Circuitry of the LGN. Relay neurons in the LGN receive signals from the retina and send signals to the visual cortex. The circuit is a feed-forward circuit. The grey arrows indicate the neglected pathways (descending inputs from the cortex and the local inhibitory network in the thalamus). B: The information measured at those synapses increases sigmoidally with the synaptic gain (the size of EPSCs). C: The energy used to convey the information from B increases linearly with the synaptic gain. D: The ratio between information (from B) and energy (from C) is a monophasic curve with one unique maximum at the physiological postsynaptic gain of the synapse. The red arrows point to the original physiological gain of the synapse. Adapted from reference [68].

The difference between these experimental results and the theoretical example showed in Figure 1.4 is that the experimental results shown in Figure 1.6 are plotted against the normalised synaptic gain g_{syn}, not the release probability. The synaptic gain is a measure of the number of glutamate receptors on the postsynaptic membrane, and, thus, of the postsynaptic peak conductance (*i.e.* the amplitude of the generated EPSCs, assuming that there is always enough glutamate released in the synaptic cleft to activate all the receptors). The release probability is a presynaptic property of a given

synapse, while the synaptic gain is a postsynaptic property. The results of Harris and colleagues shown in Figure 1.6 thus address the same question of energetic efficiency of information transfer but focus on the postsynaptic side of the synapse (number of postsynaptic receptors), while the example in Figure 1.4 focuses on energetic efficiency of information transfer with respect to the presynaptic side of the synapse (release probability).

Figure 1.6B shows that the information transmitted by this synapse increases sigmoidally with the postsynaptic gain, similarly to the theoretical case we considered earlier for the presynaptic release probability shown in Figure 1.4A. Figure 1.6C shows that the energy used to transmit the information increases linearly with the postsynaptic gain. The red arrows indicate the original physiological gain of the synapse (for a normalised synaptic gain of 1). Figure 1.6C indicates that this synapse does not transmit all the information it could, but transmitting more information would cost more energy to the system. This leads to the ratio between information flow and energy consumption to have a unique maximum (Figure 1.6D), this maximum sitting at the physiological gain of the cell (as shown with the red arrow). These results indicate that synapses impinging on LGN thalamic relay cells from the retina operate according to the energetic efficiency of information transfer principle: maximizing the ratio between information flow and energy consumption instead of simply maximizing the amount of information transferred.

As mentioned above, information transmitted by neurons is evaluated using information theory. Being able to correctly estimate the information is thus an essential task when studying energetic efficiency of information transfer. The next Section (Section 1.2.2) focuses on information theory and its applications in the cases this thesis is interested in.

1.2.2 Information theory

Information theory is a mathematical framework created to study the transmission of information through communication systems. It was primarily developed by Claude Shannon, with a first publication in 1948 [63]. This paper was focused on the transmission of information in the presence of noise in the channel of communication. It also introduced the concept of entropy, which we will describe in detail later in this Introduction. Soon after this publication, the scientific community recognized the significance and flexibility of this method and researchers started to apply information theory to fields outside of its original scope, for example in statistics, biology, behavioural science or statistical mechanics, and also, obviously, in Neuroscience [69]

The way in which neurons transmit information was debated at the time the paper from Shannon was published. Information theory thus appeared

to be the perfect tool to answer this question [70]. The first works using information theory in Neuroscience were published in the 50s and indeed showed that neurons are able to relay large quantities of information [71–73]. These works were the starting points for many subsequent studies trying to better characterize information flow in specific neural systems. Information theory has been used to address many different Neuroscience questions and in a number of different contexts. It has led to a number of advances and developments of theories for brain functions, and it is still used today in order to better understand how neural cells process information. The interested reader can see references [74], [75], [76] and [77] for reviews on the subject.

Information theory in Neuroscience

Information can have different meanings in different contexts [78]. In Neuroscience, information is often evoked when discussing stimulus encoding (information encoding), decision-making (information processing) or memory (information storage). Information of a neural variable is generally associated with a reduction in uncertainty of another variable or stimulus [79]. Information is measured in bits, which should allow for straightforward comparisons between cells, brain regions, tasks or subjects.

Data obtained from Neuroscience experiments are frequently noisy and often represent systems with non-linear interactions. They also can be composed of numerous variables (e.g. physiologic, behavioural or stimulation data) that interact. Information theory offers a number of multivariate analysis tools, it can be applied to different types of data, it can capture non-linear interactions and it is model-independent. Because of its general applicability, information theory is widely used in Neuroscience, for example on analysis of electroencephalography (EEG), magnetoencephalography (MEG), functional magnetic resonance imaging (fMRI) and trial-based data, or on studies on connectivity or sensory coding [77]. It can detect a wide range of interactions and structure, even in large or complex systems.

The results of information theoretic analysis can quantify the uncertainty of variables, the dependencies between variables and the influence some variables have on others [79]. It thus can be used to quantify encoding (how much information a neuron conveys) [80, 81], or complex encoding relationships (how much information two neurons convey together) [82, 83]. It is important to note that information theory can not produce models that describe how the system works (but it can be used to restrict the space of possible models) [77].

The following sections present the metrics of information theory used in the present thesis. Other tools to measure information exists, for more details on those, see for example reference [77].

Entropy

The concept of entropy is something physicists are utterly familiar with. Developed in Boltzmann's theories about statistical mechanics and thermodynamics, entropy measures the disorder in a system. The entropy used in information theory and developed by Shannon has a similar mathematical formulation but the reason why Shannon chose to also call it entropy is attributed to John van Neumann [84]: *" The theory was in excellent shape, except that he needed a good name for "missing information". "Why don't you call it entropy", von Neumann suggested. "In the first place, a mathematical development very much like yours already exists in Boltzmann's statistical mechanics, and in the second place, no one understands entropy very well, so in any discussion you will be in a position of advantage. "*

The entropy is the fundamental quantity in information theory. It measures the uncertainty contained in a variable. For a variable X with individual states x, the entropy $H(X)$ is defined as [63, 79]:

$$H(X) = \sum_{x \in X} p(x) log_2 \frac{1}{p(x)}. \tag{1.1}$$

This ensures $H(X) \geq 0$, as a negative uncertainty would not have any meaning. Moreover, systems with one absolutely certain state have $H(X) = 0$, as there is no uncertainty over their state. In this thesis, we will concentrate on discrete distributions, but it is interesting to note that there also exists an extension of entropy for continuous distributions [79].

For systems with two variables, X with states x and Y with states y, the joint entropy $H(X, Y)$ is defined as:

$$H(X, Y) = \sum_{\substack{x \in X \\ y \in Y}} p(x, y) log_2 \frac{1}{p(x, y)}. \tag{1.2}$$

The average uncertainty in a variable, given the state of another variable, is calculated with the conditional entropy $H(X|Y)$ (also called noise entropy) and is defined as:

$$H(X|Y) = \sum_{\substack{x \in X \\ y \in Y}} p(x, y) log_2 \frac{1}{p(x|y)}. \tag{1.3}$$

Conditional entropy can also be used to rewrite the joint entropy:

$$H(X, Y) = H(X) + H(Y|X). \tag{1.4}$$

Figure 1.7 gives a more visual interpretation of Equations 1.1 to 1.4, where it can be seen that the total area of $H(X, Y)$ is composed by the areas of $H(X)$ and $H(Y|X)$, or equivalently $H(Y)$ and $H(X|Y)$.

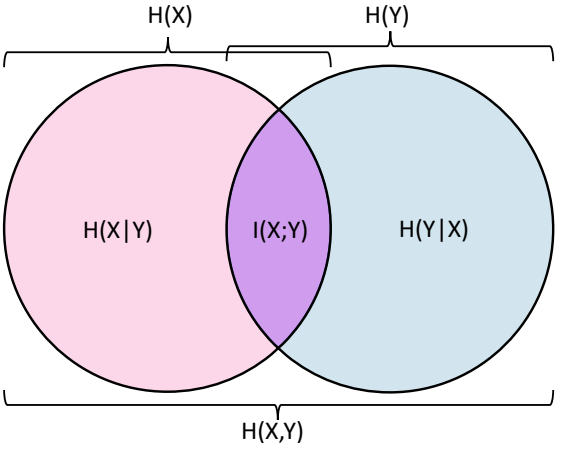

Figure 1.7: Venn diagram of entropy and mutual information. The entropies $H(X)$ and $H(Y)$ each are a circular area. The area formed with the combination of $H(X)$ and $H(Y)$ is the joint entropy $H(X,Y)$. The conditional entropy $H(X|Y)$ is the area formed by removing from $H(X)$ the part of $H(X)$ that intersects with $H(Y)$. The other conditional entropy $H(Y|X)$ is the area formed by removing from $H(Y)$ the part of $H(Y)$ that intersects with $H(X)$. The area at the intersection between $H(X)$ and $H(Y)$ is the mutual information $I(X;Y)$. It can also be constructed as the area of $H(X)$ minus the area of $H(X|Y)$, or equivalently $H(Y)$ minus $H(Y|X)$.

Mutual information

The entropy measures the uncertainty in one variable. If two variables are dependent from each other (for example the output spike trains of two connected neurons), then learning the state of one variable reduces the uncertainty in the other variable. This means that one variable will provide information about the other variable. This can also be written as:

$$H(X) = H(X|Y) + I(X;Y), \tag{1.5}$$

where $H(X|Y)$ is the entropy that remains in X, given knowledge about Y, and $I(X;Y)$ is the information provided by Y about X measured in bits. This is also illustrated in Figure 1.7. Equation 1.5 can be rewritten as:

$$\begin{aligned} I(X;Y) &= H(X) - H(X|Y) \\ &= \sum_{\substack{x \in X \\ y \in Y}} p(x,y) log_2 \frac{p(x,y)}{p(x)p(y)}. \end{aligned} \tag{1.6}$$

I(X;Y) is called the mutual information [79]. It corresponds to the reduction of uncertainty in X, given knowledge of the state of Y. $H(X) \geq 0$ implies that $I(X;Y) \leq H(X)$.

By symmetry, $I(X;Y)$ can also be rewritten as [79] (see also Figure 1.7):

$$\begin{aligned} I(X;Y) &= H(Y) - H(Y|X) \\ &= I(Y;X). \end{aligned} \tag{1.7}$$

This means that the information Y provides about X is the same as the information X provides about Y. Furthermore, if we combine Equation 1.7 with Equation 1.4, we obtain the alternate form:

$$I(X;Y) = H(X) + H(Y) - H(X,Y). \tag{1.8}$$

It is possible to expand the mutual information to take into account more than two variables [79]. An alternate form of mutual information that is very helpful for causal relations is the mutual information between two variables conditioned on a third variable, the conditional mutual information:

$$\begin{aligned} I(X;Y|Z) &= H(X|Z) - H(X|Y,Z) \\ &= \sum_{\substack{x\in X \\ y\in Y \\ z\in Z}} p(x,y,z) log_2 \frac{p(x,y|z)}{p(x|z)p(y|z)}. \end{aligned} \tag{1.9}$$

Transfer entropy

The transfer entropy was first introduced by Schreiber in 2000 [85] as a tool to measure the statistical coherence between systems evolving in time. It is defined as a conditional mutual information with assumptions about temporal order:

$$\begin{aligned} TE(X \to Y) &= I(X^-;Y^+|Y^-) \\ &= H(X^-|Y^-) - H(X^-|Y^+,Y^-) \\ &= \sum_{\substack{x^-\in X^- \\ y^+\in Y^+ \\ y^-\in Y^-}} p(x^-,y^+,y^-) log_2 \frac{p(x^-,y^+|y^-)}{p(x^-|y^-)p(y^+|y^-)}. \end{aligned} \tag{1.10}$$

This quantity measures the information about the future state of a variable Y^+ provided by another variable in the past X^- given the information provided by the past state of the first variable Y^-. The transfer entropy measures the changes caused in Y from X that cannot be derived only by the past states of Y, like it is the case for example if X was an inhibitory neuron impinging on neuron Y, i.e. the past of X is more informative about the future state of Y than the past of Y itself.

In a Neuroscience context, and in this thesis in particular, when applied to spike trains, transfer entropy measures the information flow between

neurons, and mutual information measures the encoding of stimulus and behavioural information shared by individual neurons. Mutual information is symmetric under the exchange of the variables. On the contrary, transfer entropy measures the direction of information flow and is, by definition, non-symmetric. Thus, transfer entropy can be seen as a conditional mutual information with assumptions about temporal structure. This means that depending on the temporal structure of the data, mutual information and transfer entropy could evaluate similar quantities. This is interesting, because even if the quantity evaluated is similar, the estimation methods will differ. The accuracy of those estimation methods seems then to be an important point to study.

Application and limitations of those metrics

The basic idea when conducting a Neuroscience experiment aiming to measure the information flow or encoding is to send a stimulus to a neuron (input) and to record the response of the neuron (output). Information theory and, specifically, the metrics introduced above, can then be used to calculate the information flow between the recorded input and output. But, to do that, the data used need to be discretized, even though biological processes are continuous. In a Neuroscience experiment, the recording on the computer of continuous biological processes ends up being discrete with some sampling rate determined by the experimenter and dependent on the equipment used. The recording is usually sampled at a higher frequency than necessary for data analysis with information theory, and must be downsampled by using binning or discretization methods [86]. The choice of the method used is important for a correct evaluation of the probability distributions. This is a very system-specific procedure, and a poor choice of discretization will lead to a wrong evaluation of the information theory metrics [77].

Estimating probability distributions is necessary to any information theoretic analysis, and, in an experimental context, these distributions must be determined from raw experimental data. In 1998, De Ruyter *et al.* developed a relatively straightforward way of estimating information carried by a spike sequence, as well as a new method for the application of information theory to neural spike trains [87]. This new method represented an advance from a methodological point of view and has led to many applications.

The issue one will encounter when evaluating the information from probabilities reconstructed from experimental data is the bias that appears due to a finite sample size. To calculate the information it is necessary to reconstruct the probability of each possible state (as it can be seen in the above Sections, every metrics introduced is based on probabilities). A state can be, for example, a word formed as a string of bins, such as 100 010 001 101 011 110 111 000 for words of length 3. The probability of a given word

is reconstructed by counting every occurrence of this word in the data and dividing this number by the total length of the data. But, if the dataset is not big enough, there is a risk to poorly evaluate those probabilities. For example, let us consider a very unlikely word, with a very small (but still bigger than zero!) probability of occurrence. In a small dataset, chances are that a very unlikely word is never observed. The probability of this state would thus be evaluated to be zero, an underestimation. Of course, because the sum of all the probabilities must, by definition, be equal to 1, this implies that the estimates of the other probabilities will also be incorrect, some of them will be overestimated.

This then propagates in the estimation of the entropies, resulting in the entropies being underestimated [88]. Intuitively, this bias can be understood because the entropy is a measure of variability in the dataset, which means that a smaller dataset is less likely to fully sample the full range of possible responses. A more mathematical understanding can be achieved with the help of Figure 1.8, where we can see that miss-evaluating very small probabilities will impact the evaluation of the function $h(x) = p(x)log\frac{1}{p(x)}$ the most. Indeed, a small probability underevaluated leads to a big error in the evaluation of $h(x)$ (as shown with the red arrows in Figure 1.8), as a higher probability overevaluated leads to a smaller error in $h(x)$ (see the orange arrows in Figure 1.8). This means that the entropy, calculated as $\sum_x h(x)$ (see Equation 1.1) will be underestimated. This graph also shows that the largest error will be for probabilities close to zero, which are the most difficult to correctly sample.

If we take the case of the evaluation of the mutual information (see Equation 1.6), the entropy and the conditional entropy will both suffer from this bias and be underestimated. Interestingly, the conditional entropy will often be underestimated by a larger amount that the entropy. The estimation of the entropy necessitates the evaluation of all the probabilities $p(x)$ (see Equation 1.1). For words of length L for instance, there are 2^L possible combinations of $p(x)$ that need to be evaluated. On the other hand, conditional entropy necessitates the evaluation of 2^{2L} possible combinations of $p(x|y)$ (see Equation 1.3). This means that in a limited dataset, $p(x)$ will be better sampled than $p(x|y)$, resulting in the conditional entropy being more significantly underestimated than the entropy. This usually results in the mutual information (entropy minus conditional entropy) being overestimated.

Poor evaluation of the probabilities will bias the estimation of entropies and mutual information, but enough observations can help adequately sample the space of states and reduce this bias [89]. This is however not always possible from an experimental point of view. To alleviate this issue, several correction methods for biases have been developed. The first theoretical (as in non-empirical) correction method was developed by Treves and Panzeri [90]

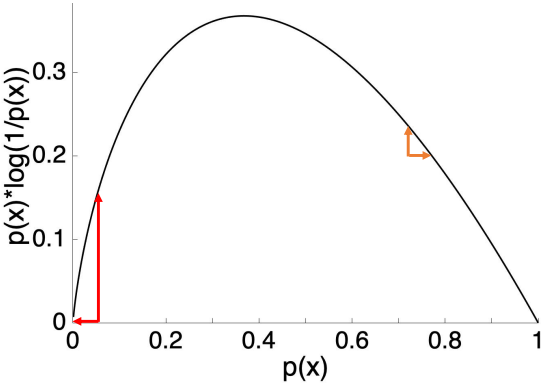

Figure 1.8: Plot of the function $h(x) = p(x)log\frac{1}{p(x)}$. The arrows show two examples were the probabilities are slightly miss-evaluated. The red arrows show a small probability (> 0) evaluated to be zero (underevaluation), leading to a big impact in the evaluation of $h(x)$. The orange arrows show a much larger probability slightly overevaluated, leading to a relatively smaller impact in the evaluation of $h(x)$. Both errors in the evaluation of the probabilities have the same magnitude.

in 1995. A few years later, in 1998, Strong and colleagues developed the so-called direct method [91], which is relatively easy to implement and is thus widely used by experimentalists. For a more thorough review on correction methods, see [88]. It is important to note that all these correction methods are designed for the mutual information. Nonetheless, because transfer entropy is also the subtraction of two entropies, its estimation could suffer from similar biases too.

This Introduction focuses on the direct method, as it is the method used in this thesis for the evaluation of the mutual information. It was also used in this thesis to adapt a correction method for the transfer entropy (see Chapter 2 for more details).

The direct method

The direct method applies two distinct corrections for the evaluation of both entropies entering the calculation of the mutual information. The first correction aims to correct the bias due to a limited dataset. The idea is to extrapolate the entropy (total entropy or conditional (noise) entropy) to an infinite dataset by plotting the entropy for fractions of the dataset. Figure 1.9 shows an example of how this correction is applied in the original work from Strong and colleagues, for words of length 10 bins. The total entropy H_{tot} is computed for the whole dataset, as well as fractions of the dataset ranging from $\frac{1}{2}$ to $\frac{1}{5}$. The points are then fitted with a function

of the form $H_{tot} = H_0 + \frac{H_1}{size} + \frac{H_2}{size^2}$, where $size$ denotes the inverse data fraction. The intercept at the fraction $\frac{1}{0}$, H_0, gives the extrapolated value of the entropy for an infinite dataset.

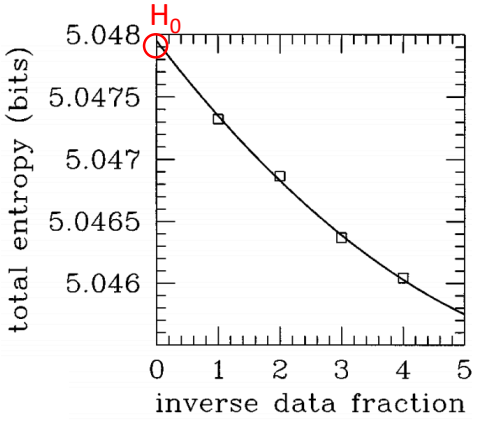

Figure 1.9: First correction of the direct method of Strong. Words used in this example have a length of 10 bins. Total entropy is evaluated over different fractions of the whole dataset. To extrapolate the value of the entropy for an infinite dataset, the points obtained is fitted with a function of the form $H_{tot} = H_0 + \frac{H_1}{size} + \frac{H_2}{size^2}$. The intercept H_0 is thus the value for an infinite dataset. Adapted from reference [91].

In addition to correcting the bias occurring for a small dataset, the direct method also offers a way to correct the bias due to finite word lengths. Ideally, the entropies should be estimated with an infinite word length to accurately take into account possible long-range correlations or stereotypical temporal patterns. Correlations and patterns reduce the information content of a spike train, but their time scales are usually not known. An infinite word length would thus ensure that they are all taken into account. For a small dataset, long words will particularly suffer form the bias described above, creating what is referred to as a sampling disaster by Strong and colleagues [91]. A sampling disaster occurs for relatively long words, when the first correction is not sufficient to correct the bias, leading to a strong underestimation of the entropy (see Figure 1.11C below). Interestingly, when the correlations have a limited range, their contribution to the entropy is usually a constant and emerges before the sampling disaster [91]. This means that entropies calculated for word lengths before the sampling disaster occurs (i.e. for relatively short words) can be used to extrapolate the estimates of the entropy and conditional entropy for words of infinite length (again, see Figure 1.11C).

A good example of the application of the direct method (and specifically the second correction) on experimental measurements can be observed in the work of Reinagel and Reid [92]. In this work, the authors used the direct method to calculate the stimulus-dependent information conveyed by single neurons in the cat lateral geniculate nucleus in response to randomly modulated visual stimuli (Figure 1.10 shows an example of how this kind of experiment is conducted). Their data were discretized accordingly.

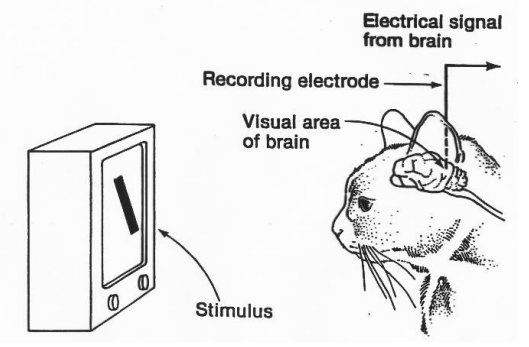

Figure 1.10: Example of a typical experiment on the cat visual pathway. A randomly modulated visual stimuli is shown to a cat and electrodes are used to record the responses of the target neurons. Adapted from reference [93].

Figure 1.11 shows how they evaluated mutual information for their measurements. As defined in the direct method, they evaluated the probability distributions for several word lengths. The entropy is evaluated from the responses to unique stimuli (see Figure 1.11A) and the conditional entropy is evaluated from the responses to a repeating stimuli (see Figure 1.11B). The mutual information is then calculated as the entropy minus the conditional entropy (see Equation 1.6).

One of the very powerful ideas introduced by the direct method (apart from the corrections) is how the conditional (or noise) entropy itself is evaluated (before applying the corrections). Instead of evaluating the joint and conditional probabilities for states between the input and output (like Equation 1.3 suggests), the conditional entropy is evaluated by comparing several outputs (called repetitions) generated by the same repeated input. The repetitions are aligned and a value of conditional entropy is calculated for each particular word position in the outputs across the repetitions (as can be seen in Figure 1.11B). The estimate of the conditional entropy is then evaluated as the average over all particular word positions. This allows the estimation of the conditional entropy to be decorrelated from the transmission processes between the input and the output. Those can indeed be a source of issues in estimating the information flow, or simply

be experimentally inaccessible. It is nonetheless important to note that this does not remove the conditionality on the input, as every repetition of the output is aligned so that each bin in every repetition corresponds to the same bin in the input (as shown in Figure 1.11B).

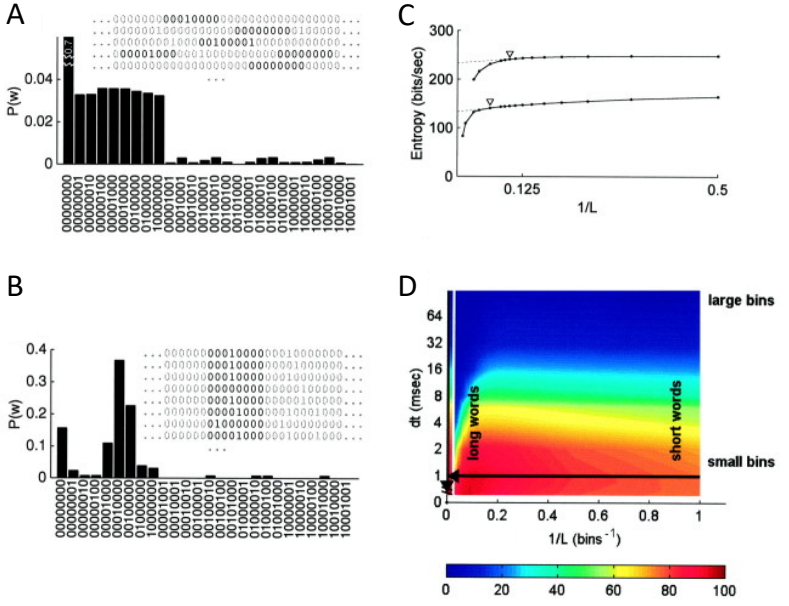

Figure 1.11: Example of the application of the direct method to estimate the mutual information. A: Probability distribution of words of length $L = 8$ for time bins of $1ms$ used for the calculation of the entropy. The inset is a sample of responses to a non repeated stimuli, with some words of length 8 highlighted. B: Probability distribution of words of length $L = 8$ for time bins of $1ms$ used for the calculation of the conditional entropy at one particular word position. The inset is a sample of responses to a repeated stimuli, with words with a fixed timing highlighted. C: Estimated entropy rate (top) and conditional entropy rate (bottom) as a function of the inverse word length $1/L$. The dashed line is the extrapolation from the linear part of the curves to words of infinite length. D: Space of parameters of the calculation of the mutual information $I(L, dt)$, indicated by color. The arrow indicates the state for the time bin $dt = 1ms$. Adapted from reference [92].

Both corrections of the direct method were applied by Reinagel and Reid [92]. Figure 1.11C shows the application of the second correction. After application of the first correction (not shown), the entropy and the conditional entropy are plotted as a function of the inverse word length. The sampling disaster can also be observed. In this case, it appears for words of length above $L \cong 10$ for the entropy and $L \cong 12$ for the conditional entropy. The correction for this bias implies to extrapolate the linear part of the

curve to words of infinite length, as it can also be seen in Figure 1.11C. The bias also appears for the evaluation of the mutual information, as shown in Figure 1.11D. The effect of the size of the discretized bin is also shown in Figure 1.11D.

The direct method is, as shown above, relatively easy to implement and efficient to evaluate the mutual information in experimental contexts. This makes it a powerful tool to study what processes can affect the information transmission. Section 1.2.3 focuses on one particular process that can modify information transmission: neuromodulation by serotonin.

