EUGENIO PIASINI

NETWORK STRUCTURE AND FUNCTION IN THE INPUT STAGE OF THE CEREBELLAR CORTEX

NETWORK STRUCTURE AND FUNCTION IN THE INPUT STAGE OF THE CEREBELLAR CORTEX

EUGENIO PIASINI

Department of Neuroscience, Physiology and Pharmacology

University College London

February 2015

Thesis submitted to UCL for the degree of Doctor of Philosophy.

I, Eugenio Piasini confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

London, February 2015

_____

Eugenio Piasini

Dedicated to the memory of Gabriele Marveggio.
1960 – 2013

ABSTRACT

It has long been recognised that neuronal networks are complex systems, whose dynamics depend on the properties of the individual synapses and neurons and the way in which they are interconnected. However, establishing clear links between network structure and function has proven difficult.

To address this question I applied tools and techniques from computational neuroscience, neuroinformatics, information theory, machine learning, spatial point process theory and network theory, deploying them on a suitable HPC infrastructure where appropriate. Moreover, access to electrophysiological and anatomical data enabled me to develop biologically accurate models and to compare my theoretical predictions with analyses of raw data.

In this work, I focused on the granule cell layer (GCL), the input stage of the cerebellar cortex. The GCL is particularly well suited to this type of analysis, as its structural characteristics are comparatively regular, well known and conserved across animal species, and several of its basic functions are relatively well understood.

I showed that the synaptic connectivity in simple feed forward networks like the GCL governs the trade-off between information transmission and sparsification of incoming signals. This suggests a link between the functional requirements for the network and the strong evolutionary conservation of the anatomy of the cerebellar GCL. Furthermore, I investigated how the geometry of the GCL interacts with the spatial constraints of synaptic connectivity and gives rise to the statistical features of the chemically and electrically coupled networks formed by mossy fibres, granule cells and Golgi cells. Finally, I studied the influence of the spatial structure of the Golgi cell network on the robustness of the synchronous activity state it can support.

Regòrdes pìsceṇa: u péṇa, u šchéṇa...

— Attilio Fomiatti (di zevìj)

ACKNOWLEDGMENTS

First of all, I would like to thank my supervisor, prof. Angus Silver, for his support and guidance. Besides his talent as a scientist, one of Angus's best quality as a supervisor has been that to know when to gently steer me away from running off a cliff and when to encourage me to follow my inspiration. He did so with attention and generosity. Working with him has been motivating and inspiring — if somewhat intense at times —, and I hope I have absorbed some of his scientific rigour and intellectual honesty.

Three people that were particularly involved in my daily work at the Silver lab deserve a special mention. Padraig Gleeson has been a constant presence from the first to the last day, never failing to find some time to help me even at the busiest times. Guy Billings displayed an incredible care and patience in mentoring me, always reminding me not to lose sight of the big picture, and pulling me out of the occasional rabbit hole when I got stuck. Bóris Marin is just an all-around amazing person; I still can't quite wrap my head around how he does it (I suspect he lives 48-hours days in some alternate dimension or something), but having him around is a unlimited source of new ideas, and stimuli, and passion. (And hardcore geekdom). I'm lucky to have him as a friend.