1.2.3 Serotonin modulation of synaptic depression

Neurotransmitters

There are several processes used by neurons to communicate (electrical synaptic transmission, ephaptic interactions or hormonal signalling to name a few), but the major mode of neuronal communication is through chemical messengers also called neurotransmitters [94].

The general mechanisms involved in chemical synaptic transmission are summarized in Figure 1.12. Neurotransmitters are synthesized in neurons. In order to protect them from the enzymes responsible for their metabolic inactivation, most of them are stored in vesicles in neurons, before they are released by those same cells. Vesicular release is triggered by physiological stimuli and involves the fusion of the vesicle with the cellular membrane. Once released, neurotransmitters interact with receptors or other proteins on the membrane of other cells, producing a functional change in the properties of that cell. They can also interact with autoreceptors present on the emitting cell as a feedback signal for the transmitter release or synthesis. To avoid overstimulation, it is important that the effect of the neurotransmitters can be stopped after an appropriate period of time. For this purpose, several termination mechanisms exist. The first one is the reuptake by neurons or glial cells with the help of transporter proteins. There also are enzymes in neurons responsible for degrading neurotransmitters. Finally, some neurotransmitters can diffuse away from the synaptic region.

The most understood and studied category of neurotransmitters so far are the so-called classical neurotransmitters, which regroups acetylcholine, biogenic amines (dopamine, norepinephrine, epinephrine and serotonin) and amino acid transmitters (γ-aminobutyric acid (GABA); the major inhibitory neurotransmitter, glutamate; the major excitatory neurotransmitter, and aspartate). Most of them were already recognized as neurotransmitters in the late 1950s. Classical neurotransmitters are synthesized in the nerve terminal, stored in small vesicles ($\sim$ 50 nm in diameter) and their reuptake is an energy-dependent process that allows the transmitters to be reused [94].

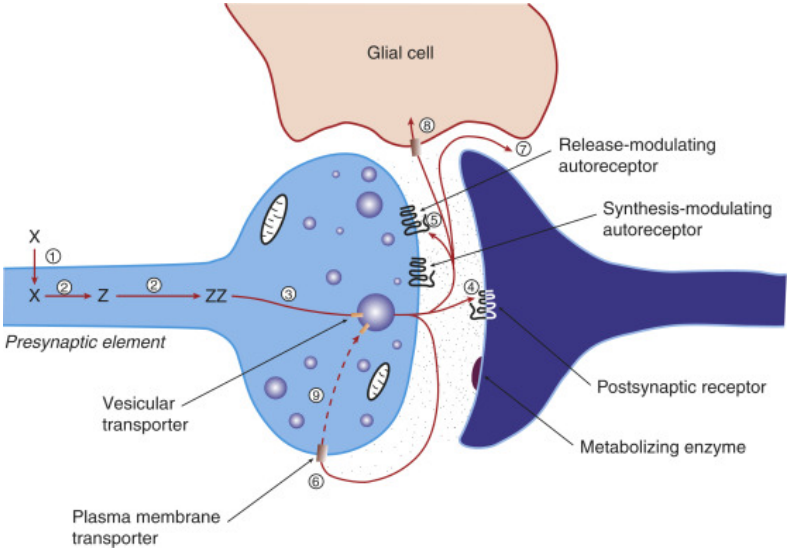

Figure 1.12: Schematic representation of the life-cycle of a neurotransmitter. The neurotransmitter (Z) is metabolized from the accumulation of a precursor amino acid (X) in the neuron (1 and 2). The vesicular transporter then accumulates the transmitter into vesicles (3). After the release, the transmitter can either interact with postsynaptic receptors (4) or autoreceptors (5). High-affinity membrane transporters (6) are responsible for terminating the transmitter action. The action of the neurotransmitter can also be terminated either by diffusion from the active sites (7) or with the help of a membrane transporter that accumulates the transmitter into glia (8). In the case where the transmitter re enters the neuron it originated from, it will undergo metabolic inactivation (9). Adapted from reference [94].

Neuropeptides are considered to be non-classical transmitters. They are synthesized in the cell-body of neurons (as opposed to the axon, where classical transmitters are synthesized). Neuropeptides are stored in bigger vesicles ($\sim$ 100 nm in diameter) that take them to the axon, where they are released. Peptides are usually released when neurons discharge at high frequencies or when they exhibit burst-firing patterns. So far, no transporters have been identified for being responsible for the reuptake of peptides, which means that they are inactivated by enzymes and by diffusion [94].

Classical transmitters were the first to be discovered because they are present in the brain in relatively high concentrations. Peptides transmitters are present in lower concentrations than classical transmitters, which makes them more difficult to study. More recently, some unconventional transmitters have also been discovered, like endocannabinoids, which are not stored into vesicles but synthesized and released from lipid precursors located in the cell membrane [95], or gases like nitric oxyde, carbon monoxyde and hydrogen sulfide [96]. Gases are not stored, their molecules diffuse freely across cellular

and vesicular membranes, which makes them messengers independent of the neuronal activity, unless they are modified or released by neuronal metabolism. Other unconventional transmitters are growth factors and neuroactive steroids [94].

There are dozens of different neurotransmitters that we currently know of and one could ask why so many transmitters are needed by the brain. There are several reasons for that. Firstly, a neuron can receive many inputs simultaneously. If those inputs are encoded by different neurotransmitters, they are easier to differentiate for the receiving neuron. Moreover, different transmitters can signal different functional states to the target cell. This is the reason why most neurons release more than one transmitter, and why different transmitters are released by different processes. Different neurotransmitters also have different temporal responses, which increases the temporal repertoire of responses that can be elicited. Finally, neurotransmitters are also used by brain cells for non-synaptic communication, for example communication between neurons and glial cells [94].

Serotonin

5-hydroxytryptamine (5-HT), also called serotonin, is a biogenic amine neurotransmitter and is depicted in Figure 1.13. The serotonin present in the brain represents only about 1% of the total body stores of serotonin [94], as it is used for several other function in the body, like gastrointestinal motility, cardiovascular function, peripheral vascular tone, cerebral vascular tone and platelet aggregation [97, 98]. Serotonin was actually first characterized by Rapport and Page in 1948, as they were studying platelets, which are the main storage of serotonin in the body [99]. It was then a few years later that two different studies showed the presence of serotonin receptors in the vertebrate brain [100, 101], which led in 1957 to the first paper proposing serotonin to be a neurotransmitter [102].

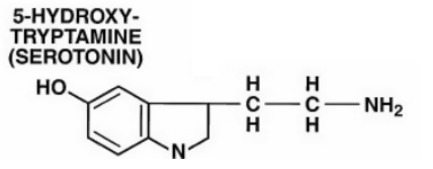

Figure 1.13: Skeletal formula of 5-hydroxytryptamine (5-HT), also called serotonin. Adapted from reference [97].

Serotonin is considered to be a classical neurotransmitter and its life cycle thus follows what is depicted in Figure 1.12. In the CNS, serotonin is synthesized and stored in axons and released into the synaptic cleft when

neuronal depolarization occurs. The main inactivation process for serotonin is reuptake into the neurons through the highly selective transporter SERT [97].

Serotonin is responsible for the modulation of many behavioural and neuropsychological processes, such as mood, perception, nociception, reward, anger, aggression, appetite, memory, sexuality and attention [98]. There exists at least 15 receptors for serotonin, grouped into 7 families [103], and all brain regions express multiple serotonin receptors [97]. Moreover, all the behaviour cited above are regulated by multiple serotonin receptors.

Serotonin modulation in the lateral geniculate nucleus

Visual information is transmitted from the retina to the cortex by thalamic relay neurons in the LGN, as can be seen in Figure 1.14 (see also Figure 1.6A for the circuitry of thalamic relay neurons in the LGN). As was discussed in Section 1.2.1, in the absence of neuromodulation, thalamic relay cells seem to operate at an energetically efficient level with respect to information transmission [68]. This means that those neurons maximize the ratio between the information they convey and the energy they use doing so. Nonetheless, neuromodulation also plays a big part in the functioning of those neurons.

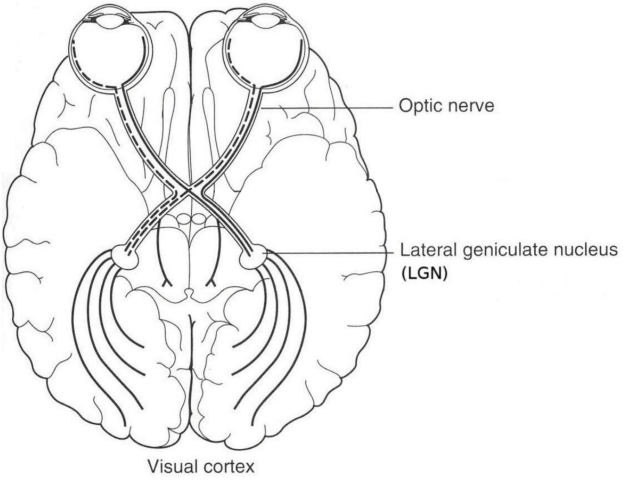

Figure 1.14: Schematic representation of the visual pathway. The optic nerve transmit visual information from retinal neurons to thalamic relay neurons in the LGN. Relay neurons in the LGN then send their signals to neurons in the visual cortex. Adapted from reference [104].

Postsynaptically, this visual information transiting at retinogeniculate synapses (the synapses between retinal neurons ans LGN neurons) can be modified by neurotransmitters released by intrinsic neurons or brainstem

projections [105, 106] and affecting the membrane conductance of relay neurons. This modification can either be a hyperpolarization of the cell and a decrease of its activity, for example by activation of $GABA_B$ receptors on the cell [107], or it can be a depolarization and an increase of activity, for example from cholinergic, noradrenergic or serotonergic projections form the brainstem [106, 108]. Presynaptic modulation can also occur at retinal axon terminals, altering the strength and short-term plasticity of retonigeniculate synapses [109, 110].

Projections from the dorsal raphe nucleus in the brainstem release serotonin in the LGN [111–113]. Serotonin is implied in the modulation of presynaptic and postsynaptic aspects of the retinogeniculate synapse [110, 114]. Postsynaptically, it results in depolarization of neurons and cancellation of spontaneous firing. It is thus thought to contribute to the activation of thamalocortical neurons when transitioning from slow-wave sleep to awake states [115–117]. Presynaptically, it inhibits action potential-evoked calcium influx into retinal axon terminals, which reduces neurotransmitters release and short-term synaptic depression [110]. This inhibition has a greater effect for low frequency stimulation, which narrows the frequency range of information transmitted [114].

When synapses are stimulated repeatedly in a short period of time, their strength is reduced, *i.e.* the amplitude of each consecutive postsynaptic current decreases [118]. This phenomenon is called synaptic depression. Synaptic depression can have short-term and long-term effects [119], allowing relative strengths of excitatory and inhibitory synapses to vary dynamically as a function of the frequency and duration of presynaptic activity [120].

Figure 1.15 demonstrates how presynaptic application of serotonin on LGN neurons affects synaptic depression. 5-carboxytryptamine (5-CT) is an agonist for the serotonin receptor 5-HT_1, which is the primary presynaptic receptor for serotonergic modulation of the retinogeniculate synapses [114]. A bath of 5-CT thus acts as application of serotonin to that synapse. Figure 1.15A shows the comparison of synaptic depression on consecutive excitatory postsynaptic currents (EPSCs) between control conditions and the case with 5-CT (50 nM in the bath). The traces shown are superimposed pairs of consecutive EPSCs separated by interspike intervals (ISIs) varying between 10 and 200 ms. As it can be seen, 5-CT reduces the peak amplitude of EPSCs. It also reduces synaptic depression, especially for short ISIs, because amplitudes of consecutive EPSCs are more similar to each other with 5-CT than in control condition. This can also be seen when looking at the paired-pulse ratio. The paired-pulse ratio (ppr) is defined as the ratio between the amplitude of the second EPSC (A2) and the amplitude of the first one (A1), $ppr = 100 \cdot \frac{A2}{A1}$, and is shown in Figure 1.15B as a function of the ISI. For example, for an ISI of 20 ms, the paired-pulse ratio for the control case is around 20%, while it stands at around 83% for the 5-CT

bath, showing again than amplitudes are more similar in the case with 5-CT and thus that synaptic depression is lower. Moreover, it can also be seen that 5-CT significantly changes the synaptic depression for pairs of stimuli separated by up to 1 second, which means that serotonin can alter the synaptic response over a wide range of presynaptic firing frequencies.

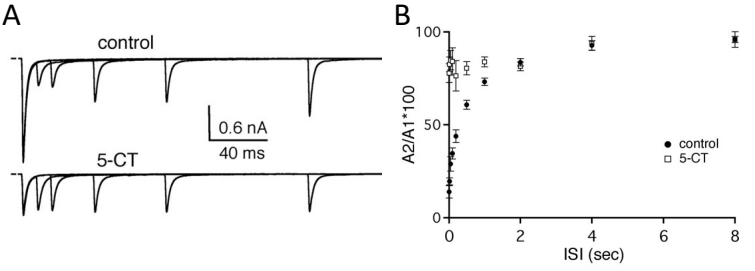

Figure 1.15: Effect of the serotonin agonist 5-CT on the amplitude of consecutive excitatory postsynaptic currents (EPSCs). A: Superimposed pairs of EPSCs separated by varying inter spike intervals (ISIs) for control conditions (top) and in the presence of 5-CT (bottom). B: Percentage paired-pulse ratio between the amplitude of the second (A2) and the amplitude of the first (A1) EPSC as a function of the ISI for control conditions (filled circles) and in the presence of 5-CT (open squares). Both panels show that the serotonin agonist 5-CT reduces synaptic depression on consecutive EPSCs. Adapted from reference [114].

When applied presynaptically, serotonin reduces the synaptic charge and the synaptic depression. EPSCs elicited are thus smaller and more similar in amplitude, which means that it will be more difficult to elicit an action potential. In other words, serotonin can potentially change the energy needed to convey information and the information conveyed itself. But what about the energetic efficiency of that synapse? In order to study this effect, it is necessary to be able to include synaptic depression in computational models.

Model for synaptic depression

A mathematical model for synaptic depression was developed by Tsodyks and Markram in 1997 [121], based on measurements in rat layer 5 pyramidal neurons. Their model assumes three possible states for synaptic resources characterizing the synaptic connection (*e.g.* synaptic vesicles or receptors): recovered, effective and inactive. Resources in the recovered state are available to be activated by presynaptic action potentials. Once activated by an action potential, they enter the effective state and then inactivate with a short time constant ($\tau_{inac} \sim$ few milliseconds). The inactivated resources then recover their original state with a longer time constant ($\tau_{rec} \sim$ seconds). This cycle is depicted in Figure 1.16.

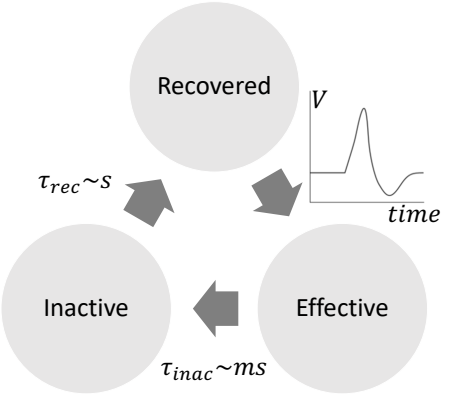

Figure 1.16: Schematic representation of the states of synaptic resources in the Tsodyks-Markram model. The resources cycle between three states: effective, inactive and recovered. Resources in the recovered state can be activated by an action potential, thus entering the effective state. Once in the effective state, they inactivate with a short time constant ($\tau_{inac} \sim ms$). Inactivated resources recover with a longer time constant ($\tau_{rec} \sim s$), going back to the recovered state.

From there, Tsodyks and Markram derived an iterative expression for successive EPSCs produced by a train of presynaptic action potentials. The current of the $(n+1)$th EPSC depends on the current of the nth EPSC as follows [121]:

$$EPSC_{n+1} = EPSC_n(1 - U_{SE})e^{-\Delta t/\tau_{rec}} + A_{se}U_{SE}(1 - e^{-\Delta t/\tau_{rec}}), \quad (1.11)$$

where Δt is the time interval between the nth and the $(n+1)$th presynaptic action potential and is assumed to be much larger than τ_{inac}, U_{SE} is the fraction of synaptic resources in the recovered state that can be activated when an action potential reaches the synapse (U_{SE} is determined by the probability that an action potential evokes vesicular release; the larger U_{SE} is, the faster synaptic resources are used, leading to more rapid depression), and A_{SE} is the maximum EPSC that can be evoked when all the resources are used at once in the effective state. U_{SE}, τ_{rec} and A_{SE} are kinetic parameters, they determine the dynamic behaviour of the synaptic transmission and, in particular, the rate of synaptic depression.

Figure 1.17 shows the results of the Tsodyks-Markram model on two different set of EPSCs: one set generated by regular spike trains (Figure 1.17A) and the other set generated by irregular spike trains (Figure 1.17B). On both panels, the bottom traces are the spike trains used as input (either injected into the cell or injected in the model), the upper traces are the experimentally measured EPSCs (averaged on 50 measures) and the middle

traces are the EPSCs generated by the model. First of all, Figure 1.17 shows that the model is efficient at reproducing the experimental traces. Secondly, the EPSCs traces reproduced in this Figure exhibit a similar behaviour to the EPSCs showed in Figure 1.15A, suggesting that this model is a good choice for simulating neuromodulation by serotonin.

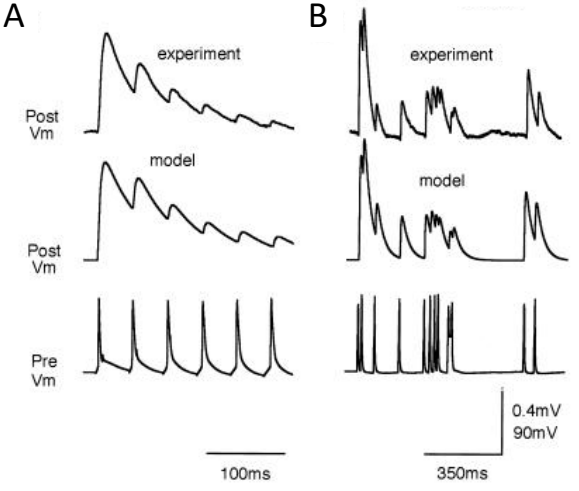

Figure 1.17: Accuracy of the Tsodyks-Markram model. The bottom traces are the spike trains used as inputs (regular spike train in A and irregular spike train in B), the upper traces are the EPSCs measured experimentally (averaged on 50 measurements) and the middle traces are the EPSCs generate with the model. There is only few differences between the measured and modelled EPSCs (upper and middle traces), showing the accuracy of the Tsodyks-Markram model. Adapted from reference [121].

1.3 Main research questions and hypotheses

As we have seen throughout this introduction, energetic efficiency of information transfer is a key feature of, at least, some synapses in the brain. In the LGN particularly, synapses do not maximize information transmission but the ratio between information transmitted and concomitant energy consumption.

The brain consumes a big fraction of the total amount of energy available in the body and this has shaped the way the brain operates. Energetic budgets of the brain have helped better understand how and by which processes energy is used. In particular, signalling processes have been shown to be the major consumers of energy in the brain. Transmitting all the information received would cost too much energy for the neurons and this

is why the vesicle release probability is rarely equal to 1. In particular, for neurons in the LGN, with large synapses releasing many vesicles, release probability is probably set, at least partially, by this energetic efficiency, or economic, principle.

Of course, to better understand energetic efficiency, it is important to be able to characterize the information that flows through synapses. Information theory, even though it was developed for communication sciences, offers a wide sets of mathematical approaches applicable in Neuroscience to evaluate information-related quantities. Two example of such metrics are mutual information and transfer entropy. Application of those metrics on experimental data can be tricky, as they rely on a number of assumptions and estimations.

Finally, energetic efficiency is not necessarily a permanent fixture of synapses. For example, neuromodulators can change synaptic strength, thus possibly affecting energy consumption and information transmission, which could also change energetic efficiency. In particular, serotonin changes the level of synaptic depression occurring in the LGN, which will have a direct effect on the generation of action potentials, and thus on the transmission of information and the energetic cost of this transmission.

This thesis aims to better understand and characterize energetic efficiency of information transfer in the visual pathway. We created models based on experimental data simulating neuronal inputs and outputs with the goal to evaluate information transfer and energy usage, as well as to simulate neuromodulation. Specifically, the papers presented in Chapter 2 address the following points:

1) Information measurements:
The evaluation of energetic efficiency necessitates the correct evaluation of information transfer. From a more general point of view, accurately measuring information-related quantities is an essential task in Neuroscience. In this thesis, we compared the relative performances of two information theoretic metrics: the mutual information and the transfer entropy. This comparison was done according to the size of the dataset (a critical point when designing experiments) and the complexity of the data. Models with different levels of biological likelihood were designed, mimicking thalamic relay cells and layer 4 spiny stellate cells (Paper 2).

2) Energetic efficiency at LGN synapses
We aimed to reproduce experimental findings about energetic efficiency in the visual pathway (Paper 1). We used the model as well as the metrics studied in 1) with the goal to reproduce the finding that thalamic relay cells maximize the ratio between information conveyed and energy used for experimentally-observed physiological conductances (Paper 2).

3) Neuromodulation by serotonin:
As neuromodulation changes the strength of synapses, it could as well alter its energetic efficiency. In this work, we concentrated on the effect of serotonin at LGN synapses. More concretely, we developed a model of paired-pulse depression and neuromodulation of serotonin based on the formalism introduced by Tsodyks and Markram and detailed in this introduction, with the goal to asses if this type of neuromodulation affects information processing and energetic efficiency, again using the information theory metrics highlighted in 1) (Paper 3).

CHAPTER 2

Publications

This chapter presents the three publications that resulted from the research performed for this thesis. Each paper is preceded by a short introductory section summarizing the methods and results and highlighting my contributions. The papers are listed in the order the studies were performed.

2.1 Paper 1

M Conrad, E. Engl, and R. B. Jolivet, "Energy use constrains brain information processing", 2017 IEEE International Electron Devices Meeting (IEDM) (Jan. 2018), pp.11.3.1-11.3.3. (7 citations on Google Scholar)

Summary

This paper is a review discussing evidence that energetic constraints influence neural cells and network properties, as well as the information processing capacity of neurons. The paper first focuses on energy expenditure by brains and neural cells. The paper also explains what processes are responsible for information transmission among neural networks, and how this information can be evaluated. The role of glial cells in information processing and energy consumption is also discussed.

The main point discussed in this paper is the fact that the majority of the brain's energy is used at synapses, even though synaptic information transmission is generally unreliable (as discussed in Section 1.2.1 of Chap-

ter 1). We show previous experimental and computational findings, and also original results, suggesting that neurons maximize the ratio between information transmission and concomitant energy consumption.

Contribution

I contributed to the writing of this manuscript.

Energy use constrains brain information processing

Mireille Conrad[1,2], Elisabeth Engl[3], and Renaud B. Jolivet[1,4]
[1]DPNC, University of Geneva, Geneva, Switzerland
[2]Lemanic Neuroscience Doctoral School, Geneva, Switzerland
[3]Lodestar Insights, London, United Kingdom
[4]CERN, Geneva, Switzerland, email: renaud.jolivet@unige.ch

***Abstract*—**The brain is an energetically expensive organ to build and operate, and a large body of literature links the evolutionary development of many of the human brain's components to the need to save energy. We, and others, have shown experimentally and through computational modelling that synapses in the brain do not maximise information transfer, but instead transfer information in an energetically efficient manner. Strikingly, this optimum implies a high failure rate in the transmission of individual information-carrying signals (action potentials or spikes). This design principle may be important when considering trade-offs between energy use and information transfer in man-made devices.

I. Introduction

The human brain is likely the most advanced biological computational device that evolved on Earth. While it lacks in raw power for arithmetic operations, it excels at processing analogue information from the natural world (sounds, images, chemical senses, pressure, temperature, etc.), seamlessly blending these stimuli into cohesive impressions, learning from them - sometimes with a single sample presentation -, storing those memories for decades, and inferring new knowledge from those memories.

Amazingly, the human brain performs these tasks with a power consumption that rarely deviates from approximately 20W in adults [1], while a desktop computer may use ~200W or more. While this may seem extraordinarily low, the brain nevertheless accounts for a disproportionate 20% of the overall body's energy budget. Therefore, while those 20W seem small for the performance achieved, they represent a significant burden on the body. Evidence from anthropology and neuroscience suggests that energy constraints might have played an important role in how we, and our brains, have evolved [2], and energy constraints have shaped the evolution of brains across species [3].

Here, we introduce recent evidence that energetic constraints determine key properties of neural networks at the individual cell level and restrict individual neurons' capacity for information processing. We highlight the notion that brains operate a trade-off between their raw information processing capacity and the concomitant energy consumption at the level of individual processing units.

II. Basic Units of Brain Circuits

A. Energy consumption in brain circuits

Our brains consist of electrically active cells called neurons. To an extent, neurons can be approximated by a resistor-capacitor circuit with multiple variable resistors arranged in parallel. The currents that flow through those resistors (or ion channels) are carried out by ions and driven by electrochemical gradients that arise across the membrane of the cell. In order to maintain those gradients, or to restore them after ion movements, cells need to actively pump ions across their membranes. This is where a cell spends most its energy currency in the form of adenosine triphosphate, which powers the cycles of these pumps.

B. Computation and information in brain circuits

Currents flowing though the membrane of neurons can give rise to a sharp (~100mV over ~2ms) depolarisation of the membrane called action potential or spike. Spikes quickly propagate along neuronal structures and form the basis of the binary-like long-range communication system among neural networks. Upon reaching the terminal structure of a neuron, spikes trigger the release of chemicals (neurotransmitters) at dedicated structures called synapses that physically connect neurons. Upon binding to the appropriate receptors, neurotransmitters give rise to trans-membrane currents in downstream neurons, which can in turn lead to the generation of spikes in those neurons.

Integration of inputs from numerous upstream neurons and generation of output spikes form the basis of neuronal cellular computation and information processing. The amplitude of total currents generated at a synapse, the synaptic strength, can be altered, modulating the relative importance of that synapse's contribution to the output. This synaptic plasticity is commonly accepted to be the mechanism through which memories are encoded.

Because the shape of spikes is stereotypical, only the timing of action potentials within sequences carries information. Thus, sequences of input and output spikes can easily be binarized and lend themselves to analysis using Shannon's information theory.

C. Energy as a gateway to a holistic understanding of brain function

As indicated above, the brain contains neurons, but these represent only ~20% of all cells in the human cortex. The other cells, collectively called glial cells (excluding the

vasculature and a few other cell types), are not electrically active in the same way neurons are, and for historical reasons, have been studied comparatively less. However, this is not to say that they are not active in other ways. Quite the contrary: recent evidence suggests that some of these cells play a role in information processing by intervening in the formation of synapses and in the modulation of synaptic strength. Additionally, these cells play a key role in linking energy consumption by neurons to energy supply from the blood flow. Thus, understanding energy constraints in brain circuits is a natural gate to a holistic understanding of brain function beyond mere neural networks [4].

III. Energetic Optimality in Brain Circuits

Applying detailed accounting to various energy-demanding biochemical processes, it is possible to calculate an energy budget for the brain. Doing so reveals that most energy in the brain is spent at synapses [1, 5]. This however poses a conundrum, as experimental evidence suggests that synaptic transmission is typically rather unreliable from an information transmission perspective. Why would such a costly feature be so unreliable?

For instance, chemical neurotransmission at synapses is a stochastic process: a spike generated and propagated in an upstream neuron gets transmitted to downstream neurons with a surprisingly low probability of as low as ~0.2 [6, 7].

Similarly, in the visual pathway, visual information is relayed to the cortex by a (essentially) one-to-one relay station connection between the retina and the visual cortex. While complete failures of signal transmission are rare in that system, the amplitude of the signal generated at that synapse is just not quite sufficient to trigger an action potential in the neuron relaying the signal to the visual cortex, rendering the transmission stochastic.

In both cases, some of us with others [8, 9] have investigated experimentally and in computational models how information transmission and concomitant energy consumption relate to each other when manipulating the overall reliability of information transfer.

In both cases, energy consumption can be modelled or experimentally measured to depend roughly linearly on the reliability of information transmission [8, 9]. Information transmission is monotonically dependent on the transmission reliability of the system and rests between 0% for a trivial system that never transmits and 100% of the input information for a perfect transmitter. Because of the presence of noise and/or of activation thresholds in neurons, the relation between information transmission and reliability is in fact best represented by a sigmoid. The consequence of these observations is that the relation between information transmission per concomitant energy consumption and transmission probability is a monophasic curve with a single optimum. Therefore there exists a single energetic optimum for information transmission in neural systems.

For synapses with a realistic structure, Harris *et al.* have shown in computational models that this optimum sits at low probabilities for information transmission, consistent with experimental observations [6, 8] (Fig. 1). A simple model of energetic optimality of information transmission additionally explains the experimentally observed relation between the reliability of synapses and their structure [7] (Fig. 2).

In the visual pathway, Harris *et al.* have shown that for almost all connections they recorded from, synapses do not maximise information transfer [9]. Indeed, it is possible to increase or decrease information transfer by manipulating those synapses experimentally (Fig. 3). However, when taking the concomitant energy consumption into account, Harris *et al.* have shown that almost every synapse in the pathway studied sits at the optimum for energetic optimality of information transmission (Fig. 4).

Therefore, in both cases, one can conclude that neural networks are not designed for optimal performance in terms of pure information transmission, but instead are forced into a trade-off between information processing power and energy consumption. We now have preliminary evidence that this trade-off might be widespread in the cortex, and are investigating how it emerges in brain circuits. We would like to argue that understanding this trade-off would be an important source of inspiration for the design of algorithmic- and power-efficient neuromorphic devices.

Acknowledgment

This work is supported by a grant from the Swiss National Science Foundation to RBJ (ID# 31003A_170079).

References

[1] D. Attwell, and S. B. Laughlin, "An Energy Budget for Signaling in the Grey Matter of the Brain," *J. Cereb. Blood Flow Met.*, vol. 21, pp. 1133-1145, Oct. 2001.

[2] L. C. Aiello, and P. Wheeler, "The Expensive-Tissue Hypothesis," *Curr. Anthropol.*, vol. 36, pp. 199-221, Apr. 1995.

[3] J. E. Niven, and S. B. Laughlin, "Energy limitation as a selective pressure on the evolution of sensory systems," *J. Exp. Biol.*, vol. 211.11, pp. 1792-1804, Jun. 2008.

[4] R. B. Jolivet, J. S. Coggan, I. Allaman, and P. J. Magistretti, "Multi-timescale Modeling of Activity-Dependent Metabolic Coupling in the Neuron-Glia-Vasculature Ensemble," *PLOS Comput. Biol.*, vol. 11, doi:10.1371/journal.pcbi.1004036, Feb. 2015.

[5] R. B. Jolivet, P. J. Magistretti, and B. Weber, "Deciphering neuron-glia compartmentalization in cortical energy metabolism," *Front. Neuroenerg.*, vol. 1, doi:10.3389/neuro.14.004.2009, Jul. 2009.

[6] T. Branco, K. Staras, K. J. Darcy, and Y. Goda, "Local Dendritic Activity Sets Release Probability at Hippocampal Synapses," *Neuron*, vol. 59, pp. 475-485, Aug. 2008.

[7] N. R. Hardingham, J. C. A. Read, A. J. Trevelyan, J. C. Nelson, J. J. B. Jack, and N. J. Bannister, "Quantal Analysis Reveals a Functional Correlation between Presynaptic and Postsynaptic Efficacy in Excitatory Connections from Rat Neocortex," *J. Neurosci.*, vol. 30, 1441-1451, Jan. 2010.

[8] J. J. Harris, R. B. Jolivet, and D. Attwell, "Synaptic energy use and supply," *Neuron*, vol. 75, pp. 762-777, Sep. 2012.

[9] J. J. Harris, R. B. Jolivet, E. Engl, and D. Attwell, "Energy efficient information transfer by visual pathway synapses," *Curr. Biol.*, vol. 25, pp. 3151-3160, Dec. 2015.

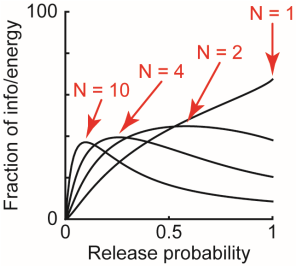

Fig. 1. Theoretical calculations show that a low presynaptic release probability maximizes the energy efficiency of information transfer between neurons connected by N > 1 release sites (adapted from [8]).

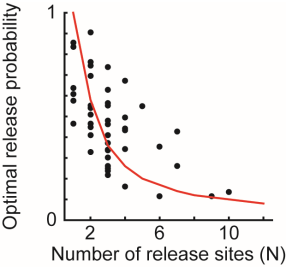

Fig. 2. Experimentally measured release probability is inversely related to the number of release sites (black symbols; adapted from [7]). Theoretical prediction of the energetically optimal release probability (red line; maxima from Fig. 1) provides an elegant explanation for this finding.

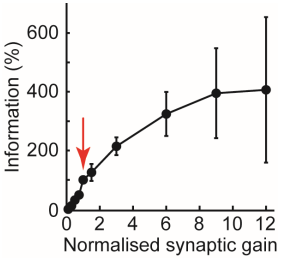

Fig. 3. Experimentally modulating the gain of relay synapses in the visual pathway in adult rat brain slices shows that those synapses are not tuned to maximize information transfer (black symbols; the arrow points to the original physiological gain of the synapse (normalized synaptic gain = 1)).

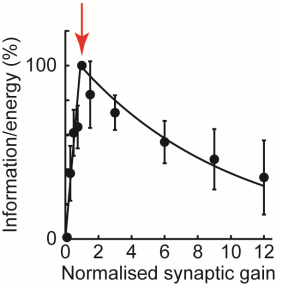

Fig. 4. Taking into account the energy consumption incurred by neurons in the visual pathway at different synaptic gains shows that those synapses maximize instead the energetic efficiency of information transfer (arrow as in Fig. 3; adapted from [9]).

2.2 Paper 2

M Conrad, and R. B. Jolivet, "Comparative performance of mutual information and transfer entropy for analysing the balance of information flow and energy consumption at synapses", in revision

Summary

In this paper, we compare the performance and trade-offs of mutual information and transfer entropy when used to measure the information transmitted between binarized input and output spike trains in the visual pathway. To do that, the two metrics are tested on spike trains generated to correspond to typical inputs and outputs of thalamic relay cells and layer 4 spiny stellate cells, with three different levels of complexity and biological likeness. The first level of complexity are inputs with a Poisson distribution and outputs generated as simple copies of the inputs, with some deletions and additions to model, respectively, failure of synaptic transmission and spontaneous firing. The next step is to simulate more realistic inputs, based on the temporal distribution of interspike intervals in *in vivo* recorded spike trains and with outputs generated as in the first step. The last complexity level are outputs simulated using an experimentally calibrated Hodgkin-Huxley-type model. The mutual information is evaluated using the direct method developed by Strong [91]. A method similar to the direct method is developed to evaluate transfer entropy and uses the package developed by Ito *et al.* [122].

The paper shows that in the two simpler cases, transfer entropy performs better at evaluating the information transmitted between the input and the output spike trains than mutual information, especially for small datasets. In the case with the Hodgkin-Huxley type model, even though the curves have similar behaviour as for the simpler cases, mutual information and transfer entropy do not converge to the same value. We show that this is the result of systematic transmission delays between inputs and outputs (time shifts), an issue to which mutual information is immune. Transfer entropy is thus less biased when used on small datasets, but it is more sensitive to temporal structures.

The paper also focuses on reproducing experimental findings of energetic efficiency of information transfer in thalamic relay cells. We compute the information (using mutual information and transfer entropy) as a function of the postsynaptic gain of the cell. We also evaluate the energy used to convey the information (also as a function of the gain of the cell) by calculating the ATP cost of synaptic transmission. This allows us to calculate the ratio between information transmitted form the input to the output spike trains and concomitant energy consumption, and to show that there is indeed a maximum in this function for the physiological gain of the cell, as showed

experimentally.

Contribution

I wrote most of the code necessary as well as runned all the simulations. I also analysed the data and was involved in writing the manuscript.

Comparative performance of mutual information and transfer entropy for analyzing the balance of information flow and energy consumption at synapses

Short title: A comparison of information theory measures for spike trains

Authors: Mireille Conrad & Renaud B Jolivet *

Affiliation: Département de Physique Nucléaire et Corpusculaire, University of Geneva, Geneva, Switzerland

Keywords: Information theory; Mutual information; Transfer entropy; Synapse; Energetic optimality;

* Correspondence to: renaud.jolivet@unige.ch

Abstract (187 words)

Information theory has become an essential tool of modern neuroscience. It can however be difficult to apply in experimental contexts when acquisition of very large datasets is prohibitive. Here, we compare the relative performance of two information theoretic measures, mutual information and transfer entropy, for the analysis of information flow and energetic consumption at synapses. We show that transfer entropy outperforms mutual information in terms of reliability of estimates for small datasets. However, we also show that a detailed understanding of the underlying neuronal biophysics is essential for properly interpreting the results obtained with transfer entropy. We conclude that when time and experimental conditions permit, mutual information might provide an easier to interpret alternative. Finally, we apply both measures to the study of energetic optimality of information flow at thalamic relay synapses in the visual pathway. We show that both measures recapitulate the experimental finding that these synapses are tuned to optimally balance information flowing through them with the energetic consumption associated with that synaptic and neuronal activity. Our results highlight the importance of conducting systematic computational studies prior to applying information theoretic tools to experimental data.

Author summary (265 words)

Information theory has become an essential tool of modern neuroscience. It is being routinely used to evaluate how much information flows from external stimuli to various brain regions or individual neurons. It is also used to evaluate how information flows between brain regions, between neurons, across synapses, or in neural networks. Information theory offers multiple measures to do that. Two of the most popular are mutual information and transfer entropy. While these measures are related to each other, they differ in one important aspect: transfer entropy reports a directional flow of information, as mutual information does not. Here, we proceed to a systematic evaluation of their respective performances and trade-offs from the perspective of an experimentalist looking to apply these measures to binarized spike trains. We show that transfer entropy might be a better choice than mutual information when time for experimental data collection is limited, as it appears less affected by systematic biases induced by a relative lack of data. Transmission delays and integration properties of the output neuron can however complicate this picture, and we provide an example of the effect this has on both measures. We conclude that when time and experimental conditions permit, mutual information – especially when estimated using a method referred to as the 'direct' method – might provide an easier to interpret alternative. Finally, we apply both measures in the biophysical context of evaluating the energetic optimality of information flow at thalamic relay synapses in the visual pathway. We show that both measures capture the original experimental finding that those synapses are tuned to optimally balance information flowing through them with the concomitant energetic consumption associated with that synaptic and neuronal activity.

Introduction (828 words)

The brain is commonly thought of as an information transmission, processing and storage biological machine, calling for comparisons to man-made devices with similar functions such as computers (see [1] for a discussion of these issues). As a consequence, tools and methods pertaining to such devices have been ported to the neurosciences for the study of neural networks. One such tool is information theory, designed in the late 1940s by Claude Shannon to formalize and find limits on signal processing and communication operations in machines [2].

Information theory has a long and successful history in neuroscience, where it has been applied to a variety of experimental data and theoretical contexts, and to address a variety of questions. One area of particular interest is the application of information theoretic concepts to spike trains, as they easily lend themselves to a reduction to binary sequences, whereupon action potentials are converted to 1s and the rest of electrophysiological traces to 0s. This reduction of spike trains to binary sequences has been used to measure information flow through synapses and neural networks, and propagation of information from the environment to the cortex through sensory pathways for instance [3-10]. The information theoretic quantity most often encountered in such contexts is the mutual information (I) [4]. The mutual information of two random variables – for instance two spike trains, or a sensory signal and a response spike train – is a measure of the mutual dependence between those two variables. By construction, I is symmetrical and quantifies how much information can be obtained about one random variable by observing the other one.

In 1998, Strong and colleagues published a procedure to measure I between a sensory input signal and the corresponding response of a neuron [11]. Through clever design, their so-called 'direct method', allows calculating I by measuring only the activity of the neuron of interest in

response to a specifically designed input sequence containing repeating portions. This method has been successfully applied in a number of contexts, for instance to quantify information flow from the retina to the primary visual cortex in primates [3], or to quantify the relation between information flow and energy consumption at synapses in the visual pathway [8]. One limitation of using *I* is the need to collect relatively large amounts of data to avoid systematic evaluation biases, and a number of methods have been devised to compensate those biases when data is limited or difficult to acquire [6]. This, however, can be tricky in an experimental context, as the time available for collecting data can be limited for a large number of various reasons.

Mutual information does not however strictly quantify directional information flow as it is symmetrical by design. To address this limitation, Schreiber has proposed a modified version of mutual information called transfer entropy (*TE*), which is explicitly built to measure how much information flows from one random variable to another one, and which is thus not symmetrical [12]. While *TE* is used widely outside of neuroscience and in systems neuroscience, it is not used very often as a replacement of *I* for analyzing spike trains specifically.

We have recently published a series of experimental and computational works on the trade-off between information flow and concomitant energy consumption in neurons and neural networks, in which we used either *I* [7-9] or *TE* [10]. This work is part of an emergent interest for energetic questions in neural information processing [13-16]. Here, we proceed to a systematic comparison of both those quantities in different biologically-relevant scenarios when comparing inputs to a synapse and the output spike train generated by the postsynaptic neuron in response, using spike train data generated from Poisson processes, or an experimentally-calibrated biophysical model of thalamic relay cells of the Hodgkin-Huxley-type [8]. We decided to focus on those two measures, and specifically on the so-called direct method to compute *I*, because of

their relative user-friendliness, which should make them popular methods among experimentalists. We report that while *TE* and *I* both allow accurate and consistent predictions of theoretical expectations in simple scenarios, *TE* is far less biased than *I* when little data is available and might thus offer more accurate measurements in experimental conditions where acquisition of large datasets is impossible or prohibitively costly. When used to analyze more realistic synthetic data (generated by a biophysical model of the Hodgkin-Huxley-type) however, the measure of *TE* can be strongly affected by systematic time frame shifts between inputs and outputs, a problem *I* is immune to by construction when using the direct method. Finally, we show how both measures perform when applied to assessing the energetic optimality of information transfer at biophysically-realistic synapses, and compare those results to results from the literature.

Our results illustrate the importance of systematically testing information theoretic measures on synthetic test data prior to designing experiments in order to fully understand how much data needs to be collected, and understand the trade-offs involved in using different measures. Our results also provide some guidance on what measure (*I* or *TE*) will perform best under different circumstances.

Results (3279 words)

In order to compare the relative performances of mutual information (*I*) and transfer entropy (*TE*) on measuring information flow at synapses, i.e. when comparing the input to a synapse to the output generated by the postsynaptic neuron, we started by generating synthetic binary input and output spike trains. To mimic the transmission of action potentials at thalamic relay cells, which we have studied and modelled before [8], we generated as input random binary Poisson spike trains at a fixed frequency matching what we had observed experimentally. That input sequence was then copied to form the output sequence with deletions occurring with a certain probability (non-transmitted spikes), and with additions (to mimic spontaneous firing of the output neuron). Numerical values derived from previous experiments for these simulations are given in Table 1 below. We have also previously studied transmission of information at the cortical synapse between thalamic relay cells and layer 4 spiny stellate cells [10]. We thus generated a second additional set of simulations using parameters matching the experimental observations for that second scenario. Numerical values for those simulations are also given in Table 1. Note that in both of these scenarios, information flows unidirectionally in a feed-forward manner (see [8] and [10] for further details).

To compute the mutual information *I* between those input and output sequences, we used the so-called direct method. The direct method requires repeating sequences (see Methods), while this isn't necessary for computing the transfer entropy (*TE*). We thus generated in each case two datasets, one with repetitions to use with the direct method to compute *I* (Figure 1A), and one without repetitions to compute *TE* (Figure 1B). In each case, the two datasets had the same duration, with the duration of the repeating sequence multiplied by the number of repetitions used

Table 1. Parameters for generation of inputs and outputs of random spike trains.

Scenario	Input frequency (Hz)	Transmission failure probability (per input spike)	Spontaneous firing probability (per output time bin)
Thalamic relay synapse	20 ¶	0.8 ¶	0.00084 ¶
Cortical layer 4 spiny stellate cell synapse	4 §	0.9 §	0.0024 §

¶ From ref. [8].

§ From ref. [10].

to compute *I* equating the total duration of the dataset generated to compute *TE*. Unless mentioned otherwise, the time bin is always 3 ms in duration, approximately the temporal extension of an individual action potential.

The mutual information (*I*) can be defined as the difference between the two entropies H_{total} and H_{noise} ($I = H_{total} - H_{noise}$; see Methods), and these entropies are typically calculated for 'words' of a certain length. For instance, the binary sequence 000110100100100101 can be segmented in words of length 1 yielding the words 0|0|0|1|1|..., or segmented in words of any other length. For words of length 3 for instance, it would yield 000|110|100|100|100|... In neuroscience, using long words is important to accurately capture and account for the information carried by stereotypical temporal patterns of spikes, if any [3, 11]. However, using longer and longer words can lead to significant biases in estimating H_{total} and H_{noise}, and eventually *I*, when using a finite dataset to build estimates, which they are always. This can be a serious limitation when trying to use *I* in an

experimental context where only limited data is available. To compensate for this, the so-called direct method includes two corrections when calculating *I*. The first correction factor is meant to extrapolate the true value of each individually computed entropies H_{total} and H_{noise} entering in the calculation of *I* to a dataset of infinite size, while the second correction is meant to extrapolate entropy estimates for both H_{total} and H_{noise} to infinitely long words. We have previously reported that these corrections did not lead to large changes in evaluating *I* when using limited experimental datasets with statistics similar to the datasets we generated here [8]. Additionally, we have had difficulties in reliably using the first above-mentioned correction (in [8] and here). As a consequence, here, we only implemented the second above-mentioned correction when using the direct method to compute *I*. Figure 2 shows how this correction was implemented. For *I*, H_{total} and H_{noise} were plotted as a function of the inverse of the word length and fitted with a linear function. We then extrapolated the fits to 0 (i.e. to words of infinite length) and took the difference of those values to calculate the mutual information (*I*). For transfer entropy, similar curves were computed, one for 'raw transfer entropy' and one for 'noise transfer entropy' (see Materials & Methods) [10, 17]. *TE* was then computed as the difference between those two curves, like for *I*. *TE* was then fitted with a linear relationship and extrapolated to infinite word lengths.

In that relatively simple scenario of action potentials being transmitted at a single relay synapse with set probabilities, $TE = I$, and it is possible to calculate that value exactly (see Methods). Figure 3 shows a comparison between that theoretical value, and *I* and *TE* calculated as described above for parameters corresponding to action potential transmission at thalamic relay cells (A) [8] or layer 4 spiny stellate cells (B) [10] (see Table 1 for details) for datasets of increasing sizes, plotted as the equivalent number of repetitions as used to calculate *I* using the direct method. The calculated theoretical value for the thalamic relay cell scenario (Figure 3A, first line in Table

1 [8]) was 15.47 bits/sec. The calculated theoretical value for the layer 4 spiny stellate cell scenario (Figure 3B, second line in Table 1 [10]) was 1.54 bits/sec. Both these values closely match the experimental and theoretical results we had previously published. While it is apparent that both *TE* and *I* eventually converge to the expected theoretical value given a dataset of sufficient size, *TE* appears to perform vastly better for small datasets, converging faster (i.e. for smaller datasets) than *I* to the correct theoretical value. Furthermore, like *I*, *TE* overestimates the expected value of transmitted information, but as illustrated in the insets in Figure 3A and B, does so to a much lesser extent than *I*, even for very short datasets.

Spike trains *in situ*, however, rarely display Poisson statistics. In the visual pathway for instance, spike trains impinging from the retina onto thalamic relay cells are characterized by non-Poissonian statistics with a high probability of spike doublets with short (~10 ms) interspike intervals. In order to test if the results of Figure 3 depend on the temporal structure of the input spike train, we generated a second set of simulations as above, but replacing the Poisson input spike trains with spike trains generated using statistics matching *in situ* recordings. To do so, we computed the interspike interval distribution of input spike trains from experimental data collected in [8] and [10] (Figure 4A and B insets). We then used these to calculate the cumulative distribution function (CDF) of interspike intervals (Figure 4 insets). Finally, we used the cumulative distribution function to generate spike trains as a series of interspike intervals with a temporal structure matching what had been observed in experiments. Under these new circumstances, there is *a priori* no expectation that *TE* should be equal to *I*, and it is not possible to calculate the expected theoretical values simply. Figure 4 shows, however, that this does not significantly change the behavior of *TE* and *I* with respect to the size of the dataset used to compute them, with both *TE* and *I* converging to values very similar to the values in Figure 3 given a large

enough dataset. Like in Figure 3, *TE* appears to significantly outperform *I* for short datasets. This prompts the question of why *TE* appears to systematically outperform *I* in the two relatively simple scenarios we tested here (Figures 3 and 4). In the Discussion below, we attempt to provide an answer to this question based on the structure of both measures.

We then wanted to test these comparisons using a more realistic model to generate output spike trains, focusing on the first scenario tested above: information transmission at thalamic relay synapses. To do so, we used an experimentally-calibrated biophysical single-compartment Hodgkin-Huxley-type model of thalamic relay cells (see [8] for details about the calibration of that model and Methods here for a detailed description of the model). To generate input for that model, we used the same procedure as described above for Figure 4, generating input spike trains with a temporal structure matching what is observed in *in vivo* experiments. We then convolved these spikes trains with unitary synaptic conductance extracted from experimental data in [8] (see Methods). Figure 5A shows a sample input spike train generated in this way, the corresponding synaptic conductance injected into the Hodgkin-Huxley-type model, the output spike train generated by that model in response to that synaptic input, and finally, the binarized output corresponding to that output spike train. We then applied *TE* and *I* on these binarized input and output spike trains in the same way as above. Figure 5C shows a plot comparing *TE* and *I* for datasets of increasing sizes (similar to what we plotted above in Figures 3A and 4A). It is immediately apparent that while the qualitative behavior of *I* and *TE* is not strikingly different than what we observed for spike trains generated using simple spiking and transmission probabilities as above, they do not appear to converge to similar values like they did previously (Figures 3A and 4A). *TE*, in particular, appears to converge to a value far lower than what we had previously

observed (compare for instance Figure 5A to Figure 4A; the values reached by *I* in these two panels are more comparable).

The MATLAB package we used (see Materials & Methods) offers the possibility to compute *TE* with various time frame shifts between the input and output sequences [17]. This is equivalent to simply shifting the frame of reference of the output binary sequence by a set number of bins. Output spikes could be significantly delayed with respect to input spikes for instance in the case of long conduction delays. Here, however, delays are more likely to be due to the specific integration properties of the postsynaptic neuron. Figure 5B shows the evolution of *TE* when plotted versus this shift. It shows that *TE* does not peak at shift 0, but rather raises from about 2bits/s for no frame shift, to about 4.5bits/s at a frame shift of about 15 ms (5 time bins), before decaying again. Note that this is different from what we have observed in a previous study at the cortical synapse between thalamic relay cells and layer 4 spiny stellate cells, where we observed instead that *TE* between the input and output sequences was maximal for no frame shift, and simply decaying for positive shifts [10] (Figure S2 therein). While we have no *a priori* explanation for this finding, this is obviously due to a relatively systematic frame shift between the timing of incoming action potentials and the timing of outgoing action potentials. The rise times of the synaptic conductances and of the membrane potential, i.e. the membrane time constant, might play a role in this observation. By comparison, in the above-mentioned study ([10]), neurons were simulated, or experimentally-recorded, in high-conductance states, which would have shortened their membrane time constant [18]. Here, conductances are only briefly opened after an incoming input action potential. However, this is unlikely to be the full explanation and we hypothesize that this systematic frame shift might also be partially related to the fact that a single action potential will often fail to elicit an output action potential (failed transmission). Instead, it has been known

for a while that two incoming action potentials in short succession are usually necessary to trigger an output action potential at those synapses [19], and the input binary sequence we generated based on *in vivo* recordings have non-Poissonian statistics with a preferred interspike interval of ~10 ms [8] (Figure 4A inset). This effect is apparent in Figure 5A, when two consecutive incoming action potentials separated by 18 ms fail to elicit an output action potential (marked by stars). Immediately after that, two consecutive incoming action potentials separated by 9 ms trigger an output action potential after the second of those has reached the synapse (marked by squares).

In Figure 5C, we plotted the two most obvious measures that can be derived from the curve in Figure 5B, the peak value of *TE* and the integral value of *TE* (calculated over frame shifts from 0 ms to 90 ms) [17]. Both display the same rapid convergence to a stable value, with respect to the size of the dataset, than *TE* in Figures 3 and 4, but they both converge to quite different values (4.4 bits/s and 41.4 bits/s respectively), both different than the value *I* converges to (23.1 bits/s).