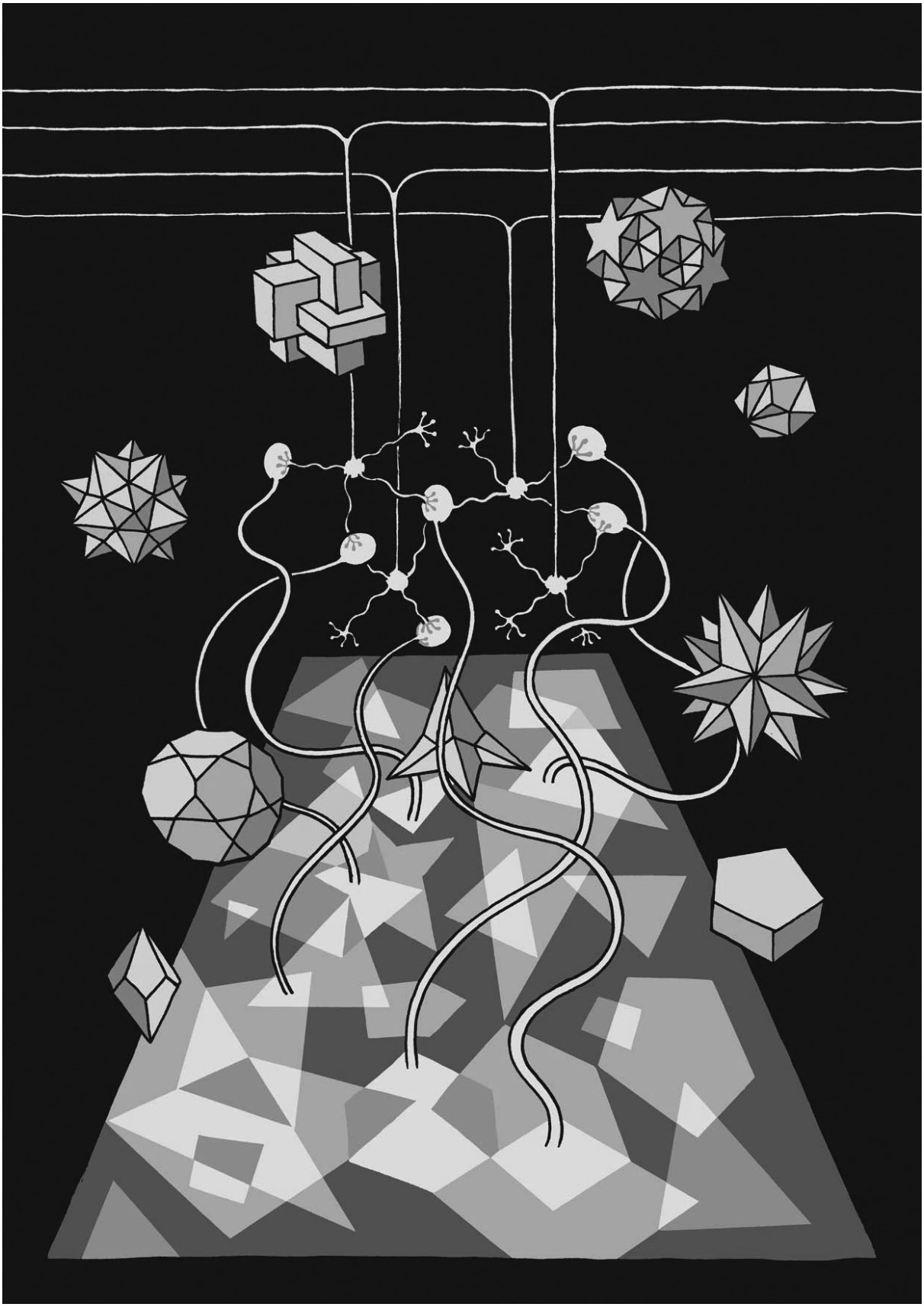
Speaking of triplets: three other mentors, from life before the Silver lab, made it possible for me to understand where I wanted to go, and ultimately allowed me to get there. Sergio Solinas, Gianpiero Cattaneo and Gabriele Marveggio. To them I owe much.

My gratitude also goes to all those colleagues and friends who helped me in so many ways over the years (presented here in no particular order). To Matteo Farinella, for showing me how not to be overcome by life in London and (let's not forget) for providing the artworks that can be found in this thesis; to Daniel Ward, for helpfully providing constant dialectic opposition in the role of the "sceptical biologist", and for his honest friendship; to Matteo Cantarelli, for bringing creative disruption to the lab, and for the *filu 'e ferru* (but not the *Mamuthones*. That still gives me nightmares); to Jason Rothman, for providing tons of useful data, tolerating me asking the same things over and over again and for assistance in model archaeology (necromancy?). To Antoine Valera, for his enthusiasm and openness to new ideas, that allowed an exciting project to pop out of nowhere

(and for feedback on this thesis!). To everyone in the lab, and particularly to Chiara Baragli, László Bicskei, Yates Buckley, Geoffrey Evans, Frédéric Lanore, Hana Roš and Koen Vervaeke for the numerous helpful discussions and precious ideas they shared with me. To George Konstantinou for that project that didn't quite materialise, but still was fun to dream up (perhaps in the future...); to Capucine Le Guen for sharing the joys and sorrows of a Marie Curie programme. To Diccon Coyle, and to Adrian Quintana, for being there to catch some bad falls — even off belay. And for being really good friends.