In order to test the hypothesis that the strong discrepancy we observed here between matching *I* and *TE* predictions in the simple scenarios graphed in Figures 3 and 4, where action potentials are transmitted from the input to the output sequence in the same corresponding time bin with no temporal frame shift (unless transmission fails), and mismatching *I* and *TE* predictions when using a Hodgkin-Huxley-type model with non-trivial transmission properties (Figure 5A-C), we reproduced the simulations of Figure 3A using simple transmission probabilities and Poissonian inputs. However, this time, we implemented a random shift between every input and output action potentials. This random (positive only) shift followed a Gaussian distribution with mean = 16 ms and standard deviation = 3.7 ms (see Materials & Methods). In the case of *I*, that random shift was systematically reproduced for each pair of inputs and outputs action potentials in each repetition. The inset in Figure 5B shows how the random shift broadens the distribution

of TE versus the frame shift. Figure 5D shows that this was sufficient to reproduce the discrepancies described above (compare Figures 3A, 5C and 5D), suggesting that this is indeed the reason behind this observation. Interestingly, by construction, when using the direct method, *I* is immune to that issue (again, compare Figures 3A, 5C and 5D).

Thus, while *TE* offers great stability in its predictions even for relatively small datasets, which will be of interest to experimentalists who might not be able to collect very large datasets due to various technical constraints, the values obtained through that method might not be directly comparable to the mutual information in non-trivial scenarios, in which any kind of systematic frame shift can be expected.

Finally, we wanted to compare the application of these two information theoretic measures to a biophysically relevant scenario. We have recently demonstrated in computational models and experiments that a number of synaptic features can be explained as a trade-off between information flow and energy consumption: the low release probability of weak central synapses [7, 9], the postsynaptic conductance at strong thalamic relay synapses [8], and the postsynaptic conductance at weak cortical synapses [10]. All these studies demonstrate that synapses and neurons, appear to be designed to maximize information flow per energy units (bits/ATP) rather than per time units (bits/sec). More recently, similar results have also been obtained at hippocampal synapses [13]. These results suggest this as a widespread principle in the brain. The existence of an information flow over energy optimum is not surprising. It stems from the basic principle that energy consumption scales roughly linearly with the biophysical parameters we studied (release probability or postsynaptic conductance), while information flow scales sigmoidally with those parameters, because of the necessity to overcome noise [7, 9], or energetic barriers (the threshold for action potential generation) [8, 10]. Thus, information flow over energy, dividing a sigmoid

by a linear function, leads to a single well-defined optimum. In particular, we have demonstrated that very mechanism for the postsynaptic conductance at strong thalamic relay synapses, using the same experimentally-calibrated biophysical Hodgkin-Huxley-type model we have used here [8]. Note however that in that original study, we injected in the Hodgkin-Huxley-type model experimentally recorded conductances with a rather limited number of repetitions (N = 5), to estimate *I* with the direct method. Here, using the approach highlighted above to generate *in vivo*-like input spike trains and synaptic conductances (see Figure 5A), we can stimulate that model with any number of repetitions of any length.

We thus proceeded to reproduce here the finding that thalamic relay synapses maximize information flow per energy units (bits/ATP) at experimentally-observed physiological conductances, quantifying information flow using both the mutual information and transfer entropy as above. Specifically, we injected *in vivo*-like conductances generated as in Figure 5A in our experimentally-calibrated biophysical Hodgkin-Huxley-type model of thalamic relay cells. We then varied the postsynaptic conductance by applying a multiplying factor (gain) between 0 and 10 to the injected conductance, like we have done in previous studies [8, 10], the experimentally-observed physiological conductance corresponding to gain = 1. Figure 6A shows information flow across the thalamic relay synapse when modulating the synaptic conductance (gain), quantified as before with *I* (using the direct method) or *TE*. In the latter case, both *TE* measured at its peak frame shift (TE_{peak}; see Figure 5C), or integrated over all frame shifts (TE_{sum}) are reported. As expected from Figure 5C, each measure (*I*, TE_{peak} and TE_{sum}) yields different results, but all three grow sigmoidally with the gain. We also additionally quantified the corresponding energy consumption by counting the ATP molecules necessary to fuel the Na,K-ATPase electrogenic pump that restores the ion gradients disturbed by ions flowing through

postsynaptic receptors and ion channels (see Methods). Figure 6B shows the energy consumption concomitant to neuronal activity, when accounting only for ionic flows at the modulated postsynaptic conductance, or when accounting for all ionic flows in the postsynaptic neuron (i.e. including also the ions flowing through the voltage-gated channels that underlie action potentials; total energy budget). In both cases, we observed a roughly linear relationship between the modulated postsynaptic conductance and the energy consumption expressed in ATP molecules consumed per second. Note however that in the case of the total energy budget, the relationship is piecewise linear, with a different, smaller slope, for low postsynaptic gains at which no output action potential is generated.

We can then evaluate how information flow, quantified using either I or TE, relates to the concomitant energy consumption at different postsynaptic gains. Figure 6C shows that this relationship (normalized to its peak) has a single well-defined peak close to the physiological gain of the synapse (gain = 1), when quantifying information flow using I or TE_{sum}. In both those cases, this 'energetic efficiency of information flow' curve peaks close to gain = 1 and closely resembles what has been reported earlier, either for experimental data or in computational models [7-10]. When quantifying the value of information flow using TE_{peak} however, this relationship appears much broader with no clear discernable peak between gains ~1 and ~5. Finally, Figure 6D shows the same results, but using the total energy budget to quantify the energy consumption of the neuron (higher traces in Figure 6B), rather than only the energy consumption imparted by the modulated postsynaptic conductance. The results are however broadly similar to those displayed in Figure 6C. Again, the curves corresponding to I and TE_{sum} match each other and match what has been reported earlier, while the curve corresponding to TE_{peak} has a broad profile with no clear peak. Therefore, it appears that even though I and TE_{sum} yield different raw values (see Figures

5C and 6A), their predictions can be compared in the current scenario when normalized, unlike TE_{peak}. TE_{sum} might therefore be a better alternative than TE_{peak} for comparing to mutual information across studies, although this will have to be systematically verified on a case by case basis.

Discussion (1096 words)

Here, we set out to evaluate systematically the performance and trade-offs of mutual information and transfer entropy, when applied to binary spike trains in the context of information flowing between individual neurons and across synapses. While these information theoretic measures are popular among theoretical and computational neuroscientists, it can be argued that they have found only limited usage among experimentalists due to the relative complexity in applying them. This is especially true in experimental contexts where acquisition of large datasets is prohibitive. The systematic biases these measures suffer from when applied to limited datasets has led to the development of a number of corrective techniques [6], which, while they improve their performances, do not necessary help in making these techniques more widely accessible. However, it can be argued that more widespread use of information theory will be essential going forward in neuroscience, especially when considering the development of *normative* theories [20] linking information processing to the energetic capacity of the brain [7-10, 21].

Here, we have used a MATLAB package readily available from online repositories [17] to calculate transfer entropy (*TE*) in simple scenarios. We show in these scenarios that *TE* outperforms mutual information when little data is available (Figures 3 and 4) with little need to apply corrective measures. Note that we did alternatively perform the simulations displayed in Figures 3 and 4, but without applying the correction to infinite word lengths (see Figure 2), and that did not change qualitatively the results (not shown). This is however to be expected when using Poisson spike trains (Figure 3), where each input action potential is generated independently from preceding action potentials and no long-range correlations are present in the input sequence [3, 11].

The mutual information is defined as the difference between the two entropies H_{total} and H_{noise}. For words of length N, H_{total} calls for the evaluation of 2^N independent probabilities from the dataset. H_{noise} on the other hand calls for the evaluation of 2^{2N} independent probabilities from the same dataset. It is thus common that many of these independent probabilities appearing in the calculation of H_{noise} will be evaluated to be null because no corresponding event will be observed in a limited dataset. As a consequence, it is common to underevaluate H_{noise}. The same issue is also true for H_{total}, but since far fewer probabilities need to be estimated to calculate H_{total}, it is commonly less underestimated than H_{noise}. The common outcome is then that the mutual information is grossly overestimated for limited datasets. This overestimate decreases in amplitude as the size of the dataset increases (see Figures 3 and 4, and ref. [6] for an excellent discussion of these issues).

Transfer entropy, on the other hand, does not seem to suffer from that problem, at least not in a similar amplitude (Figures 3 and 4). We do not have at this time a definitive explanation as to why that is. Transfer entropy can also be written as the difference between the two entropies $H(\mathbf{X}^-|\mathbf{Y}^-)$ and $H(\mathbf{X}^-|Y^+,\mathbf{Y}^-)$ (see Materials and Methods), but these are two conditional entropies, i.e. they both call for the evaluation of 2^{2N} independent probabilities for words of length N. We prudently speculate that this contributes to balancing the systematic errors in evaluating each entropy, and that this leads to a better overall estimation of *TE*, even with limited data (see Figures 3 and 4).

These results suggest that *TE* might be a better choice for experimentalists over *I* due to its lower sensitivity to the size of the dataset, and due to its relative simplicity of use. However, the situation gets more complex when considering a more realistic biophysical scenario. Figure 5 shows that when systematic frame shifts occur between input and output sequences, i.e. when the

input and output action potentials do not happen in matching time bins due to transmission delays or due to the integrative properties of the neuron under consideration, the use of *TE* requires a more careful examination of the detailed biophysics at play. It also becomes difficult to directly compare results obtained with *I* and *TE*. In Figure 5, using a Hodgkin-Huxley-type model for thalamic relay cells, we show that the specific integration properties of that modelled neuron lead to transfer entropy being 'distributed' over time shifts between the input and output sequences (see Figure 5B). As noted above however, this is not always the case, as we have observed in a previous study (in a different setting) that *TE* between the input and output sequences was maximal for no frame shift, and simply decaying for positive shifts [10]. We additionally provide convincing evidence that this observation is due to a systematic shift by recapitulating the results of Figure 5C using systematically shifted Poisson spike trains. By construction, the mutual information estimated using the 'direct' method is immune to that issue.

The fact that *TE* appears to be distributed over time thus poses the question of what feature of *TE* to actually use. The two most obvious features of the curve in Figure 5B are its peak and its integral. Both appear very stable with respect to the size of the dataset but none of them matches the value predicted by mutual information (Figure 5C), making systematic comparisons between these measures difficult. In Figure 6 we apply both features of *TE* (peak and integral) and *I* to the evaluation of the energetic optimality of information transfer at thalamic relay synapses [8]. Our results recapitulate the original experimental and computational finding that those synapses appear to maximize not the information flow (bits/sec), but the ratio of information per concomitant energy use (bits/ATP; see Figure 6). In that context, the integral of *TE* is the feature that matches best the curve obtained using *I*. Both predict the energetically optimal gain of the synapse to be

close to 1, the physiological gain of the synapse. The peak of *TE* on the contrary yields markedly different predictions.

These results suggest that a detailed computational study of the system under investigation should be systematically performed prior to applying transfer entropy to experimental data, and that when experimental conditions permit, mutual information estimated using the 'direct' method might provide a more straightforward applied measure. In particular, it is important to test for the presence of transmission delays or of integrative properties in the system under investigation that might lead to the kind of effects described here. When it cannot be demonstrated, as we have done in [10], that output action potentials occur mostly in the same time bin than impinging action potentials, it might be best to use mutual information instead.

Materials & Methods (1945 words)

We are interested in characterizing the properties and performance of two information theoretic measures, mutual information (*I*) and transfer entropy (*TE*), in the context of energy consumption and efficacy of information transfer at individual synapses. In the following, we will apply both these measures to assess information flow from the binary input spike train of a synapse to the binary output spike train generated by the postsynaptic neuron. Below, we start by describing how the test data for this characterization were generated.

Synthetic spike trains

The first dataset we tested transfer entropy and mutual information on are synthetic Poisson spike trains. In this scenario, the input was a randomly generated Poisson spike train, and the output was created by copying that input spike train, applying for each transmitted spike a certain probability of transmission failure, and for each time bin a certain probability of spontaneous firing, even in the absence of an input spike in the matching input time bin. Unless specified otherwise, all time bins in this manuscript are 3 ms. This simple scenario matches the propagation of individual spikes in the visual pathway, and numerical values for the probabilities were derived from experiments measuring the propagation of spikes at thalamic relay cells in the lateral geniculate nucleus and between thalamic relay cells and layer 4 spiny stellate cells in the primary visual cortex (see numerical values in Table 1 in the Results section above). Unless stated otherwise, each result is the average of 10 independent simulations.

Additionally, in Figure 5, we tested adding a random frame shift between input and output action potentials. We added to the timing of output action potentials a random gaussian delay centered at +16 ms with a standard deviation of ±3.7 ms.

Input based on biological recordings

The first step towards a more realistic model was to generate inputs similar to those that can be observed impinging onto the cells of interest. To do that, input spike trains were generated according to the distribution of input interspike intervals recorded *in vivo*. The *in vivo* distribution of interspike intervals was used to generate the cumulative distribution function of the intervals, and this function was in turn used as the generative function for the input (see insets in Figure 3). The output was generated as before, with probabilities of failure of transmission and spontaneous firing. 10 simulations were performed, and the mean was taken. Numerical values follow the values given in Table 1.

Hodgkin-Huxley type model

While using transmission and spontaneous firing probabilities as above is expedient, this can never fully capture the complexity of the biophysical processes that lead to spiking. In order to remediate to that issue, we adapted the single-compartment Hodgkin-Huxley-type model of thalamic relay cells by Harris *et al*. [8]. Details about how that model was carefully calibrated onto experimental data can be found in [8]. Our adjustments are detailed below.

Briefly, the Hodgkin-Huxley-type model of thalamic relay cells in the lateral geniculate nucleus was adapted from earlier models [22-24], and follows the formalism devised by Hodgkin and Huxley with:

$$C_m \frac{dV}{dt} = -\sum_j i_j - i_{Hold} - i_{Syn}, \quad (1)$$

where C_m = 1 µF/cm^2 is the membrane capacitance, V is the membrane voltage (in mV), i_{Hold} is the injected current, i_{Syn} is the synaptic current and i_j are the intrinsic currents. The cell surface area was $1.52 \cdot 10^{-4}$ cm^2. All currents and conductances are subsequently reported per unit surface area (cm^2). Following Bazhenov and colleagues [22], the intrinsic currents included a leak current i_L, a potassium leak current i_{KL}, an A-type potassium current i_A, a T-type low threshold calcium current i_T, an h-current i_h, a fast sodium current i_{Na} and a fast potassium current i_K. All the intrinsic currents had the same general form:

$$i = g m^M h^N (V - E), \quad (2)$$

where for each current i, g is the maximal conductance, $m(t)$ is the activation variable, $h(t)$ is the inactivation variable, E is the reversal potential and M and N are the number of independent activation and inactivation gates.

The i_h current was given by:

$$i_h = g_{\max} O (V - E_h), \quad (3)$$

with E_h = –43 mV [25]. $g_{\max}$ = 0.0254 mS/cm^2 was set to match experimental data (see [8] for further details). The time dependence of the gating variable O was defined by:

$$\frac{dO}{dt} = \frac{1}{\tau_O}(O_\infty - O), \quad (4)$$

with time constant $\tau_O = 1/\left[e^{(-14.59 - 0.086 \cdot V)} + e^{(-1.87 + 0.0701 \cdot V)}\right]$ (in ms) and steady-state variable $O_\infty = 1/\left[1 + e^{((V+75)/5.5)}\right]$ [25].

The leak currents were given by:

$$i_L = g_{\mathrm{L}}(V - E_L) \tag{5}$$

and:

$$i_{KL} = g_{\mathrm{KL}}(V - E_K), \tag{6}$$

with E_L = –70 mV [24]. E_K was set to match the effective potassium reversal potential used in the experiments in [8]: E_K = –105 mV, while g_{L} = 0.025 mS/cm^2 and g_{KL} = 0.025 mS/cm^2 were manually adjusted to match both the average input resistance at the resting membrane potential and the resting membrane potential as recorded in experiments.

The A-type potassium current was given by:

$$i_A = g_A m^M h^N (V - E_K), \tag{7}$$

with M = 4 and N = 1. The time dependence for m and h was defined as for O, with:

$$m_\infty = 1/\left[1 + e^{(-(V+60)/8.5)}\right], \tag{8}$$

$$\tau_m = 0.1 + 0.27/\left[e^{((V+35.8)/19.7)} + e^{(-(V+79.7)/12.7)}\right], \tag{9}$$

$$h_\infty = 1/\left[1 + e^{((V+78)/6)}\right] \tag{10}$$

and:

$$\tau_h = 0.27/\left[e^{((V+46)/5)} + e^{(-(V+238)/37.5)}\right], \tag{11}$$

if V < –63 mV, and τ_h =5.1 ms otherwise [22, 23].

The T-type calcium current was given by:

$$i_T = g_T m^M h^N (V - E_T), \tag{12}$$

with M = 2 and N = 1. The time dependence for m and h was defined as for O, with:

$$m_\infty = 1/\left[1 + e^{(-(V+57)/6.2)}\right], \tag{13}$$

$$\tau_m = 0.13 + 0.22/\left[e^{(-(V+132)/16.7)} + e^{((V+16.8)/18.2)}\right], \tag{14}$$

$$h_\infty = 1/\left[1 + e^{((V+83)/4)}\right] \tag{15}$$

and:

$$\tau_h = 8.2 + \left[56.6 + 0.27 \cdot e^{((V+115.2)/5)}\right]/\left[1 + e^{((V+86)/3.2)}\right]. \tag{16}$$

E_T is given by $E_T = RT/2F \cdot log\left(\mathrm{Ca}_0^{2+}/\mathrm{Ca}^{2+}\right)$ with F = 96489 C/mol the Faraday constant, R = 8.314 J mol^{-1} K^{-1} the gas constant, T = 309°K the temperature and Ca_0^{2+} = 2 mM the extracellular calcium concentration. The intracellular calcium dynamics were defined by:

$$\frac{d\mathrm{Ca}^{2+}}{dt} = -\frac{1}{\tau_{\mathrm{Ca}}}\left(\mathrm{Ca}^{2+} - \mathrm{Ca}_i^{2+}\right) - A\, i_T, \tag{17}$$

with Ca_i^{2+} = 2.410^{-4} mM, the baseline intracellular calcium concentration, and A = 5.1810^{-5} mM cm^2 ms^{-1} μA^{-1}, a constant.

The fast sodium current was defined by:

$$i_{Na} = g_{Na} m^3 h (V - E_{Na}), \tag{18}$$

with E_{Na} = +90 mV. The maximal conductance g_{Na} = 4.4 mS/cm^2 was also set to match experimental data [8]. The time dependence for m and h was defined by:

$$\frac{dx}{dt} = \alpha_x(1 - x) - \beta_x x, \tag{19}$$

where x stands for either h or m and with [26]:

$$\alpha_m = 0.32\left[13.1 - V + V_{\mathrm{shift}}^{\mathrm{Na}}\right]/\left[e^{((13.1-V+V_{\mathrm{shift}}^{\mathrm{Na}})/4)} - 1\right], \tag{20}$$

$$\beta_m = 0.28\left[V - V_{\mathrm{shift}}^{\mathrm{Na}} - 40.1\right]/\left[e^{((V-V_{\mathrm{shift}}^{\mathrm{Na}}-40.1)/5)} - 1\right], \tag{21}$$

$$\alpha_h = 0.128\, e^{((17-V+V_{\mathrm{shift}}^{\mathrm{Na}})/18)} \tag{22}$$

and:

$$\beta_h = 4/\left[1 + e^{((40-V+V_{\mathrm{shift}}^{\mathrm{Na}})/5)}\right]. \tag{23}$$

The fast potassium current was given by:

$$i_K = g_K n^4 (V - E_K). \tag{24}$$

The maximal conductance g_K = 3.3 mS/cm^2 was set to match experiments (again see [8] for further details). The time dependence for *n* was defined as for the sodium gating variables *m* and *h* with:

$$\alpha_n = 0.032\left[15 - V + V_{\text{shift}}^{\text{K}}\right] / \left[e^{((15-V+V_{\text{shift}}^{K})/5)} - 1\right] \tag{25}$$

and:

$$\beta_n = 0.5\, e^{((10-V+V_{\text{shift}}^{\text{K}})/40)}. \tag{26}$$

V_{shift}^{Na}= –60.1 mV and V_{shift}^{K} = –62.5 mV were manually adjusted to allow the model to be depolarized to –55 mV without spontaneously spiking. i_{Hold} = -2.05 µA/cm^2 was set in subsequent simulations so as to hold the model at –55 mV. For a cell surface area of 1.5210^{-4} cm^2, this corresponds to an injected current of ~310 pA, similar to experimentally measured values of 30-550 pA [8]. g_A = 3 mS/cm^2 and g_T = 1.8 mS/cm^2 were set so that the model achieved an output frequency, when stimulated with the synaptic conductance recorded in [5], similar to the average frequency observed in experiments.

Synaptic currents

Individual synaptic currents experimentally recorded in ref. [8] were fitted by sums of exponentials:

$$g(t) = w\left(e^{-t/\tau_1} - e^{-t/\tau_2}\right), \tag{27}$$

yielding the parameters given in Table 2. The synaptic current is then given by:

$$i_{Syn} = -g_{\text{AMPA}}\left(V - E_{\text{excitatory}}\right) - g_{\text{NMDA}}\left(\frac{9.69}{1+0.1688\, e^{-0.0717\, V}}\right)\left(V - E_{\text{excitatory}}\right), \tag{28}$$

with $E_{\text{excitatory}}$ = 0 mV, and where g_{AMPA} and g_{NMDA} describe the time courses of individual AMPA and NMDA synaptic currents.

Table 2. Parameters of individual AMPA and NMDA synaptic conductances.

Conductance	*w*	τ_1 [ms]	τ_2 [ms]
AMPA	86.63	3.03	2.98
NMDA	0.31	45.04	3.79

Information theory

For two coupled physical systems that produce realizations *x* and *y* of random variables *X* and *Y*, the mutual information *I* is defined by [27]:

$$I(X,Y) = H(X) - H(X|Y), \tag{29}$$

where $H(X) = \sum_x p(x)$ is the Shannon entropy for the probability distribution *p(x) = p(X=x)* of the outcome *x* of the random variable *X* and $H(X|Y) = \sum_y p(y) \sum_x p(x|y) \log_2(1/p(x|y))$ is the conditional entropy. The mutual information can then be written as:

$$I(X,Y) = \sum_x \sum_y p(x,y) \log_2 \frac{p(x,y)}{p(x)p(y)}. \tag{30}$$

Because *H(X)* is the total average information in the variable *X* and *H(X|Y)* is the average information that is unique to *X*, the mutual information represents the shared information between the two processes or, in other words, the deviation from independence of the two processes. Equations (29) and (30) are symmetric under the exchange of the two variables and thus do not contain any directional indication.

The transfer entropy TE of the same random variables is defined as [27]:

$$TE(X \to Y) = I(X^-, Y^+|Y^-) = H(X^-|Y^-) - H(X^-|Y^+, Y^-) \tag{31}$$

$$= \sum_{y^-} \sum_{y^+} \sum_{x^-} p(x^-, y^+, y^-) \log_2 \frac{p(y^-)p(x^-, y^+, y^-)}{p(x^-, y^-)p(y^+, y^-)} \quad (32)$$

where $\boldsymbol{X}^-$ and $\boldsymbol{Y}^-$ denote the past state of the two processes with outcomes x^- and y^-, and Y^+ is the future random variable of Y with outcome y^-.

The transfer entropy represents the amount of predictive information actually transferred from process X to process Y. As can be seen in Equation (31), the transfer entropy is not usually symmetric under the exchange of X and Y.

Equivalence between I and TE

In a scenario where $\boldsymbol{X}^-$ and Y^+ are both independent of $\boldsymbol{Y}^-$ (see Equation (31) above), which is for instance realized in the scenarios displayed in Figure 3, the transfer entropy TE can be written as:

$$\begin{aligned} TE(X \to Y) &= I(\boldsymbol{X}^-, Y^+ | Y^-) \\ &= \sum_{y^-} \sum_{y^+} \sum_{x^-} p(x^-, y^+) p(y^-) \log \frac{p(y^-) p(x^-, y^+) p(y^-)}{p(x^-) p(y^-) p(y^+) p(y^-)} \\ &= \underbrace{\sum_{y^-} p(y^-)}_{1} \sum_{y^+} \sum_{x^-} p(x^-, y^+) \log \frac{p(x^-, y^+)}{p(x^-) p(y^+)} \\ &= I(\boldsymbol{X}^-, Y^+) \end{aligned} .$$

In this simplified case, TE is equivalent to I. In Figure 3, we have used this fact, together with the fact that with knowledge of all the probabilities entering in Equations (29)-(30), we can directly calculate – rather than estimate from the data – the theoretical value of $TE = I$.

Estimation methods

Poisson spike trains are already binary. The output spike trains generated with the Hodgkin-Huxley-type model were binarized using a time bin of 3 ms using the function *findpeaks* in MATLAB (The Mathworks, Natick MA).

Mutual information was evaluated using the so-called 'direct' method devised by Strong and colleagues [11]. As noted in the Results section, we only applied the second of the two corrections from the original method, i.e. the extrapolation to infinite word lengths (see Figure 2, Results and ref. [8] for further details).

Transfer entropy was calculated using the MATLAB package by Ito and colleagues [17], following similar procedures as in [10]. In particular, a baseline value TE_{noise} was calculated by randomly shuffling words in the output and calculating the transfer entropy between the input and the shuffled output. The transfer entropy values reported in the manuscript are $TE = TE_{raw} - TE_{noise}$ after extrapolation to words of infinite lengths similar to the correction applied for calculations of the mutual information.

Repetitions of 128 s were used for calculations of the mutual information using the 'direct' method. Each data point is the average of N = 10 independent simulations. For direct comparison, we used datasets of the same length for calculations of the transfer entropy. For instance, if we had 50 repetitions of 128 s for the 'direct' method (6400 s total), we used a single input spike train of 6400 s for calculation of the transfer entropy.

Information and energy

To calculate information transfer at the simulated synapse, output spike trains were processed as detailed above. To calculate the metabolic cost incurred by the modelled cell, the Na^+ component of i_{Syn} was integrated and converted to the corresponding ATP consumption per unit time, while

the same procedure was followed for i_{Na}, the Na^+ component of i_h and Ca^{2+} entry via i_T, and this was added to the ATP used on i_{Syn}. For i_{Syn}, the conductance was scaled by 7/13 (derived from the reversal potentials $E_{excitatory}$ = 0 mV, E_{Na} = +90 mV and E_K = -105 mV) and multiplied by V - E_{Na} to calculate the contribution of sodium ions. For i_h, the conductance was scaled by $(E_K - E_h)/(E_K - E_{Na})$ and multiplied by V - E_{Na} to isolate the contribution of sodium ions. For i_T, we assumed that each calcium ion is exchanged for 3 sodium ions [28]. For each gain in Figure 6, N = 10 simulations were performed, and the mean was taken.

Simulations

Simulations were run using custom-written MATLAB scripts (The Mathworks, Natick MA). Differential equations were integrated using the built-in solver *ode15s* with an integration time step dt = 0.05 ms. All results presented are the mean of N = 10 independent simulations.

References

1. Cobb M. The Idea of the Brain: A History: Profile Books; 2020. 480 p.

2. Shannon CE. A mathematical theory of communication. Bell Syst Tech J. 1948;27:379–423.

3. Reinagel P, Reid RC. Temporal coding of visual information in the thalamus. J Neurosci. 2000;20(14):5392-400. PubMed PMID: 10884324.

4. Dayan P, Abbott LF. Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems. Cambridge MA: MIT Press; 2001. 576 p.

5. London M, Schreibman A, Hausser M, Larkum ME, Segev I. The information efficacy of a synapse. Nat Neurosci. 2002;5(4):332-40. doi: 10.1038/nn826. PubMed PMID: 11896396.

6. Panzeri S, Senatore R, Montemurro MA, Petersen RS. Correcting for the sampling bias problem in spike train information measures. J Neurophysiol. 2007;98(3):1064-72. doi: 10.1152/jn.00559.2007. PubMed PMID: 17615128.

7. Harris JJ, Jolivet R, Attwell D. Synaptic Energy Use and Supply. Neuron. 2012;75(5):762-77. doi: 10.1016/j.neuron.2012.08.019. PubMed PMID: WOS:000308684300005.

8. Harris JJ, Jolivet R, Engl E, Attwell D. Energy-Efficient Information Transfer by Visual Pathway Synapses. Curr Biol. 2015;25(24):3151-60. doi: 10.1016/j.cub.2015.10.063. PubMed PMID: WOS:000367233400014.

9. Conrad M, Engl E, Jolivet R. Energy use constrains brain information processing. Technical Digest - International Electron Devices Meeting. 2018:11.3.1-.3.3.

10. Harris JJ, Engl E, Attwell D, Jolivet RB. Energy-efficient information transfer at thalamocortical synapses. PLoS Comput Biol. 2019;15(8):e1007226. doi: 10.1371/journal.pcbi.1007226.

11. Strong SP, Koberle R, de Ruyter van Stevenick R, Bialek W. Entropy and Information in Neural Spike Trains. Physical review letters. 1998;80(1):197-200.

12. Schreiber T. Measuring information transfer. Physical review letters. 2000;85(2):461-4. Epub 2000/09/16. doi: 10.1103/PhysRevLett.85.461. PubMed PMID: 10991308.

13. Mahajan G, Nadkarni S. Local design principles at hippocampal synapses revealed by an energy-information trade-off. bioRxiv. 2019. doi: 10.1101/748400.

14. Kostal L, Shinomoto S. Efficient information transfer by Poisson neurons. Math Biosci Eng. 2016;13(3):509-20. Epub 2016/04/24. doi: 10.3934/mbe.2016004. PubMed PMID: 27106184.

15. Kostal L, Lansky P, McDonnell MD. Metabolic cost of neuronal information in an empirical stimulus-response model. Biol Cybern. 2013;107(3):355-65. Epub 2013/03/08. doi: 10.1007/s00422-013-0554-6. PubMed PMID: 23467914.

16. Kostal L, Lansky P. Information capacity and its approximations under metabolic cost in a simple homogeneous population of neurons. Biosystems. 2013;112(3):265-75. Epub 2013/04/09. doi: 10.1016/j.biosystems.2013.03.019. PubMed PMID: 23562831.

17. Ito S, Hansen ME, Heiland R, Lumsdaine A, Litke AM, Beggs JM. Extending transfer entropy improves identification of effective connectivity in a spiking cortical network model. PLoS One. 2011;6(11):e27431. Epub 2011/11/22. doi: 10.1371/journal.pone.0027431. PubMed PMID: 22102894; PubMed Central PMCID: PMCPMC3216957.

18. Destexhe A, Rudolph M, Paré D. The high-conductance state of neocortical neurons in vivo. Nature Reviews Neuroscience. 2003;4:739-51.

19. Carandini M, Horton JC, Sincich LC. Thalamic filtering of retinal spike trains by postsynaptic summation. J Vis. 2007;7(14):20 1-11. Epub 2008/01/26. doi: 10.1167/7.14.20. PubMed PMID: 18217815; PubMed Central PMCID: PMCPMC2901808.

20. Levenstein D, Alvarez VA, Amarasingham A, Azab H, Gerkin RC, Hasenstaub A, et al. On the role of theory and modeling in neuroscience. arXiv. 2020. doi: https://arxiv.org/abs/2003.13825.

21. Aiello LC, Wheeler P. The Expensive-Tissue Hypothesis. Current Anthropology. 1995;36(2):199-221.

22. Bazhenov M, Timofeev I, Steriade M, Sejnowski TJ. Cellular and network models for intrathalamic augmenting responses during 10-Hz stimulation. Journal of Neurophysiology. 1998;79(5):2730-48. PubMed PMID: 9582241.

23. Bazhenov M, Timofeev I, Steriade M, Sejnowski TJ. Computational models of thalamocortical augmenting responses. Journal of Neuroscience. 1998;18(16):6444-65. PubMed PMID: 9698334.

24. McCormick DA, Huguenard JR. A Model of the Electrophysiological Properties of Thalamocortical Relay Neurons. Journal of Neurophysiology. 1992;68(4):1384-400. PubMed PMID: ISI:A1992JV73300034.

25. Huguenard JR, McCormick DA. Simulation of the currents involved in rhythmic oscillations in thalamic relay neurons. Journal of Neurophysiology. 1992;68(4):1373-83. PubMed PMID: 1279135.

26. Traub RD, Miles D. Neuronal Networks of the Hippocampus. Cambridge: CUP; 1991.

27. Wibral M, Vicente R, Lindner M. Transfer entropy in neuroscience. In: Wibral M, Vicente R, Lizier JT, editors. Directed Information Measures in Neuroscience: Springer; 2014. p. 3-36.

28. Attwell D, Laughlin SB. An energy budget for signaling in the grey matter of the brain. Journal of Cerebral Blood Flow and Metabolism. 2001;21:1133-45.

Figures

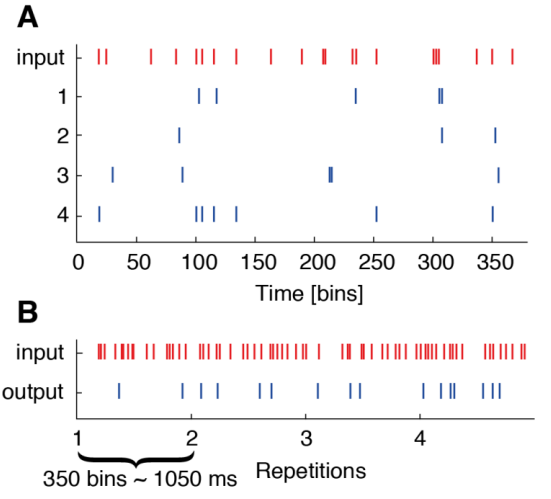

Figure 1. Sample input and output sequences for computing *I* and *TE*.

Examples of input and output data generated, with repetitions for mutual information calculations (*I*; to use with the direct method) (**A**), and without repetitions for transfer entropy calculations (*TE*) (**B**). Here, outputs are generated using transmission statistics derived from experiments in the thalamic relay cell scenario (see Methods and Table 1). In each case, two length-matched datasets are generated to compare the relative performances of *I* and *TE* for recordings of a certain duration. 1 time bin = 3 ms.

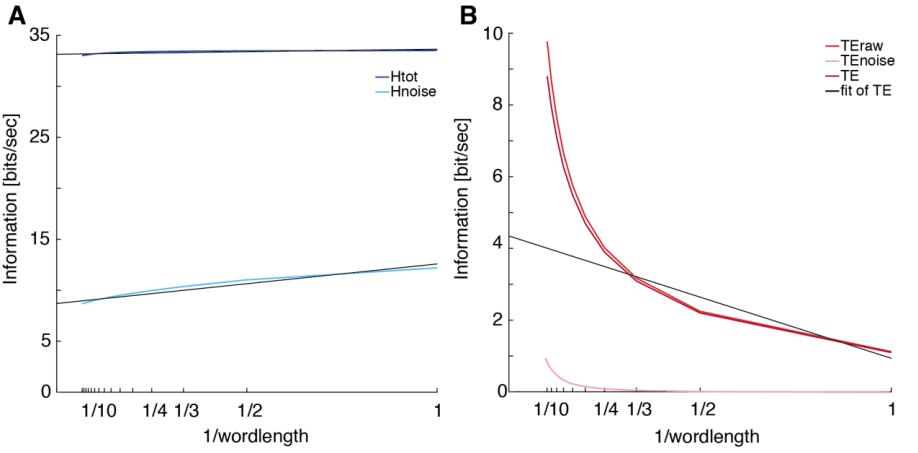

Figure 2. Extrapolation of mutual information and transfer entropy values to infinite word lengths.

Example of extrapolation to infinite word lengths for mutual information (**A**) and transfer entropy calculations (**B**). Mutual information is calculated using the 'direct' method and the second correction of that method is applied to extrapolate both H_{tot} and H_{noise} to words of infinite lengths. The mutual information $I = H_{tot} - H_{noise}$ at the intercept for 1/wordlength = 0. A similar method was used to evaluate the transfer entropy. In both cases, curves were generated using spike trains simulated using the Hodgkin-Huxley-type model.

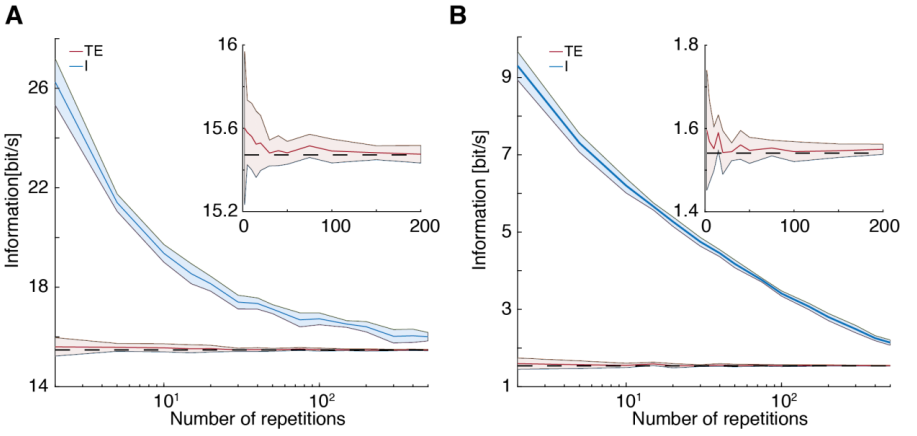

Figure 3. Comparative performance of *I* and *TE* in evaluating information flow at synapses driven by Poisson synthetic spike trains.

Comparison between mutual information (*I*; blue; calculated following [8, 11]) and transfer entropy (*TE*; red; calculated following [10, 17]) as a function of the size of the dataset for randomly generated spike trains based on thalamic relay cells characteristics (**A**) [8] or layer 4 spiny stellate cells characteristics (**B**) [10] (see also Table 1). In each case, the black line indicates the theoretical value (see Methods). In both A and B, the inset zooms on *TE* for a low number of repetitions. In each case, shaded areas indicate the standard error of the mean.

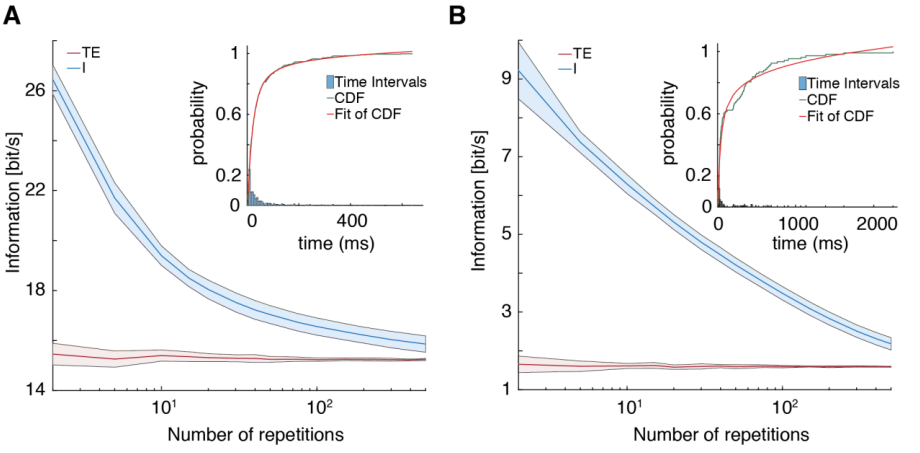

Figure 4. Comparative performance of *I* and *TE* in evaluating information flow at synapses driven by synthetic spike trains with realistic biological temporal structures.

Comparison between mutual information (*I*; blue; calculated following [8, 11]) and transfer entropy (*TE*; red; calculated following [10, 17]) as a function of the size of the dataset for randomly generated using the cumulative distribution function (CDF) of the interspike intervals (Insets) based on experimental data recorded impinging onto thalamic relay cells (**A**) [8] or layer 4 spiny stellate cells (**B**) [10]. In each case, the inset shows the CDF of the biological interspike interval used to generate input sequences. In each case, shaded areas indicate the standard error of the mean.

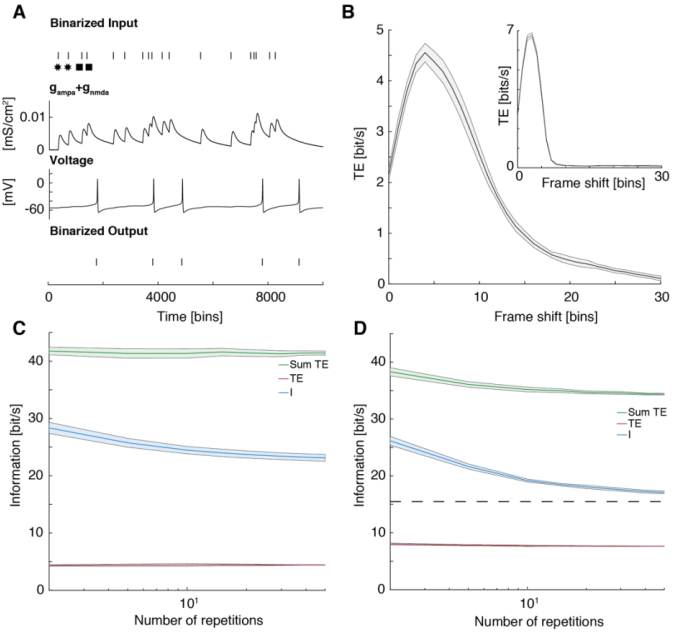

Figure 5. Application to a biophysical Hodgkin-Huxley-type model.

(**A**) Poisson spike trains and transmission probabilities are replaced by an experimentally-calibrated Hodgkin-Huxley-type model for thalamic relay cells that recapitulates transmission properties at thalamic relay synapses. 1 time bin = 3 ms. (**B**) Synaptic and neuronal dynamics can lead to output action potentials being generated in different time bins than the incoming input ('frame shift'). This leads to transfer entropy being positive over multiple temporal frame shifts between the binarized input and output sequences. Mutual information calculated following [11] is immune to that issue. (**C**) Because of this, direct comparison between mutual information (*I*; blue) and transfer entropy (*TE*; calculated using [17]; red: peak value from B; green: integral over *TE* in B) can be difficult. (**D**) This effect can be recapitulated in simple Poisson spike trains by adding gaussian temporal jitter to the timing at which an output spike is generated (see inset in B). *I* is unaffected when the temporal jitter is preserved over repetitions (compare with Figure 2A), while the results obtained for *TE* recapitulate what is observed with the Hodgkin-Huxley-type model (compare with C). In each case, shaded areas indicate the standard error of the mean.

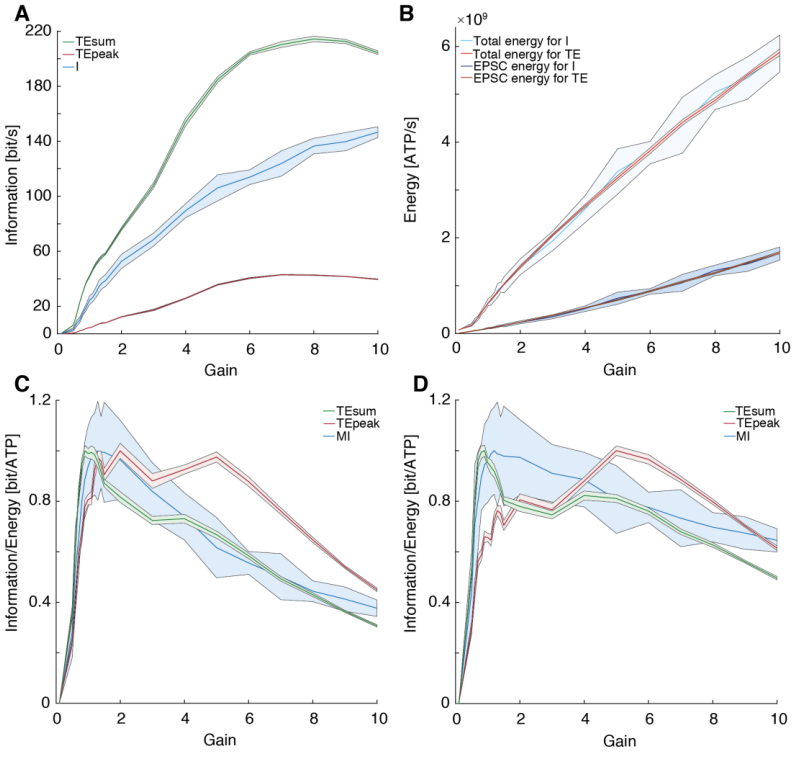

Figure 6. Information flow vs. energy.

(**A**) Information transfer at modelled thalamic relay synapses calculated using the Hodgkin-Huxley-type model as a function of the gain of the synapse (normalized synaptic conductance). (**B**) Total energy consumption and EPSC energy consumption used to convey the information calculated in A. (**C**) Normalized ratio between information (A) and EPSC energy (B) as a function of the normalized synaptic conductance. (**D**) Normalized ratio between information (A) and total energy consumption (B) as a function of the normalized synaptic conductance.

2.3 Paper 3

M Conrad, and R. B. Jolivet, "Modelling neuromodulated information flow and energetic consumption at thalamic relay synapses", Lectures Notes in Computer Science (to appear, 2020)

Summary

This paper presents an improvement of the Hodgkin-Huxley-type model developed in Paper 2 to take into account neuromodulation of transmission properties by serotonin at thalamic relay synapses. At these synapses, serotonin causes modulation of the release probability of glutamate-containing vesicles and thus of the amplitude of the postsynaptic conductance. This neuromodulation of the amplitude is simulated with a modified version of the Tsodyks-Markram model [121]. We show that this neuromodulation changes the voltage trajectory in the postsynaptic cell, and consequently the binary output sequence in response to the same input. This indicates that serotonin can change the encoding of visual information.

The paper also assesses the effect of neuromodulation by serotonin on the energetic efficiency of information transfer. For this, we use mutual information and calculate energy consumption like in Paper 2, as a function of the postsynaptic gain of the cell. We show that neuromodulation by serotonin reduces both the information conveyed and the energy consumed by the synapse to convey this information. The neuromodulated ratio between information and energy still has a maximum around the physiological gain, but the peak is slightly broadened compared to the case without serotonin. These results seems to indicate that energetic efficiency of information transfer is a generic design principle in the brain.

Contribution

I wrote most of the code necessary as well as runned all the simulations. I also analysed the data and was involved in writing the paper.

Modelling Neuromodulated Information Flow and Energetic Consumption at Thalamic Relay Synapses*

Mireille Conrad[1] and Renaud B. Jolivet[1][0000−−0002−5167−0851]

Department of Nuclear and Corpuscular Physics,
University of Geneva, Geneva, Switzerland
{mireille.conrad,renaud.jolivet}@unige.ch

Abstract. Recent experimental and theoretical work has shown that synapses in the visual pathway balance information flow with their energetic needs, maximising not the information flow from the retina to the primary visual cortex (bits per second), but instead maximising information flow per concomitant energy consumption (bits of information transferred per number of adenosine triphosphate molecules necessary to power the corresponding synaptic and neuronal activities) [10, 5, 11]. We have previously developed a biophysical Hodgkin-Huxley-type model for thalamic relay cells, calibrated on experimental data, and that recapitulates those experimental findings [10]. Here, we introduce an improved version of that model to include neuromodulation of thalamic relay synapses' transmission properties by serotonin. We show how significantly neuromodulation affects the output of thalamic relay cells, and discuss the implications of that mechanism in the context of energetically optimal information transfer at those synapses.

Keywords: Brain energetics · Information theory · Energetic optimality · Neuromodulation.

1 Introduction

The brain consumes an inordinate amount of energy with respect to its size. It is responsible for about 20% of the whole body baseline energy metabolism at rest, while representing usually only 2% of its mass [9]. Over the last couple of years, attempts at theoretically or experimentally determining an energetic budget for the brain have all pointed to synapses as the locus where most of brain energy is being spent [7], with estimates putting their share of the brain's signalling energy budget at roughly 60% [1, 13, 9]. A better understanding of brain energetics is essential because abnormal energy metabolism is an early

* This work was supported by grants from the Swiss National Science Foundation (31003A_170079), the European Commission (H2020 862882 IN-FET), and the Australian Research Council (DP180101494) to RBJ. MC is enrolled in the Lemanic Neuroscience Doctoral School.

hallmark of numerous pathologies of the central nervous system [9], and because neuroenergetics offers a lens through which one can easily address the complex heterocellular complexity of the brain [4, 12, 13].

Synapses are also the locus where electrophysiological 'information' is transmitted from neuron to neuron, and Shannon's information theory [18] has been used to great effectiveness in neuroscience, to measure information flow at synapses, in neural networks, or between different brain areas [15, 16, 9].

Given that synapses are a key mechanism in interneuronal communication and appear by all estimates to be responsible for a large fraction of the brain's energy consumption, it is natural to think that they would be reliable information transmission devices. A large body of experimental evidence, however, shows that synapses can be remarkably unreliable. For instance, the release probability for presynaptic vesicles is often measured to be in the few tens of percents for cortical neurons [8, 3]. Similarly, action potential transmission at thalamic relay synapses, which relay information from sensory modalities to the primary sensory cortices, can be astonishingly low (see references in ref. [10]).

Recently, we have shown that this apparent paradox can be resolved when considering the energetic optimality of information transmission. In other words, synapses and neurons maximise the energetic efficiency of information transfer, measured in bits of information transferred through a synaptic connection, or from the input to the output of a cell, per number of adenosine triphosphate molecules necessary to power this synaptic or neuronal activity [5]. In particular, we have shown that this trade-off between information flow and concomitant energetic consumption can explain low release probability at cortical synapses [9], and action potential transmission characteristics in the visual pathway, at thalamic relay synapses between retinal ganglion cells and thalamic relay neurons [10], as well as at the next synapse in that pathway, the synapse that thalamic relay cells form on layer 4 spiny stellate cells in the primary visual cortex [11].

In parallel to experiments, we have developed biophysical Hodgkin-Huxley-type models of these systems, all carefully calibrated on experimental data (see refs. [10, 11] for further details). Here, we resume our study of these questions at thalamic relay synapses to investigate neuromodulation of these synapses' transmission properties by serotonin. The next section introduces the Hodgkin-Huxley-type model and the newly experimentally-calibrated model of neuromodulated synaptic input.

2 Mathematical model of information transmission and concomitant energy consumption in thalamic relay cells

2.1 Hodgkin-Huxley formalism

We have previously published an experimentally calibrated biophysical single-compartment model of the Hodgkin-Huxley-type for thalamic relay cells [10]. Briefly, the model is written as follows: The Hodgkin-Huxley formalism describes

the dynamics of the membrane voltage V as:

$$C_m \frac{dV}{dt} = -\sum_j i_j - i_{Hold} - i_{syn}, \tag{1}$$

with C_m the membrane capacitance, i_j the intrinsic currents, i_{Hold} an experimentally injected holding current (see ref. [10] for further details) and i_{syn} the synaptic currents. Following Bazhenov and colleagues [2], the intrinsic currents include a leak current i_L, a potassium leak current i_{KL}, an A-type potassium current i_A, a T-type low threshold calcium current i_T, an h-current i_h, a fast sodium current i_{Na} and a fast potassium current i_K. All the intrinsic currents have the same general form:

$$i = g m^M h^N (V - E), \tag{2}$$

where for each current i, g is the maximal conductance, $m(t)$ is the activation variable, $h(t)$ is the inactivation variable, E is the reversal potential, and M and N are the number of independent activation and inactivation gates.

The intracellular calcium dynamics is defined by:

$$\frac{d\mathrm{Ca}_i^{2+}}{dt} = -\frac{1}{\tau_{\mathrm{Ca}}}(\mathrm{Ca}_i^{2+} - \mathrm{Ca}_{i,0}^{2+}) - A\, i_T, \tag{3}$$

with $\mathrm{Ca}_{i,0}^{2+} = 2.4\,10^{-4}\,\mathrm{mM}$, the baseline intracellular calcium concentration, and $A = 5.18\,10^{-5}\,\mathrm{mM\,cm^2\ ms^{-1}\,\mu A^{-1}}$, a constant. The time dependence for m and h is defined by:

$$\frac{dx}{dt} = \alpha_x(1 - x) - \beta_x x, \tag{4}$$

where x stands for either h or m.

We refer the interested reader to ref. [10] for all further details of the model, and for a detailed description of experimental and calibration procedures.

2.2 Modelling synaptic input, synaptic depression and neuromodulation

In order to model the strong paired-pulse depression and neuromodulation that is known to happen at thalamic relay synapses, we use the formalism introduced by Tsodyks and Markram [20].

First, a sequence of binarized input action potentials is generated with a temporal resolution $\Delta t = 3\,\mathrm{ms}$. In order to generate sequences with the same temporal statistics than recorded *in vivo*, we use sequences recorded *in vivo* (available from ref. [10]) to calculate the non-Poissonian *in vivo* inter-spike interval distribution. From this distribution, we calculate the cumulative distribution function of inter-spike intervals, and in turn, use that cumulative distribution function to generate synthetic binary sequences.

Each input is then used to trigger an AMPA and a NMDA conductance with the generic form:

$$g(\delta t) = A\ (\exp(-\delta t/\tau_{\text{rise}}) - \exp(-\delta t/\tau_{\text{decay}}))\,, \tag{5}$$

with τ_{rise} and τ_{decay} some time constants, δt the time elapsed since the input action potential and A an amplitude. Consecutive contributions are summed up.

In each case, the effective amplitude A of the triggered conductance is modulated by synaptic depression. To model this, we use a slight adaptation of the Tsodyks-Markram model [20], whereby the amplitude of the conductance is given by:

$$A_{n+1} = A_n\,(1-U)\exp(-\Delta t/\tau_{\text{rec}}) + A\,U\,(1-\exp(-\Delta t/\tau_{\text{rec}})), \tag{6}$$

with U and τ_{rec} some parameters. Fitting that model on experimental data from ref. [10] yields $U = 0.7$ and $\tau_{\text{rec}} = 620\,\text{ms}$ (to be described in details somewhere else). These parameters predict a paired-pulse depression of ~ 0.4 for consecutive pulses at 100 ms interval, in excellent agreement with what was observed in electrophysiological recordings [10]. Additionally, that procedure yields $\tau_{\text{rise}} = 0.75\,\text{ms}$ and $\tau_{\text{decay}} = 2\,\text{ms}$ for the AMPA conductance, and $\tau_{\text{rise}} = 9\,\text{ms}$ and $\tau_{\text{decay}} = 22\,\text{ms}$ for the NMDA conductance, and a ratio between the peak amplitudes of the NMDA and AMPA conductances of 0.1.

Thus, i_{syn} (see Eq. 1) is given by:

$$\begin{aligned} i_{syn} = &-g_{\text{AMPA}}(V - E_{\text{excitatory}}) \\ &- g_{\text{NMDA}}\left(\frac{9.69}{1+0.1688\,e^{-0.0717V}}\right)(V - E_{\text{excitatory}}), \end{aligned} \tag{7}$$

with $E_{\text{excitatory}} = 0\,\text{mV}$, the reversal potential of AMPA and NMDA receptors. The additional term in the description of the NMDA conductance is added to describe the nonlinear I-V relation of NMDA receptors due to the Mg^{2+} block [10]. In each case, g_{AMPA} and g_{NMDA} are determined by the procedure mentioned above combining Equations [5] and [6]. An example of what that procedure yields can be observed in the top two panels of Fig. 1 below.

Activation of serotonin receptors at thalamic relay synapses modulates the release probability of presynaptic vesicles. Specifically, the release probability is reduced and while this tends to lead to smaller postsynaptic potentials (PSPs), it makes consecutive PSPs more similar to each other in amplitude. The literature and preliminary experimental data (courtesy of D. Attwell, E. Engl and J.J. Harris) show that the presence of serotonin receptor agonists experimentally lead to reduced paired-pulse depression with the ratio of consecutive PSPs at 100 ms interval to be about ~ 0.8 (instead of ~ 0.4 in control conditions). This can be easily achieved in the model presented here by changing the value of the parameter U to 0.2, yielding an elegant and simple framework to study the effect of neuromodulation at thalamic relay synapses. An example of what that procedure yields can be observed in Fig. 2 below and can be directly compared with the results in Fig. 1.

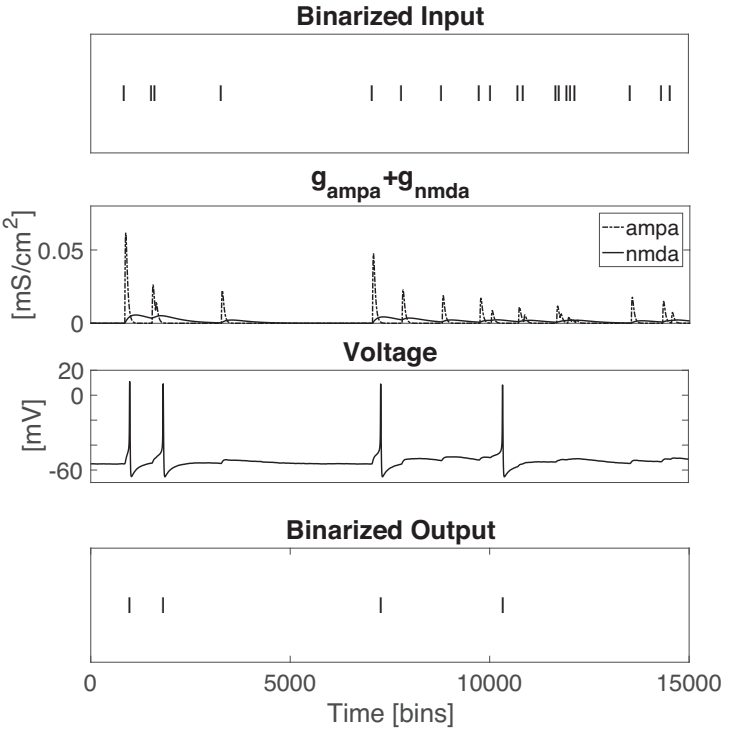

Fig. 1. Model dynamics in absence of neuromodulation ($U = 0.7$). **Top row:** Binarized input sequence at time resolution $\Delta t = 3$ ms. Input action potentials are generated at approximately 20 Hz and their temporal dynamics follows experimental data recorded *in vivo* in rodents, i.e. their inter-spike interval distribution matches the inter-spike interval distribution observed *in vivo*. **Second row:** The dynamics of the AMPA and NMDA conductances with parameter values for amplitudes, time constants and synaptic depression derived from experimental recordings and following the Tsodyks-Markram model [20]. With $U = 0.7$, synaptic conductances display significant depression. A paired-pulse depression of ~ 0.4 is predicted in these circumstances (consecutive pulses at 100 ms interval), matching what was observed in electrophysiological recordings [10]. **Third row:** Dynamics of the membrane voltage of the thalamic relay cell predicted in response to the input sequence. **Bottom row:** Binarized output sequence at time resolution $\Delta t = 3$ ms. The thalamic relay cell model only produces 4 output action potentials in response to the top input sequence (the average output frequency is ~ 4 Hz). In general, those are synchronous with the first input following a relatively long silent period.

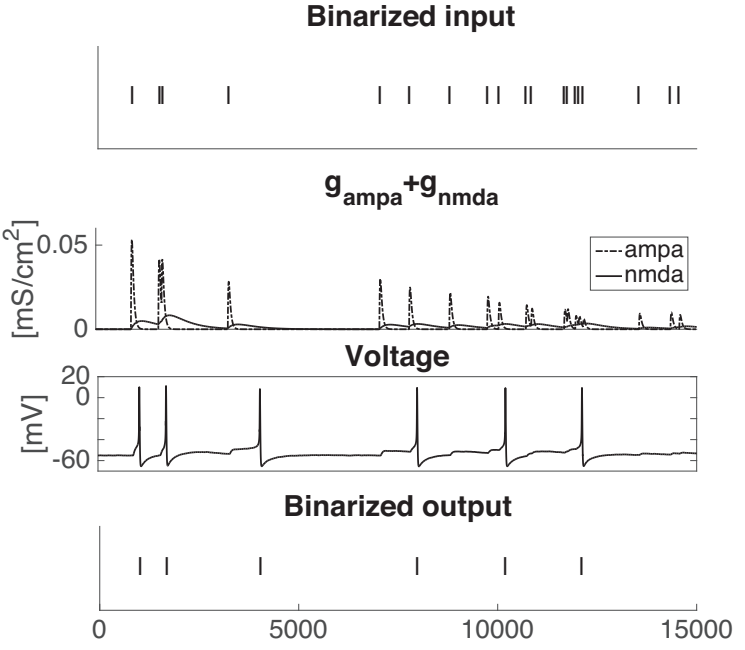

Fig. 2. Model dynamics with strong neuromodulation by serotonin ($U = 0.2$). **Top row:** Binarized input sequence at time resolution $\Delta t = 3$ ms. Input action potentials are generated at approximately 20 Hz and their temporal dynamics follows experimental data recorded *in vivo* in rodents. This is the same sequence as in Fig. 1. **Second row:** The dynamics of the AMPA and NMDA conductances with parameter values for amplitudes, time constants and synaptic depression derived from experimental recordings and following the Tsodyks-Markram model [20]. With $U = 0.2$, synaptic conductances display much less depression than in Fig. 1. A paired-pulse depression of ~ 0.8 is predicted in these circumstances (consecutive pulses at 100 ms interval), matching what was observed in preliminary electrophysiological recordings. **Third row:** Dynamics of the membrane voltage of the thalamic relay cell predicted in response to the input sequence. **Bottom row:** Binarized output sequence at time resolution $\Delta t = 3$ ms. The thalamic relay cell model now produces 6 output action potentials in response to the top input sequence. That output sequence is significantly different than the one at the bottom of Fig. 1.

2.3 Information flow and neuroenergetics

In order to assess information flow at the modelled feed-forward synapse, we collect the binarized input and output sequences with a temporal resolution of $\Delta t = 3\,\text{ms}$ (see Fig. 1 top and bottom panels). We then apply the so-called direct method by Strong *et al.* [19] to compute the *mutual information* between those input and output sequences, similar to what has been done in [16] and [10]. A detailed description of how to use this method and others for the analysis of spike trains can be found in ref. [15]. Note also that it is possible to use the so-called *transfer entropy* to measure information flow between neurons [17], instead of the mutual information as we do here. We refer the reader to refs. [11, 6] for a comparative discussion of these measures in a context similar to the one discussed here.

The energy consumption in thalamic relay cells in this scenario arises from presynaptic activity and from the generation of output action potentials. Transport of ions across membranes during neural activity leads to the activation of the Na,K-ATPase electrogenic pump, which consumes adenosine triphosphate (ATP) molecules to maintain and reestablish normal ionic gradients [9, 1, 13]. It is thus possible to compute the energetic cost of neuronal activity (number of ATP molecules consumed in response to that activity) using biophysics as described in [1, 10, 11].

3 Modulation of transmission properties by the neuromodulator serotonin

Figure 1 shows typical data generated by the model in the scenario corresponding to the control experimental situation described in ref. [10], i.e. with strong paired-pulse depression ($U = 0.7$). Strong depression is apparent in the second panel from the top, where each input action potential following the first action potential after a long period of silence only evokes a much reduced conductance. As a result, the model, like the cells it is based on, tends to generate outputs only when two input action potentials come in close succession to each other (this is not always the case, however, as a single input action potential can be seen to trigger an output spike at time bin 7000). While the model receives input action potentials at a frequency of $\sim 20\,\text{Hz}$, it generates output action potentials at only $\sim 4\,\text{Hz}$.

Figure 2 shows typical data generated by the model in the scenario corresponding to application of serotonin, i.e. with weak paired-pulse depression ($U = 0.2$). The binary input sequence is the same as the one used in Fig. 1. Weak depression is apparent in the second panel from the top, where each input action potential triggers in average smaller conductances, but with amplitudes more evenly distributed over time. A comparison between Figs. 1 and 2 reveals that even though the model is driven in both cases by the same binary input sequence, the voltage trajectory of the cell is very significantly affected by changing the value of U, and the binary output sequence of the cell is now largely different.

These results suggest that neuromodulation at these synapses could have a significant effect on the quantity and type of information that reaches the primary visual cortex.

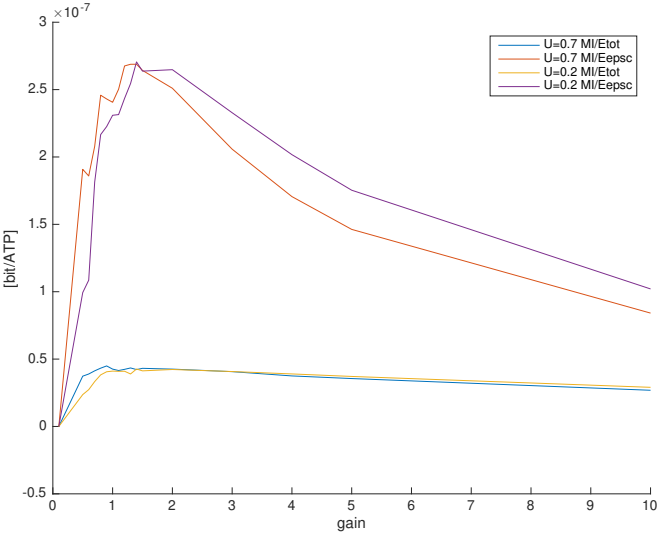

Fig. 3. Energetic optimality of information transfer at thalamic relay synapses. Mutual information (bits/sec) divided by concomitant energetic costs (ATP/sec), factoring the cost of reverting ionic flows across the cellular membrane resulting from the activation of postsynaptic receptors (top curves), or resulting from the activation of postsynaptic receptors and from the generation of action potentials (bottom curves). The parameter U is chosen to be either $U = 0.7$ to match control experimental conditions or $U = 0.2$ to model the application of serotonin receptor agonists. An overall gain factor is applied to the synapse with gain = 1 matching the experimental physiological gain in control conditions. The curves reveal the presence of an optimum at, or slightly above, gain = 1 [10]. Shifting from $U = 0.7$ to $U = 0.2$ appears to shift the peak of each curve slightly to the right, and to slightly broaden the peak.

We then tested whether this type of neuromodulation affects the energetic optimality of information transmission at these synapses. To this end, we ran simulations varying the overall gain of the thalamic relay synapse and measured the mutual information between the input and output sequences [10, 11] using the direct method [19]. We additionally computed the equivalent energetic budget using standard biophysical methods developed in ref. [1]. We observed that, while changing the value of U does affect absolute values, the energetic

consumption (measured in ATP/sec) associated with the generation of postsynaptic potentials, or with the generation of postsynaptic potentials *and* action potentials, scales more or less linearly with the gain of those synapses, whatever the value of U (not shown). This matches what has been observed elsewhere [9–11]. We also observed that, while changing the value of U does affect absolute values, information flow across the relay synapse (measured in bits/sec) scales sigmoidally with the gain of those synapses (not shown). Again, this matches what has been observed elsewhere [9–11].

We then computed in each scenario the ratio of information flowing through the synapse to the concomitant energy consumption necessary to power the synaptic and neuronal activity of the thalamic relay neuron. Preliminary results displayed in Figure 3 show that this results in curves with a relatively well-defined energetic optimum for information transmission, whether $U = 0.7$ or 0.2, and whether the energy budget includes the cost of postsynaptic potentials alone (top curves), or also includes the cost of postsynaptic potentials and action potentials (bottom curves). In all cases, the energetic optimum stood at, or close to, gain = 1, the physiological gain of the synapse in control conditions. This is in excellent accordance with experimental findings and a previous version of this model (driven by experimentally-recorded conductances). Despite its significant effect on the actual output sequences generated by the thalamic relay neuron, neuromodulation (shifting from $U = 0.7$ to $U = 0.2$) only appears to shift the peak of each curve slightly to the right, and to slightly broaden said peak.

4 Discussion

Here, we have introduced a carefully-calibrated *mechanistic* [14] model of synaptic depression and neuromodulation by serotonin at thalamic relay synapses. We have described how to build and calibrate such a model using experimental data, and together with the work in ref. [10], we have established that it qualitatively, and to some extent quantitatively, captures the behaviour of biological thalamic relay neurons, in particular here, with respect to modelling *in vivo*-like synaptic inputs, including their modulation by serotonin.

The results presented here suggest that neuromodulation by serotonin does not very significantly affect the energetic optimality of information transmission at thalamic relay synapses. The fact that neuromodulation does not very strongly affect the position of the peak in the information over energy curves, and the fact that this peak sits at the experimentally observed physiological gain for those synapses (gain = 1), reinforces the notion that this principle might be a relatively generic design principle in the brain. This model thus also contains *normative* (energetic) aspects [14]. It is our contention that synaptic activity has evolved under, and is to an extent shaped by, energetic constraints [9–11, 5].

Neuromodulation does, however, very significantly affect what output sequences are sent out to the primary visual cortex in response to a given input sequence. In other words, it appears to change the encoding of visual information. Our model thus opens now the possibility to systematically investigate what

kind of input sequences will maximise under different circumstances information flowing from the retina to the primary visual cortex.

References

1. Attwell, D., Laughlin, S.B.: An energy budget for signaling in the grey matter of the brain. Journal of Cerebral Blood Flow & Metabolism **21**(10), 1133–1145 (2001)
2. Bazhenov, M., Timofeev, I., Steriade, M., Sejnowski, T.J.: Cellular and network models for intrathalamic augmenting responses during 10-hz stimulation. Journal of Neurophysiology **79**(5), 2730–2748 (1998)
3. Branco, T., Staras, K., Darcy, K.J., Goda, Y.: Local dendritic activity sets release probability at hippocampal synapses. Neuron **59**(3), 475–485 (2008)
4. Coggan, J.S., Calì, C., Keller, D., Agus, M., Boges, D., Abdellah, M., Kare, K., Lehväslaiho, H., Eilemann, S., Jolivet, R.B., et al.: A process for digitizing and simulating biologically realistic oligocellular networks demonstrated for the neuro-glio-vascular ensemble. Frontiers in Neuroscience **12**, 664 (2018)
5. Conrad, M., Engl, E., Jolivet, R.: Energy use constrains brain information processing. In: Technical Digest - International Electron Devices Meeting. pp. 11–3 (2018)
6. Conrad, M., Jolivet, R.B.: Comparative performance of mutual information and transfer entropy for analysing the balance of information flow and energy consumption at synapses. Submitted (2020)
7. Engl, E., Jolivet, R., Hall, C.N., Attwell, D.: Non-signalling energy use in the developing rat brain. Journal of Cerebral Blood Flow & Metabolism **37**(3), 951–966 (2017)
8. Hardingham, N.R., Read, J.C., Trevelyan, A.J., Nelson, J.C., Jack, J.J.B., Bannister, N.J.: Quantal analysis reveals a functional correlation between presynaptic and postsynaptic efficacy in excitatory connections from rat neocortex. Journal of Neuroscience **30**(4), 1441–1451 (2010)
9. Harris, J.J., Jolivet, R., Attwell, D.: Synaptic energy use and supply. Neuron **75**(5), 762–777 (2012)
10. Harris, J.J., Jolivet, R., Engl, E., Attwell, D.: Energy-efficient information transfer by visual pathway synapses. Current Biology **25**(24), 3151–3160 (2015)
11. Harris, J.J., Engl, E., Attwell, D., Jolivet, R.B.: Energy-efficient information transfer at thalamocortical synapses. PLOS Computational Biology **15**(8), e1007226 (2019)
12. Jolivet, R., Coggan, J.S., Allaman, I., Magistretti, P.J.: Multi-timescale modeling of activity-dependent metabolic coupling in the neuron-glia-vasculature ensemble. PLoS Computational Biology **11**(2) (2015)
13. Jolivet, R., Magistretti, P.J., Weber, B.: Deciphering neuron-glia compartmentalization in cortical energy metabolism. Frontiers in Neuroenergetics **1** (2009)
14. Levenstein, D., Alvarez, V.A., Amarasingham, A., Azab, H., Gerkin, R.C., Hasenstaub, A., Iyer, R., Jolivet, R.B., Marzen, S., Monaco, J.D., Prinz, A.A., Quraishi, S., Santamaria, F., Shivkumar, S., Singh, M.F., Stockton, D.B., Traub, R., Rotsteim, H.G., Nadim, F., Redish, A.D.: On the role of theory and modeling in neuroscience (2020)
15. Panzeri, S., Senatore, R., Montemurro, M.A., Petersen, R.S.: Correcting for the sampling bias problem in spike train information measures. Journal of Neurophysiology **98**(3), 1064–1072 (2007)

16. Reinagel, P., Reid, R.C.: Temporal coding of visual information in the thalamus. Journal of Neuroscience **20**(14), 5392–5400 (2000)
17. Schreiber, T.: Measuring information transfer. Physical Review Letters **85**(2), 461 (2000)
18. Shannon, C.E.: A mathematical theory of communication. Bell System Technical Journal **27**(3), 379–423 (1948)
19. Strong, S.P., Koberle, R., Van Steveninck, R.R.D.R., Bialek, W.: Entropy and information in neural spike trains. Physical Review Letters **80**(1), 197 (1998)
20. Tsodyks, M.V., Markram, H.: The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. Proceedings of the National Academy of Sciences **94**(2), 719–723 (1997)

CHAPTER 3

Discussion

The brain is a very sophisticated structure, as well as a very powerful computational device. And like every computational device, the brain needs energy to be able to function correctly. This quite small organ is one of the main consumers of energy in the human body, as it consumes by itself around 20% of the total amount of energy needed by the body. But, the energy the body can allocate to the brain is not infinite and this has probably shaped the development of brain function. Studies focused on energetic budgets in the brain have shown that the majority of the brain's energy is consumed by synapses, the connections between neurons. Strikingly, even when energy supplies are normal, synapses are mostly unreliable connections. For example, in the CNS, the probability that an action potential leads to the release of a vesicle of neurotransmitters is estimated to be around 25 to 50 %. Action potentials are the most important signals neurons exchange, they are the vectors of information transiting along those neurons. Such a low probability of transmission means that neurons do not maximize the information they could convey or transmit. A number of recent studies suggest that instead, they maximize the ratio between information and energy used to convey this information: the energetic efficiency of information transfer.

This PhD thesis includes 3 papers focused on better understanding energetic efficiency of information transfer at synapses from a computational point of view. The papers themselves as well as a short summary of each can be found in Chapter 2. Paper 1 reviews existing findings about energetic efficiency. In order to study energetic efficiency of information transfer, it is important to be able to correctly evaluate the information flow between

circuit components. Paper 2 focuses on the comparison of two metrics of information theory, mutual information and transfer entropy. Those two metrics are then used to try to reproduce previous experimental findings of energetic efficiency in the visual pathway. Information flow and energy consumption are not stable quantities through time, as many factors can influence the state of neurons. Paper 3 thus concentrates on the modelling of synaptic depression and, specifically, on the effect of serotonin on energetic efficiency of information transfer at LGN synapses.

The following sections of this Chapter focus on the three research questions highlighted in Chapter 1: how to estimate the information at the synaptic level? Can we reproduce experimental findings about energetic efficiency in the LGN computationally? And what is the effect of neuromodulation by serotonin on energetic efficiency of information transfer?

3.1 Information measurements

Information theory is a useful tool when studying how information is managed by neurons and neuronal networks. But, using information theory is not necessarily straightforward, especially in a Neuroscience context. Here, the performances of two metrics of information theory (mutual information I and transfer entropy TE) were compared on binary spike trains when used to measure the information flow between neurons. In most experimental settings, the length of the data that can be collected is an issue, as animals might have limited focus during an experiment, or cells die after a few hours *in vitro*. In Paper 2, we chose to compare I and TE as a function of the length of the dataset to see exactly how the amount of information available influences those two metrics. In both cases, we used a correction method.

We generated input-output pairs in three different scenarios, progressively adding complexity:

1. At first, we generated inputs as spike trains with a Poisson statistic and generated the outputs by copying each bin from the corresponding input with some probability of making an error (failure of transmission) and some probability of spontaneous firing. These two probabilities were chosen to reproduce probabilities found in thalamic relay cells and layer 4 spiny stellate cells.

2. Obviously, the input thalamic relay cells and layer 4 spiny stellate cells receive do not have a Poisson statistic. So, in order to be more realistic, we generated inputs with statistics based on biological interspike intervals. The output was generated as in scenario 1.

3. Finally, we studied a more complex scenario by using a Hodgkin-Huxley-type model calibrated on thalamic relay cells to generate the

output. The input was generated as in scenario 2.

3.1.1 Bias and correction methods

Both I and TE are calculated as the subtraction of two entropies (see Equations 1.6 and 1.10). As explained in Chapter 1, the evaluation of those entropies can lead to biases, especially with limited datasets. In this thesis, we chose to use the direct method [91] in the evaluation of the mutual information. The direct method includes two correction, one that extrapolates the value of the entropies for a dataset of infinite size, and one that extrapolates the value of the entropies for infinite word lengths (see Section 1.2.2).

As stated in Paper 2, we were not able to obtain convincing results for the first correction. Instead of a monotonically increasing curve like the one plotted in the original paper by Strong and colleagues (Figure 2 of reference [91], or Figure 1.9 of Chapter 1), ours was essentially flat, for both the total entropy (see Equation 1.1) and the conditional (or noise) entropy (see Equation 1.3). Figure 3.1 shows an example for the total entropy H_{tot} in scenario 1, with words of length 7. Interestingly, we can observe in the work of Strong that this correction only influences the third digit after the decimal point of the estimate of the total entropy, suggesting that this correction does not have a lot of impact. It is nonetheless important to note that these results are for a specific visual system in the fly, and for several hours of recording.

On the other hand, we obtained convincing results with the second correction method. Thus, in order to be consistent, we chose to apply a similar method for the correction of the bias due to long words to the transfer entropy. There is no explicit "noise entropy" in the definition of the transfer entropy, like there is for the mutual information. Nonetheless, it is important to verify that random correlations in the dataset are not affecting the results. This can be an issue especially when considering long words. We thus evaluated a noise for the transfer entropy TE_{noise} by randomly shuffling words in the output and calculating the transfer entropy between the input and the shuffled output. Randomly shuffling the words allows to create a spike train with the same frequency but with no other temporal structure, allowing to calculate the intrinsic noise and thus the contribution of random correlations. The estimate for the transfer entropy TE is then calculated as the raw transfer entropy TE_{raw} (the transfer entropy of the dataset) minus the noise transfer entropy: $TE = TE_{raw} - TE_{noise}$. The value obtained for the estimate of the transfer entropy TE is then extrapolated to words of an infinite length, in a similar fashion as the correction used in the direct method (see Figure 2B of Paper 2). For computational purposes, we chose to first subtract the noise before extrapolating the result to words of an infinite

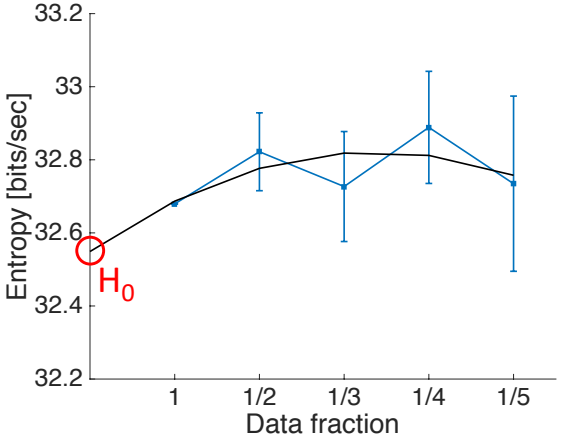

Figure 3.1: Application of the first correction of the direct method of Strong [91], for words of length of 7 bits with input and output generated as in scenario 1 (Poisson statistics). Total entropy is evaluated over different fractions of the whole dataset. To extrapolate the values of the entropy for infinite data, the entropy is fitted with a function of the form $H_{tot} = H_0 + \frac{H_1}{size} + \frac{H_2}{size^2}$, with $size$ denoting the inverse data fraction. The intercept H_0 is thus the value for an infinite dataset.

length, but extrapolating first TE_{raw} and TE_{noise} and then subtracting would yield the exact same outcome.

It is nonetheless important to note the differences between the method we developed for the transfer entropy and the direct method applied to mutual information, even though both aim to correct the same bias. Mainly, the direct method is applied to the individual total and conditional entropies, while our method is applied directly on the transfer entropy. This choice was made for the sake of simplicity. However, investigating a correction method for the transfer entropy directly based on the entropies would be an interesting study.

Correctly locating the length of the longest word before the sampling disaster is determinant to correctly use those two correction methods. The examples shown in Chapter 1 [92], as well as in the original paper from Strong *et al.* [91], show a behaviour essentially linear before the sampling disaster, which makes it relatively easy to determine an inflection point by visual inspection of the data. This was not always the case in our work, especially for the transfer entropy (see Figure 2B of Paper 2). In this case, the choice of the part of the curve considered to be part of the sampling disaster had to be made with caution, as it can obviously have an effect on the results. This is something to take into account when designing an

experiment using this correction method as well, as this could depend on the behaviour of the neurons studied.

3.1.2 Mutual information and transfer entropy comparison

Scenarios 1 and 2 show that transfer entropy might be a better choice to evaluate information flow than mutual information, especially for small datasets, as it converges to its asymptotic value faster. In scenario 1, all the probabilities of the system are known and can be used in Equations 1.6 and 1.10 to calculate the theoretical values that the estimation method should reach. Moreover, in this scenario we have shown that the theoretical value is the same for both metrics, a useful property to compare their efficiency. Both metrics overestimate the theoretical value for a very small dataset before converging to the theoretical value as the size of the dataset increases. Interestingly, TE overestimate the theoretical value far less than I, even for short recordings (as can be seen in the insets of Figure 3 of Paper 2). Even though we cannot compute an exact theoretical value, we observe the same behaviour in scenario 2.

Even though we do not have a definitive answer as to why TE is less biased than I in this scenario, we advance a tentative explanation. When applying the corrections discussed above, we estimate the probabilities with words of length L. In the case of the mutual information, as explained in Chapter 1, the evaluation of $H(\mathbf{X})$ necessitates the estimation of 2^L probabilities and the evaluation of $H(\mathbf{X}|\mathbf{Y})$ necessitates the estimation of 2^{2L} probabilities. We use here bold notation for the variables $\mathbf{X}$ and $\mathbf{Y}$ to denote that they represent words of length L.

In the case of TE, it is important to note that Y^+ (see Equation 1.10) represents one bin, and not a word of length L. Otherwise, it would break causality, meaning that we are receiving information from the future. In this case, the evaluation of $H(\mathbf{X}^-|\mathbf{Y}^-)$ necessitates 2^{2L} probabilities to be estimated, and $H(\mathbf{X}^-|Y^+, \mathbf{Y})$ necessitates 2^{2L+1} probabilities (and not 2^{3L}!). We thus suspect that the explanation as to why TE converges faster is that the downward biases made when estimating $H(\mathbf{X}^-|\mathbf{Y}^-)$ and $H(\mathbf{X}^-|Y^+, \mathbf{Y})$ are more similar in amplitude than the ones made when evaluating $H(\mathbf{X})$ and $H(\mathbf{X}|\mathbf{Y})$. Because TE is calculated as $H(\mathbf{X}^-|\mathbf{Y}^-) - H(\mathbf{X}^-|Y^+, \mathbf{Y})$, biases similar in amplitude would thus give a result closer to the asymptotic value, reducing the overall upward bias for TE.

In both scenarios 1 and 2, TE and I converge to the same value. This is not a surprise, as in both scenarios, $\mathbf{X}^-$ and Y^+ are independent of $\mathbf{Y}^-$. Those two first scenarios were based on quite simple input statistics, not necessarily representative of realistic neuronal input statistics. Scenario 3 was helpful to show the limitations of this scenario.

In this scenario, the typical behaviour of the curve information vs. size of the dataset is still the same for both I and TE. TE converges faster to an asymptotic value with respect to the size of the dataset used for the evaluation and the amplitude of the overestimation is lower that what is observed for I. However, the two metrics do not converge to the same value. I still converges to a value close to the ones obtained in scenarios 1 and 2, but not TE.

Paper 2 showed that the Hodgkin-Huxley-type model transmission properties affect the evaluation of transfer entropy. In scenarios 1 and 2, spikes are transmitted from the input to the output in the matching time bin, with no transmission delay. This leads to the information carried by one given spike to be transmitted in one bin as well. In scenario 3, our Hodgkin-Huxley-type model adds systematic frame shifts between input and output. We show that this is due to integration of synaptic currents by the postsynaptic neuronal membrane. Intracellular dynamics and the time course of integration of presynaptic inputs can lead output action potential to be generated with a significant delay after an input action potential has impinged on the synapse. This means that the information carried by one spike can, in this scenario, be distributed over several consecutive time bins. This impacts the evaluation of the transfer entropy, because it also distributes it over several consecutive time bins (as shown in Figure 5 of Paper 2). Figure 3.2 shows an example (not included in Paper 2) of the transfer entropy evaluated in scenario 1. It can be seen that all the information is transmitted in the time bin corresponding to the timing of the spike in the input (frame shift = 0).

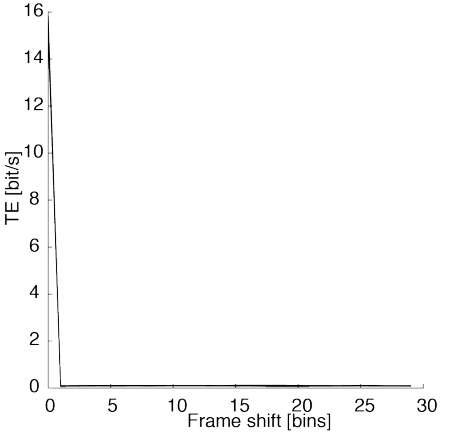

Figure 3.2: Transfer entropy plotted as a function of the frame shift between input and output (*i.e.* when shifting the input and the output sequences by a set number of time bins, 0 indicating that the bin considered in the output is in the same position as the spike in the input) for scenario 1 (Poisson statistics) for words of length 5 bins.

It is important to note here that using the direct method to calculate the mutual information makes it immune to the issue posed by frame shifts. $H(X)$ is only calculated on the output and thus cannot be impacted by the frame shift. $H(X|Y)$ is evaluated by comparing repetitions. If the frame shifts are the same across repetitions, then $H(X|Y)$ will not be affected either. We could thus expect mutual information (evaluated using the direct method) to be a better evaluation of the information conveyed by the neuron than transfer entropy in this scenario. But this is the case essentially because we used the direct method. Other methods to evaluate the mutual information exists, and they would not necessarily be immune to temporal frame shifts (an example being the method described in [123]).

The fact the the information is evaluated on multiple repetitions is one of the reasons why the direct method is so powerful. As shown by Mainen and Sejnowski [124], presenting the same strongly fluctuating input to a neuron several times will produce outputs that are very similar (as shown in Figure 3.3A). This can also be observed in the paper from Strong and colleagues about the direct method [91] (Figure 3.3B). We can see in panels A and B of Figure 3.3 that in both those studies, most spikes are present in the output spike train at the same timings, with only a few discrepancies across repetitions. This is also what we observe with our model, as can be observed in Figure 3.3C, where the few differences across repetitions are highlighted in green. This model, like a neuron, can add a delay between input and output action potentials. This delay is mostly the same across repetitions, thus not being an issue when using the direct method.

Paper 2 introduces two measures of the transfer entropy, called TE_{peak} and TE_{sum}. The value these two estimates of transfer entropy and the estimate of mutual information converge to are not the same in this last scenario. Apart from the issue caused by delayed transmission discussed above, it is also important to note that the approximation that $\mathbf{X}^-$ and Y^+ are independent of $\mathbf{Y}^-$ does not necessarily hold anymore, because this Hodgkin-Huxley-type model generates excitatory postsynaptic potential that can last more than the duration of one time bins (3 ms). This means that the integration of the presynaptic signal has a time constant that can be greater than 3 ms, thus affecting more than just one bin in the output. This implies that transfer entropy and mutual information do not estimate the same theoretical quantity anymore. Actually, this case of dependency between $\mathbf{X}^-$, Y^+ and $\mathbf{Y}^-$ also means that transfer entropy should be greater than in the independent case (more information is contained in a dependent state). This is what we observe with TE_{sum}. This is also a very important point as, in neurons, this assumption of independence is usually not correct. In this sense, TE_{sum} seems to be a better estimation than TE_{peak}, because it is a sum of all the information.

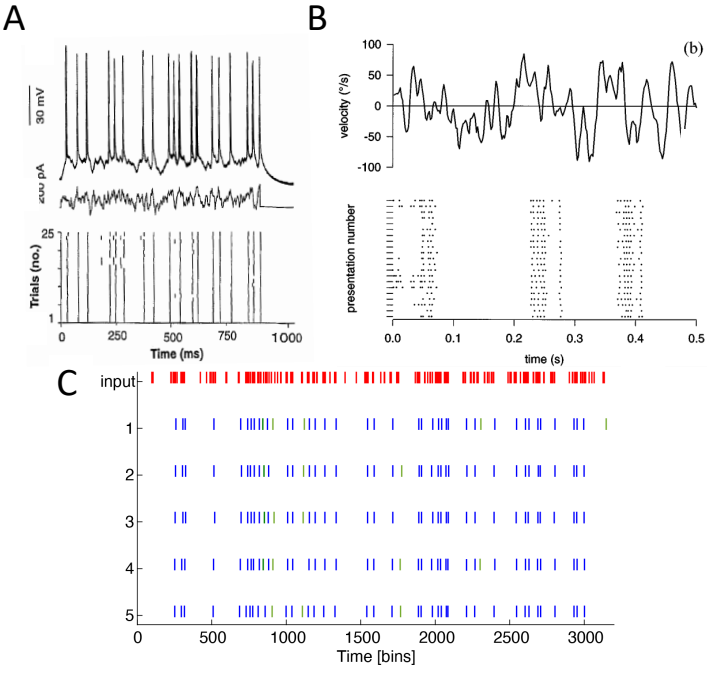

Figure 3.3: Examples of the reliability of neurons. In each cases, the response of the neurons over several trials are very similar to each other. A: A fluctuating current (middle) presented during 900 ms evokes trains of action potentials in a regular-firing layer-5 neuron. The top panel shows 10 superimposed responses. The bottom panel shows a raster plot of response spike times of 25 consecutive trials. Adapted from reference [124]. B: The input is a pattern moving across a fly's visual field. The top panel shows the angular velocity of that pattern. The bottom panel shows the repetition of responses spikes recorded in the visual cortex (H1). Adapted from reference [91]. C: The same input (upper row, red) is injected 5 times in our Hodgkin-Huxley-type model to generate the outputs in blue. The few spikes that are different across repetitions are highlighted in green.

When naively looking at the results displayed in Figures 3 and 4 of Paper 2, one could think that using transfer entropy is better than using mutual information, as it needs less data to converge to its asymptotic value. Nonetheless, we showed in Figure 5 of Paper 2 that in more complex (an biologically accurate) scenarios, where frame shifts could be expected, transfer entropy becomes more difficult to use accurately. Again, this does not mean that transfer entropy is wrong, but rather that mutual information and transfer entropy do not measure exactly the same quantities. A choice of metric should thus be made with care according to the experimental system investigated. However the method showed in Figures 3 and 4 of Paper 2

could be used to estimate the error made with mutual information when working with a small dataset, or to estimate the amount of data that need to be collected to have an acceptable estimation.

The Hodgkin-Huxley-type model we have used here was designed to mimic the integration properties of thalamic relay cells. Of course, a large variety of neurons exists, and with that also a variety of ways to integrate signals. The results presented in this thesis should thus be considered with these limitations in mind.

3.2 Energetic efficiency of information transfer at LGN synapses

Synapses are the major locus of energy consumption in the brain. Moreover, the brain has limited energy supply and reliable information transmission has a high energetic cost for the brain. These energetic constraints probably have shaped the information processing capability of neurons and this could explain why some neurons have a probability of synaptic transmission as low as 20%. A low probability of transmission means that neurons do not maximize the information they could convey. They seem to instead maximize the ratio between the information they do convey and the energy necessary to convey this information. This is referred to as energetic efficiency of information transfer. The efficacy of transmission can be linked to presynaptic properties, like the vesicle release probability, or postsynaptic properties, like the synaptic gain (for instance the number of glutamate receptors in the postsynaptic membrane, which relates directly to the amplitude of EPSCs).

As explained in Paper 1, information transmission is sigmoidally dependent on the transmission probability of neurons while the energy used to convey this information depends essentially linearly on this probability. The ratio between information and energy as a function of the probability of information transmission is thus a monophasic curve with a single optimum. This maximum represents the energetic optimum for information transfer and usually sits at a low probability of transmission. This was shown computationally at generic synapses [17] and experimentally at thalamic relay synapses [68].

In Paper 2, we wanted to investigate if we could use the model we developed to reproduce those previous findings of energetic efficiency in thalamic relay cells, using the metrics we studied to quantify the information. We injected varying *in vivo*-like conductances in our Hodgkin-Huxley-type model, represented by a varying postsynaptic gain, gain = 1 being the experimentally-observed physiological conductance. This model is powerful in this case because it can be used to generate *in vivo*-like spike trains with any numbers of repetitions and length.

3.2.1 Energetic efficiency with our model

As shown in Paper 2, the information conveyed by the neuron calculated with I, TE_{peak} and TE_{sum} varies sigmoidally with the postsynaptic gain and the energy consumed varies linearly with the gain. This gives a curve for the energetic efficiency (information divided by energy) with a single well-defined peak close to the physiological synaptic gain for the information calculated with I and TE_{sum}. These results are in agreement with what was published earlier about energetic efficiency in thalamic relay cells [17, 68].

On the other hand, when TE_{peak} is used, the peak is much broader with no discernable maximum for postsynaptic gains between $\sim$ 1 - 5. As discussed in Section 3.1.2, it is likely that TE_{sum} is a better approximation of the information than TE_{peak}. This result also seems to point to this conclusion.

In Paper 2, we calculated the energetic efficiency of information transfer when taking into account the energy consumption of EPSCs only, and also when taking into account the total energy consumption at the synapse (this means essentially to also take into account the energy used by output action potentials). EPSCs are the major consumers of energy at synapses, but as shown in Figure 6B of Paper 2, the total energy budget makes a significant difference on the energy used versus the postsynaptic gain. However, Figure 6C and D show that the curves obtained when calculating the ratio between the information flow (either calculated with MI, TE_{peak} or TE_{sum}) and either the energy used by the EPSCs or the total energy are qualitatively equivalent.

The synapse studied here (the retinogeniculate synapse) is a particular synapse in the brain. As shown by Budisantoso and colleagues [125], the synapse at thalamic relay cells is formed by a synaptic appendage with several tens of release sites. This is very different from synapses formed in the cortex for example, which are usually composed of a few distinct release sites [56]. In this thesis, we focused on energetic efficiency of information transfer in LGN neurons, but more recent studies have also demonstrated energetic efficiency in other neurons, with very different synapses. It was for example shown with respect to postsynaptic gain at cortical synapses [126] and with respect to presynaptic release probability at hippocampal synapses [127].

Energetic efficiency of information transfer could thus be widespread in the cortex, even maybe a generic principle in the brain. Although our model aims to correctly reproduce neuronal behaviour in terms of information transmission, it is based on only one type of cells and of course is made with approximations. Nonetheless, our approach combined with models appropriate for other types of neuron or circuits could be used to better understand and study energetic efficiency of information transmission across the brain. For example, this approach is the one used to study energetic efficiency in layer 4 spiny stellate cells in [126].

Interestingly, most of the studies focusing on energetic efficiency of information transfer focus either on the presynaptic properties (mainly the release probability) or the postsynaptic properties (mainly the synaptic gain and the size of EPSCs). Is is nonetheless possible that energetic efficiency of information transfer at synapses is a mixture of both pre- and postsynaptic properties. Indeed, releasing more neurotransmitters would not have an effect if receptors are not present in sufficient numbers on the postsynaptic membrane, and, inversely, adding receptors in the membrane would not change anything if there was not sufficient neurotransmitters released to bind to them all. As we will discuss in Section 3.3.2, Paper 3 was our first attempt at linking pre- and postsynaptic sides.

3.3 Neuromodulation by serotonin

When neuromodulation and synaptic depression occur, they can change the release probability and the amplitudes of EPSCs, thus possibly altering information transmission and energy consumption. What is then the effect on energetic efficiency of information transfer? Paper 3 focused on studying the effect of the neuromodulator serotonin at LGN synapses by adding synaptic depression in the Hodgkin-Huxley-type model developed in Paper 2.

Paired-pulse depression and neuromodulation at thalamic relay synapse were added in our Hodgkin-Huxley-type model with an adaptation of the model of synaptic depression developed by Tsodyks and Markram [121], where the amplitudes of the conductances are modulated by synaptic depression. The model, fitted on experimental data, showed very good agreement with electrophysiological recordings and allowed us to run simulations of any length and with any synaptic gains. The model thus offers a simple yet efficient framework to study the effect of serotonin on thalamic relay cells by varying one unique parameter (see Paper 3).

3.3.1 Modelling of neurmodulation by serotonin

Paper 3 showed that, with the exact same input injected in the model, outputs generated with or without serotonin are largely different. There is generally fewer spikes in the case without serotonin, and the timing of output spikes can be affected. This happens because the activation of serotonin receptors at thalamic relay synapses modulates the release probability of vesicles, resulting in smaller EPSCs, but with more similar amplitudes over time.

This means that when the same stimuli is detected by the retina, the signal reaching the primary visual cortex can be significantly different depending on the level of serotonin in the LGN, affecting the quantity and type of information reaching the cortex. In other words, the information processing

capacity of the LGN is affected by serotonin. In the next Section, we will discuss exactly how the information transferred changes, but this result already shows that information transmission is affected by neuromodulation, as expected. This also means that different input sequences will maximize the information transmission between retina and primary visual cortex in different circumstances. This thesis focuses on energetic efficiency of information transfer, and this is what will be discussed in the next Section, but this is a widely interesting subject and our model will be the basis for further studies on the question.

3.3.2 Effect of serotonin on energetic efficiency

Release of serotonin at LGN synapses changes the values of information and energy when plotted against the gain, but the shape of the curves are still very similar to the case without serotonin. Those results were not published in Paper 3 and can be found here instead. In the preliminary results we show here and in Paper 3, we use the mutual information to evaluate the information. We chose to start with mutual information because it is immune to frame shifts, as shown in Paper 2. Simulations using transfer entropy to evaluate the information are being performed at the time this thesis is written. As can be seen in Figure 3.4, information and energy are slightly reduced with serotonin compared to the case without serotonin. It is not surprising that the presence of serotonin reduces energy consumption, as serotonin reduces the vesicular release probability. The reduction of information transmitted is maybe less intuitive, because, as it can be seen in Paper 3, serotonin actually increases the number of spikes in the output response. But, because inputs are generated in the same way in the cases with and without serotonin, more output spikes are not necessarily more informative.

Like in Paper 2, results obtained here also show that the curves of energetic efficiency of information transfer are qualitatively equivalent when considering only the energy consumption of EPSCs, or when considering the total energy consumption (see Figure 3 of Paper 3).

As shown in Figure 3 of Paper 3, the curve of the ratio between information transmitted and energy consumed is very similar in the cases with and without serotonin. In the case with serotonin, the maximum is slightly shifted to the right and the overall peak is slightly broadened. However, the maximum of the ratio between information and energy is still located at a postsynaptic gain close to the physiological gain. This means that the release of serotonin at this synapse does not alter the energetic efficiency principle.

This could also be an explanation of why less information is transmitted in the case with serotonin. As discussed above, energy consumption is reduced because fewer vesicles are released. If the synapse wants to stay at an energetic efficiency level, it thus has to reduce its information transmission

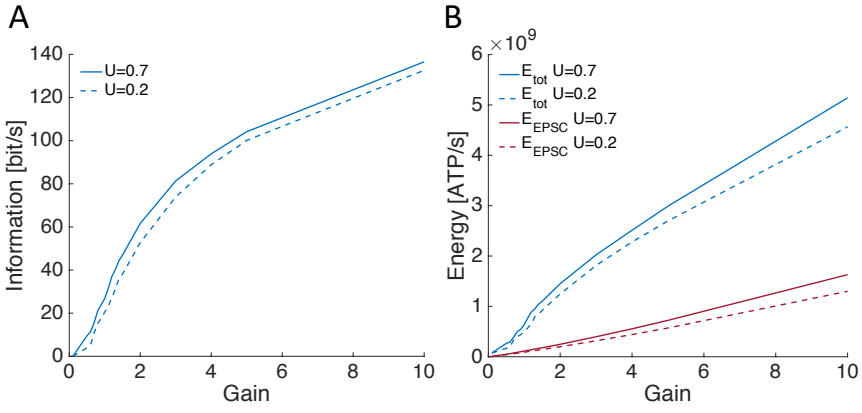

Figure 3.4: Comparison between the cases without serotonin ($U = 0.7$; see Paper 3 for a definition of U) and with serotonin ($U = 0.2$) for the information conveyed at the synapse (calculated with the mutual information; A) and the energy consumed (B) as functions of the postsynaptic gain of the cell. B shows the energy when accounting only for currents modulating the postsynaptic conductance (E_{EPSC}), and the total energy when accounting for all currents in the postsynaptic cell (*i.e.* including currents that underlie action potentials; E_{tot}).

as well, and this is achieved through synaptic depression. As explained in Chapter 1, synaptic depression is a consequence of a decrease of the number of vesicle released. This could indicate that synaptic depression and synaptic activity more generally were designed through evolution of the brain to respect the principle of energetic efficiency of information transfer.

As stated above, in this study, we are considering the pre- and post-synaptic properties of the synapse together. Indeed, neuromodulation by serotonin changes the presynaptic release probability, and we study the energetic efficiency of information transfer with respect to the postsynaptic gain. This study is thus a first step towards a more general understanding of energetic efficiency of information transfer with a more holistic approach of synaptic properties.

3.4 Limitations of the model

Designing a model made to reproduce biophysics is a very complex question. Moreover, calibrating a model from experimental biological data can be very difficult. As an example, Figure 3.5 shows how the AMPA (α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid; Figure 3.5A) and NMDA (N-methyl-D-aspartate; Figure 3.5B) conductances used in the Hodgkin-Huxley-type model of Paper 2 were determined from experimental measurements. Both conductances were recorded separately, and every black lines in both

panels of Figure 3.5 is a synaptic conductance triggered by an individual input spike. Figure 3.5 illustrates the level of noise one deals with in such circumstances.

The AMPA and NMDA conductances used in our Hodgkin-Huxley-type model are the fit (in red) of the mean (in blue) of every individual synaptic conductances (in black), as shown in Figure 3.5. Even though the mean curves are similar to typical AMPA and NMDA conductances, it is important to remember that they are indeed the mean over a broad range of recorded conductances. Moreover, Figure 3.5 also shows that the fits obtained are not perfect, especially the one for the AMPA conductance. We suspect that the AMPA and NMDA conductances were not perfectly decorrelated during measurements, and that the bump appearing around 0.01 s for the AMPA conductance is a contribution of the NMDA conductance.

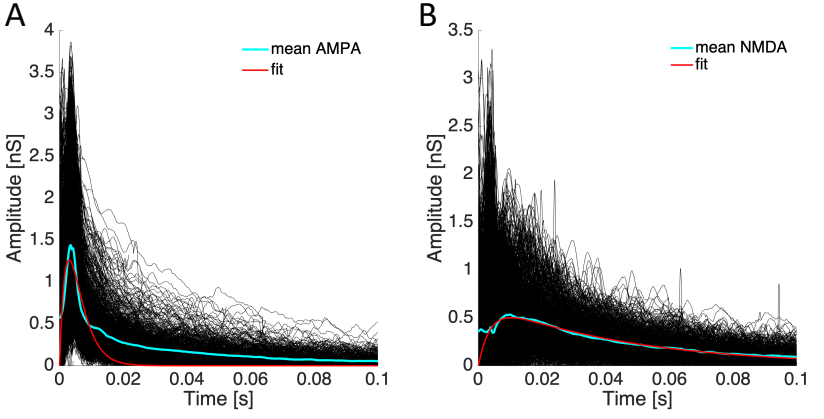

Figure 3.5: Experimentally recorded AMPA (A) and NMDA (B) conductances used to determine the conductances to be used in our Hodgkin-Huxley-type model. Each black line represents an individual experimental postsynaptic conductance. The blue line is the mean of all the black lines and the red line is the fit of the blue line. Both fits are of the form $g(t) = w(e^{-t/\tau_1} - e^{-t/\tau_2})$. Data are from reference [68].

Nonetheless, these conductances and our Hodgkin-Huxley-type model were helpful to study energetic efficiency. It is important to remember that biological data are noisy, and to keep that in mind when building a model. The model we built helped better understand the neuronal processes we studied, but it is far from perfect.

CHAPTER 4

Conclusion and Perspectives

The energy the body can allocate to the brain is limited, and this has probably shaped how neurons process information. Information is transmitted from a neuron to another when an action potential reaches the synapse, triggering the release of vesicles containing neurotransmitters. Surprisingly, in the CNS, the probability that vesicles are released when an action potential reaches the synapse is quite low. Studies suggest that this low release probability is, among other possible explanations, a consequence of the limited amount of energy available and that, instead of maximizing the information they convey, neurons maximize the ratio between the information they convey and the concomitant energy usage. This principle is referred to as energetic efficiency of information transfer. In this PhD thesis, we set out to better understand this energetic efficiency from a computational point of view, focusing on neurons in the visual pathway. The circuit studied in this thesis is a feed-forward circuit, with the retina transmitting information to thalamic relay cells in the LGN, with a one-to-one connection between those cells. Thalamic relay cells then convey information to the visual cortex.

We first studied the relative performance of two metrics of information theory (mutual information and transfer entropy) on datasets of different sizes. Data collection can be limited in Neuroscience experiments and those metrics can suffer form biases when applied to small datasets. We showed that, in simple cases of input-output generation (that are not necessarily biologically realistic), transfer entropy is more accurate to approximate the information than mutual information, especially for small dataset. But, when moving to a more complex (and biologically realistic) scenario, *via* a

Hodgkin-Huxley-type model, we showed that the transmission properties of the system can be an issue when working with transfer entropy and we discussed two different ways of estimating the transfer entropy. On the other hand, the correction method used for the mutual information (the direct method [91]) makes it immune to this kind of issues. Thus, mutual information seems to be a more robust choice, but an estimation of the error in the case of a small dataset would be needed.

We then used those metrics and our model to reproduce experimental findings about energetic efficiency of information transfer in thalamic relay cells. Our simulations showed, as expected, that the energy used to convey information varies linearly with the postsynaptic gain of the cell and that the information conveyed by the cell varies sigmoidally with this gain. This led to the ratio of information over energy to be a monophasic curve with a single optimum, in accordance to what was observed in experimental studies [68]. Moreover, the maximum of this function sat at a synaptic gain very close to the physiological gain of the cell, showing that the neuron we modelled respect the principle of energetic efficiency of information transfer.

Finally, we assessed what happens to energetic efficiency when neuromodulation occurs, by simulating serotonin induced paired-pulse depression at our thalamic relay synapses, using a variation of the Tsodyks-Markram model [121]. We showed that information and energy have a behaviour very similar to the case without serotonin, but that they are both reduced when serotonin is injected. Interestingly, this leads to almost no changes when looking at the ratio between information and energy, with the maximum still sitting close to the physiological gain. Those results indicate that energetic efficiency of information transfer could be a generic design in the brain and that energetic constraints have probably shaped brain function during evolution. The results we show in this thesis have been computed using the mutual information to evaluate the information flow, but the same study using transfer entropy is in progress.

More studies will be necessary to determine if energetic efficiency of information transfer at synapses is truly a generic design in the brain and how it is achieved in neural circuits. For this purpose, the approach used in this thesis could be generalized to other types of cells, with appropriate models. Better understanding energetic efficiency of information transfer could indeed be easier if studied computationally rather than experimentally, as it is very difficult to evaluate experimentally energy consumption at the cellular level. All the ideas studied in this thesis will thus be needed to be tested on other scenarios, like for example what was done in layer 4 spiny stellate cells with a multi-compartment neuron model [126].

Our approach could also be used to better understand why some neurons do not work at an energetically efficient level with regard to information transfer. In some pathways, it is possible that reliable information trans-

mission is more important than energy savings. The calyx of Held and the neuromuscular junction are known to be highly reliable in terms of information transmission. Nonetheless, researches on the calyx of Held suggest that the release probability and the number of postsynaptic receptors are kept at a low level by other processes, thus limiting energy consumption [68, 128, 129]. It is also possible that some cells alternate between energetically efficient states and states with a highly reliable information transmission and thus more demanding in energy (like reference [130] suggests), to adapt, for example, to different kind of stimuli or brain activity states.

Moreover, one natural question that arises in regard with energetic efficiency is how does it emerges in neural networks? Information transfer at newly formed synapses is not necessarily immediately energetically efficient. Synaptic plasticity could be a way for the neurons to learn to operate at an energetically efficient state. Are there other synaptic features that need to be taken into account? Can all types of neurons and synapses apply the principle of energetic efficiency of information transfer? Our model could be used to study these questions, as it could be extended to transmission between more than two cells. Moreover, better understanding how energetic efficiency emerges in networks could be a very useful tool to help design neuromorphic devices (and maybe also machine learning algorithms) that would be efficient in regards to computation and power consumption.

Our simulations with neuromodulation by serotonin showed that, when release probability is changed, information flux and energy consumption are modified in a way that allows the information transfer to still be energetically efficient in regards to information transmission. Other phenomenons can change synaptic properties during the life of a neuron, for example short- and long-term plasticity. Our model could also be used to asses what happens to energetic efficiency in those cases. This approach could also be used to study other neuromodulators than serotonin.

Bibliography

[1] J. W. Mink, R. J. Blumenschine, and D. B. Adams, "Ratio of central nervous system to body metabolism in vertebrates: its constancy and functional basis", American Journal of Physiology-Regulatory, Integrative and Comparative Physiology **241**, R203–R212 (1981).

[2] J. S. Rigden, *Macmillan Encyclopedia of Physics* (Simon & Schuster Macmillan, 1996).

[3] D. Silver et al., "Mastering the game of Go with deep neural networks and tree search", Nature **529**, 484–489 (2016).

[4] J. Mattheij, *Another way of looking at Lee Sedol vs AlphaGo · jacques mattheij*, https://jacquesmattheij.com/another-way-of-looking-at-lee-sedol-vs-alphago/.

[5] M. Le Roux and P. Mollard, *Game over? New AI challenge to human smarts (Update)*, https://phys.org/news/2016-03-game-ai-human-smarts.html.

[6] L. C. Aiello and P. Wheeler, "The expensive-tissue hypothesis: the brain and the digestive system in human and primate evolution", Current Anthropology **36**, 199–221 (1995).

[7] S. Herculano-Houzel, "Scaling of brain metabolism with a fixed energy budget per neuron: implications for neuronal activity, plasticity and evolution", PLOS ONE **6**, e17514 (2011).

[8] A. Navarrete, C. P. van Schaik, and K. Isler, "Energetics and the evolution of human brain size", Nature **480**, 91–93 (2011).

[9] C. W. Kuzawa et al., "Metabolic costs and evolutionary implications of human brain development", Proceedings of the National Academy of Sciences **111**, 13010–13015 (2014).

[10] C. A. Gleason, C. Hamm, and M. D. Jones, "Cerebral blood flow, oxygenation, and carbohydrate metabolism in immature fetal sheep in utero", American Journal of Physiology-Regulatory, Integrative and Comparative Physiology **256**, R1264–R1268 (1989).

[11] K. L. Leenders et al., "Cerebral blood flow, blood volume and oxygen utilizationnormal values and effect of age", Brain **113**, 27–47 (1990).

[12] J. J. Harris, C. Reynell, and D. Attwell, "The physiology of developmental changes in BOLD functional imaging signals", Developmental Cognitive Neuroscience **1**, 199–216 (2011).

[13] M. Erecińska and I. A. Silver, "ATP and brain function", Journal of Cerebral Blood Flow & Metabolism **9**, 2–19 (1989).

[14] D. Attwell and S. B. Laughlin, "An energy budget for signaling in the grey matter of the brain", Journal of Cerebral Blood Flow and Metabolism: Official Journal of the International Society of Cerebral Blood Flow and Metabolism **21**, 1133–1145 (2001).

[15] P. Lennie, "The cost of cortical computation", Current Biology **13**, 493–497 (2003).

[16] R. Jolivet, P. J. Magistretti, and B. Weber, "Deciphering neuron-glia compartmentalization in cortical energy metabolism", Frontiers in Neuroenergetics **1** (2009) `10.3389/neuro.14.004.2009`.

[17] J. J. Harris, R. Jolivet, and D. Attwell, "Synaptic energy use and supply", Neuron **75**, 762–777 (2012).

[18] L. Sokoloff, "Handbook of Physiology Section 1: Neurophysiology. Vol. 3", Postgraduate Medical Journal **38**, 1843–64 (1962).

[19] M. T. T. Wong-Riley, "Cytochrome oxidase: an endogenous metabolic marker for neuronal activity", Trends in Neurosciences **12**, 94–101 (1989).

[20] D. T. W. Chang, A. S. Honick, and I. J. Reynolds, "Mitochondrial trafficking to synapses in cultured primary cortical neurons", Journal of Neuroscience **26**, 7035–7045 (2006).

[21] J. T. Sakata and T. A. Jones, "Synaptic mitochondrial changes in the motor cortex following unilateral cortical lesions and motor skills training in adult male rats", Neuroscience Letters **337**, 159–162 (2003).

[22] A. F. MacAskill et al., "Miro1 is a calcium sensor for glutamate receptor-dependent localization of mitochondria at synapses", Neuron **61**, 541–555 (2009).

[23] A.-L. Lin et al., "Nonlinear coupling between cerebral blood flow, oxygen consumption, and ATP production in human visual cortex", Proceedings of the National Academy of Sciences **107**, 8446–8451 (2010).

[24] C. N. Hall et al., "Oxidative phosphorylation, not glycolysis, powers presynaptic and postsynaptic mechanisms underlying brain information processing", Journal of Neuroscience **32**, 8940–8951 (2012).

[25] A. F. MacAskill, T. A. Atkin, and J. T. Kittler, "Mitochondrial trafficking and the provision of energy and calcium buffering at excitatory synapses", European Journal of Neuroscience **32**, 231–240 (2010).

[26] Z.-H. Sheng and Q. Cai, "Mitochondrial transport in neurons: impact on synaptic homeostasis and neurodegeneration", Nature Reviews Neuroscience **13**, 77–93 (2012).

[27] L. A. Ligon and O. Steward, "Role of microtubules and actin filaments in the movement of mitochondria in the axons and dendrites of cultured hippocampal neurons", Journal of Comparative Neurology **427**, 351–361 (2000).

[28] A. Wieraszko, "Changes in the hippocampal slices energy metabolism following stimulation and long-term potentiation of Schaffer collaterals-pyramidal cell synapses tested with the 2-deoxyglucose technique", Brain Research **237**, 449–457 (1982).

[29] W. B. Potter et al., "Metabolic regulation of neuronal plasticity by the energy sensor AMPK", PLOS ONE **5**, e8996 (2010).

[30] P. Isope and B. Barbour, "Properties of unitary granule cell→Purkinje cell synapses in adult rat cerebellar slices", Journal of Neuroscience **22**, 9668–9678 (2002).

[31] N. Brunel et al., "Optimal information storage and the distribution of synaptic weights: perceptron versus Purkinje cell", Neuron **43**, 745–757 (2004).

[32] C. Howarth, C. M. Peppiatt-Wildman, and D. Attwell, "The energy use associated with neural computation in the cerebellum", Journal of Cerebral Blood Flow & Metabolism (2009) `10.1038/jcbfm.2009.231`.

[33] M. T. Scharf et al., "The energy hypothesis of sleep revisited", Progress in Neurobiology **86**, 264–280 (2008).

[34] M. Dworak et al., "Sleep and brain energy levels: ATP changes during sleep", The Journal of Neuroscience: The Official Journal of the Society for Neuroscience **30**, 9007–9016 (2010).

[35] G. Turrigiano, "Homeostatic synaptic plasticity: local and global mechanisms for stabilizing neuronal function", Cold Spring Harbor Perspectives in Biology **4**, a005736 (2012).

[36] G. Tononi and C. Cirelli, "Sleep and the price of plasticity: from synaptic and cellular homeostasis to memory consolidation and integration", Neuron **81**, 12–34 (2014).

[37] R. Sims et al., "Astrocyte and neuronal plasticity in the somatosensory system", Neural Plasticity **2015** (2015) 10.1155/2015/732014.

[38] R. J. Youle and D. P. Narendra, "Mechanisms of mitophagy", Nature Reviews Molecular Cell Biology **12**, 9–14 (2011).

[39] M. Damiano et al., "Mitochondria in Huntington's disease", Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease, Mitochondrial Dysfunction **1802**, 52–61 (2010).

[40] I. Maurer, S. Zierz, and H. Möller, "A selective defect of cytochrome c oxidase is present in brain of Alzheimer disease patients", Neurobiology of Aging **21**, 455–462 (2000).

[41] M. T. Lin and M. F. Beal, "Mitochondrial dysfunction and oxidative stress in neurodegenerative diseases", Nature **443**, 787–795 (2006).

[42] C. Vande Velde et al., "The neuroprotective factor Wlds does not attenuate mutant SOD1-mediated motor neuron disease", Neuro-Molecular Medicine **5**, 193–203 (2004).

[43] Hofmeijer Jeannette and van Putten Michel J.A.M., "Ischemic cerebral damage", Stroke **43**, 607–615 (2012).

[44] K. Friston, J. Kilner, and L. Harrison, "A free energy principle for the brain", Journal of Physiology-Paris, Theoretical and Computational Neuroscience: Understanding Brain Functions **100**, 70–87 (2006).

[45] K. Friston, "The free-energy principle: a unified brain theory?", Nature Reviews Neuroscience **11**, 127–138 (2010).

[46] P. C. Fletcher and C. D. Frith, "Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia", Nature Reviews. Neuroscience **10**, 48–58 (2009).

[47] P. Schwartenbeck et al., "Optimal inference with suboptimal models: Addiction and active Bayesian inference", Medical Hypotheses **84**, 109–117 (2015).

[48] K. Friston et al., "The dysconnection hypothesis (2016)", Schizophrenia Research **176**, 83–94 (2016).

[49] J. E. Niven and S. B. Laughlin, "Energy limitation as a selective pressure on the evolution of sensory systems", The Journal of Experimental Biology **211**, 1792–1804 (2008).

[50] E. R. Kandel, J. H. Schwartz, and T. M. Jessell, eds., *Principles of neural science*, 4th ed (McGraw-Hill, Health Professions Division, New York, 2000).

[51] A. Zador, "Impact of synaptic unreliability on the information transmitted by spiking neurons", Journal of Neurophysiology **79**, 1219–1229 (1998).

[52] M. S. Goldman, "Enhancement of information transmission efficiency by synaptic failures", Neural Computation **16**, 1137–1162 (2004).

[53] L. R. Varshney, P. J. Sjöström, and D. B. Chklovskii, "Optimal information storage in noisy synapses under resource constraints", Neuron **52**, 409–423 (2006).

[54] W. B. Levy and R. A. Baxter, "Energy-efficient neuronal computation via quantal synaptic failures", The Journal of Neuroscience: The Official Journal of the Society for Neuroscience **22**, 4746–4755 (2002).

[55] V. Braitenberg and A. Schüz, *Cortex: statistics and geometry of neuronal connectivity* (Springer Science & Business Media, Mar. 2013).

[56] A. M. Zador, "Synaptic connectivity and computation", Nature Neuroscience **4**, 1157–1158 (2001).

[57] R. R. de Ruyter van Steveninck and S. B. Laughlin, "The rate of information transfer at graded-potential synapses", Nature **379**, 642–645 (1996).

[58] A. Manwani and C. Koch, "Detecting and estimating signals over noisy and unreliable synapses: information-theoretic analysis", Neural Computation **13**, 1–33 (2001).

[59] H.-T. Xu et al., "Choice of cranial window type for in vivo imaging affects dendritic spine turnover in the cortex", Nature Neuroscience **10**, 549–551 (2007).

[60] W. B. Levy and R. A. Baxter, "Energy efficient neural codes", Neural Computation **8**, 531–543 (1996).

[61] S. B. Laughlin, R. R. de Ruyter van Steveninck, and J. C. Anderson, "The metabolic cost of neural information", Nature Neuroscience **1**, 36–41 (1998).

[62] V. Balasubramanian, D. Kimber, and M. J. Berry, "Metabolically efficient information processing", Neural computation **13**, 799–815 (2001).

[63] C. Shannon, "A mathematical theory of communication", The Bell System Technical Journal **27**, 379–423 (1948).

[64] J. A. Perge et al., "How the optic nerve allocates space, energy capacity, and information", Journal of Neuroscience **29**, 7917–7928 (2009).

[65] V. S. Dani et al., "Reduced cortical activity due to a shift in the balance between excitation and inhibition in a mouse model of Rett Syndrome", Proceedings of the National Academy of Sciences **102**, 12560–12565 (2005).

[66] N. R. Hardingham et al., "Quantal analysis reveals a functional correlation between presynaptic and postsynaptic efficacy in excitatory connections from rat neocortex", The Journal of Neuroscience: The Official Journal of the Society for Neuroscience **30**, 1441–1451 (2010).

[67] T. Branco et al., "Local dendritic activity sets release probability at hippocampal synapses", Neuron **59**, 475–485 (2008).

[68] J. J. Harris et al., "Energy-efficient information transfer by visual pathway synapses", Current Biology **25**, 3151–3160 (2015).

[69] L. Martignon, "Information theory", in *International Encyclopedia of the Social & Behavioral Sciences*, edited by N. J. Smelser and P. B. Baltes (Pergamon, Oxford, Jan. 2001), pp. 7476–7480.

[70] M. Cobb, *The idea of the brain: the past and future of Neuroscience* (Basic Books, New York, Apr. 2020).

[71] D. M. MacKay and W. S. McCulloch, "The limiting information capacity of a neuronal link", The bulletin of mathematical biophysics **14**, 127–135 (1952).

[72] W. S. McCulloch, "An upper bound on the informational capacity of a synapse", in Proceedings of the 1952 ACM national meeting (Pittsburgh), ACM '52 (May 1952), pp. 113–117.

[73] A. Rapoport and W. J. Horvath, "The theoretical channel capacity of a single neuron as determined by various coding systems", Information and Control **3**, 335–350 (1960).

[74] A. G. Dimitrov, A. A. Lazar, and J. D. Victor, "Information theory in neuroscience", Journal of computational neuroscience **30**, 1–5 (2011).

[75] E. Piasini and S. Panzeri, "Information theory in Neuroscience", Entropy **21**, 62 (2019).

[76] J. D. Victor, "Approaches to information-theoretic analysis of neural activity", Biological Theory **1**, 302–316 (2006).

[77] N. M. Timme and C. Lapish, "A tutorial for information theory in Neuroscience", eNeuro **5** (2018) 10.1523/ENEURO.0052-18.2018.

[78] P. Adriaans, "Information", in *Stanford Encylopedia of Philosophy*, edited by E. N. Zalta, U. Nodelman, and C. Allen (Oct. 2012), p. 66.

[79] T. M. Cover and J. A. Thomas, *Elements of information theory* (John Wiley & Sons, Nov. 2012).

[80] W. Bialek et al., "Reading a neural code", Science (New York, N.Y.) **252**, 1854–1857 (1991).

[81] N. Brenner et al., "Synergy in a neural code", Neural Computation **12**, 1531–1552 (2000).

[82] L. M. A. Bettencourt, V. Gintautas, and M. I. Ham, "Identification of functional information subgraphs in complex networks", Physical Review Letters **100**, 238701 (2008).

[83] N. M. Timme et al., "High-degree neurons feed cortical computations", PLoS computational biology **12**, e1004858 (2016).

[84] J. Avery, *Information theory and evolution* (World Scientific, 2003).

[85] n. Schreiber, "Measuring information transfer", Physical Review Letters **85**, 461–464 (2000).

[86] C. S. Daw, C. E. A. Finney, and E. R. Tracy, "A review of symbolic analysis of experimental data", Review of Scientific Instruments **74**, 915–930 (2003).

[87] R. de Ruyter van Steveninck, W. Bialek, and H. B. Barlow, "Real-time performance of a movement-sensitive neuron in the blowfly visual system: coding and information transfer in short spike sequences", Proceedings of the Royal Society of London. Series B. Biological Sciences **234**, 379–414 (1988).

[88] S. Panzeri et al., "Correcting for the sampling bias problem in spike train information measures", Journal of Neurophysiology **98**, 1064–1072 (2007).

[89] A. M. T. Ramos and E. E. N. Macau, "Minimum sample size for reliable causal inference using transfer entropy", Entropy **19**, 150 (2017).

[90] A. Treves and S. Panzeri, "The upward bias in measures of information derived from limited data samples", Neural Computation **7**, 399–407 (1995).

[91] S. P. Strong et al., "Entropy and information in neural spike trains", Physical Review Letters **80**, 197–200 (1998).

[92] P. Reinagel and R. C. Reid, "Temporal coding of visual information in the thalamus", The Journal of Neuroscience: The Official Journal of the Society for Neuroscience **20**, 5392–5400 (2000).

[93] D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex", The Journal of Physiology **160**, 106–154.2 (1962).

[94] A. Y. Deutch, "Chapter 6 - Neurotransmitters", in *Fundamental Neuroscience (Fourth Edition)*, edited by L. R. Squire et al. (Academic Press, San Diego, Jan. 2013), pp. 117–138.

[95] B. E. Alger and J. Kim, "Supply and demand for endocannabinoids", Trends in Neurosciences **34**, 304–315 (2011).

[96] A. K. Mustafa, M. M. Gadalla, and S. H. Snyder, "Signaling by gasotransmitters", Science Signaling **2**, re2–re2 (2009).

[97] L. F. Mohammad-Zadeh, L. Moses, and S. M. Gwaltney-Brant, "Serotonin: a review", Journal of Veterinary Pharmacology and Therapeutics **31**, 187–199 (2008).

[98] M. Berger, J. A. Gray, and B. L. Roth, "The expanded biology of serotonin", Annual Review of Medicine **60**, 355–366 (2009).

[99] M. M. Rapport, A. A. Green, and I. H. Page, "Crystalline serotonin", Science (New York, N.Y.) **108**, 329–330 (1948).

[100] B. M. Twarog and I. H. Page, "Serotonin content of some mammalian tissues and urine and a method for its determination", The American Journal of Physiology **175**, 157–161 (1953).

[101] A. H. Amin, T. B. Crawford, and J. H. Gaddum, "The distribution of substance P and 5-hydroxytryptamine in the central nervous system of the dog", The Journal of Physiology **126**, 596–618 (1954).

[102] B. B. Brodie and P. A. Shore, "A concept for a role of serotonin and norepinephrine as chemical mediators in the brain", Annals of the New York Academy of Sciences **66**, 631–642 (1957).

[103] W. K. Kroeze, K. Kristiansen, and B. L. Roth, "Molecular biology of serotonin receptors structure and function at the molecular level", Current Topics in Medicinal Chemistry **2**, 507–528 (2002).

[104] E. B. Goldstein, *Sensation and perception, 4th ed*, Sensation and Perception, 4th Ed (Thomson Brooks/Cole Publishing Co, Belmont, CA, US, 1996), pp. xxii, 681.

[105] D. A. McCormick, "Neurotransmitter actions in the thalamus and cerebral cortex and their role in neuromodulation of thalamocortical activity", Progress in Neurobiology **39**, 337–388 (1992).

[106] M. Steriade and D. Contreras, "Spike-wave complexes and fast components of cortically generated seizures. i. role of neocortex and thalamus", Journal of Neurophysiology **80**, 1439–1455 (1998).

[107] V. Crunelli and N. Leresche, "A role for GABAB receptors in excitation and inhibition of thalamocortical cells", Trends in Neurosciences **14**, 16–21 (1991).

[108] D. A. McCormick and T. Bal, "Sensory gating mechanisms of the thalamus", Current Opinion in Neurobiology **4**, 550–556 (1994).

[109] R. J. Miller, "Presynaptic receptors", Annual Review of Pharmacology and Toxicology **38**, 201–227 (1998).

[110] C. Chen and W. G. Regehr, "Presynaptic modulation of the retinogeniculate synapse", Journal of Neuroscience **23**, 3130–3135 (2003).

[111] A. D. De Lima and W. Singer, "The serotoninergic fibers in the dorsal lateral geniculate nucleus of the cat: distribution and synaptic connections demonstrated with immunocytochemistry", The Journal of Comparative Neurology **258**, 339–351 (1987).

[112] G. C. Papadopoulos and J. G. Parnavelas, "Distribution and synaptic organization of serotoninergic and noradrenergic axons in the lateral geniculate nucleus of the rat", The Journal of Comparative Neurology **294**, 345–355 (1990).

[113] A. Dinopoulos, I. Dori, and J. G. Parnavelas, "Serotonergic innervation of the lateral geniculate nucleus of the rat during postnatal development: a light and electron microscopic immunocytochemical analysis", The Journal of Comparative Neurology **363**, 532–544 (1995).

[114] D. P. Seeburg, X. Liu, and C. Chen, "Frequency-dependent modulation of retinogeniculate transmission by serotonin", Journal of Neuroscience **24**, 10950–10962 (2004).

[115] D. A. McCormick and H. C. Pape, "Noradrenergic and serotonergic modulation of a hyperpolarization-activated cation current in thalamic relay neurones", The Journal of Physiology **431**, 319–342 (1990).

[116] D. A. McCormick, "Neurotransmitter actions in the thalamus and cerebral cortex", Journal of Clinical Neurophysiology: Official Publication of the American Electroencephalographic Society **9**, 212–223 (1992).

[117] J. E. Monckton and D. A. McCormick, "Neuromodulatory role of serotonin in the ferret thalamus", Journal of Neurophysiology **87**, 2124–2136 (2002).

[118] S. B. Nelson and G. G. Turrigiano, "Synaptic depression: a key player in the cortical balancing act", Nature Neuroscience **1**, 539–541 (1998).

[119] J. A. Varela et al., "A quantitative description of short-term plasticity at excitatory synapses in layer 2/3 of rat primary visual cortex", Journal of Neuroscience **17**, 7926–7940 (1997).

[120] M. Galarreta and S. Hestrin, "Frequency-dependent synaptic depression and the balance of excitation and inhibition in the neocortex", Nature Neuroscience **1**, 587–594 (1998).

[121] M. V. Tsodyks and H. Markram, "The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability", Proceedings of the National Academy of Sciences **94**, 719–723 (1997).

[122] S. Ito et al., "Extending transfer entropy improves identification of effective connectivity in a spiking cortical network model", PLOS ONE **6**, e27431 (2011).

[123] M. London et al., "The information efficacy of a synapse", Nature Neuroscience **5**, 332–340 (2002).

[124] Z. F. Mainen and T. J. Sejnowski, "Reliability of spike timing in neocortical neurons", Science (New York, N.Y.) **268**, 1503–1506 (1995).

[125] T. Budisantoso et al., "Mechanisms underlying signal filtering at a multisynapse contact", Journal of Neuroscience **32**, 2357–2376 (2012).

[126] J. J. Harris et al., "Energy-efficient information transfer at thalamocortical synapses", PLOS Computational Biology **15**, e1007226 (2019).

[127] G. Mahajan and S. Nadkarni, "Local design principles at hippocampal synapses revealed by an energy-information trade-off", bioRxiv, 748400 (2019).

[128] H. Taschenberger et al., "Optimizing synaptic architecture and efficiency for high-frequency transmission", Neuron **36**, 1127–1143 (2002).

[129] M. J. Fedchyshyn and L.-Y. Wang, "Developmental transformation of the release modality at the calyx of Held synapse", The Journal of Neuroscience: The Official Journal of the Society for Neuroscience **25**, 4131–4140 (2005).

[130] T. Ngodup et al., "Activity-dependent, homeostatic regulation of neurotransmitter release from auditory nerve fibers", Proceedings of the National Academy of Sciences of the United States of America **112**, 6479–6484 (2015).