I am indebted to all the people who supported me with their friendship and love, even from far away. My friends from Italy (these come in alphabetical order to prevent complaints): Cristian, Gaël, Gabriele, Gioele, Giovanni, Michele, Nicola and Sofia. Clélia, who managed to put up with me during the hardest months. My family, that made it all possible. *Grazie*.

Opposite page: *Sparse coding*. Courtesy of Matteo Farinella.



CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

1.1 STRUCTURE AND DYNAMICS IN COMPLEX SYSTEMS

The interplay between the structure and the collective behaviour of complex systems is a central theme in neuroscience and in many related fields. In the following I will give a brief overview of some influential ideas, originated at the interface between physics, mathematics and computer science, that constitute the broader context within which this work has developed.

1.1.1 *Artificial neural networks*

Historically, research on distributed information processing systems has progressed along two main directions: the quest to prove, in principle, *universality* (for example, in the sense of Turing [1936]) of specific approaches to computation, and the practical task of finding viable methods for configuring the structure of a given system to make its dynamics conform to a specific functional requirement. In particular, the first proof that an ensemble of "elementary" binary units can *function* as a universal computer if equipped with an appropriate connectivity *structure* dates back to McCulloch and Pitts [1943]. This fundamental insight gave birth to the field of *artificial neural networks* (ANNs), whose popularity rose and fell several times in the following decades as new network architectures and techniques for employing them were introduced (see Hertz et al. [1991] and Cowan and Sharp [1988] for an overview of the "classic period" of ANNs, and Lukoševičius and Jaeger [2009] and Schmidhuber [2015] for more recent developments).

artificial neural networks

One of the most important early contributions is due to Frank Rosenblatt and his group, who introduced the *perceptron* [Rosenblatt, 1962], modelled after an idealised mathematical representation of a neural system. The perceptron is a *linear classifier* equipped with an algorithm that can adjust its internal ("network") structure to configure it for the execution of a specific classification task. The configuration process is called *learning* or *training* of the neural network. Unfortunately, despite the initial enthusiasm for its potential applications in artificial intelligence, it soon turned out that the practical usefulness of the perceptron was severely limited by the fact that it could only learn to solve linear classification problems. This limita-

perceptron

learning

tion was shared by *adalines*, a similar ANN introduced in the same period by Widrow and Hoff [1960].

In the following years, the attention shifted to *content-addressable* and *associative memories*, built around the idea of partitioning the phase space of a dynamical system in a finite number of attractor basins corresponding to the patterns to be stored. The most successful ANN designs from this period are *Hopfield networks* [Hopfield, 1982] and *Boltzmann machines* [Hinton and Sejnowski, 1983], in which deterministic or stochastic units are arranged in a recurrent network that can be trained — again, by adjusting the strenghts of the connections constituting the network structure — to produce the desired partitioning of the phase space. Of particular interest to neuroscience is a series of works by David Marr, in which the concept of associative memory was used to formulate functional models for cerebellum, neocortex and hippocampus [Marr, 1969, 1970, 1971]. Marr's model of the cerebellum, in particular, is still one of the most influential theories of cerebellar function to date, and will be discussed in detail in section 1.3.

associative memories

Hopfield networks

Boltzmann machines

Another important development came in the 80s, when Rumelhart and Zipser [1985] and Parker [1985] independently re-discovered the *backpropagation algorithm*, first proposed in the previous decade by Werbos [1974], that makes it feasible to train a *multilayer perceptron* (MLP). An MLP is a network composed by many perceptrons connected in series. The main functional advantage it derives from the increased structural complexity with respect to a single-layer perceptron is being capable of performing nonlinear classification tasks. Training an MLP is a nonconvex optimization problem, and the backpropagation algorithm allows to compute the gradient of a loss function for the network, thus enabling gradient descent-like methods to be applied for the optimisation. The discovery of a viable training algorithm for this architecture immediately put ANNs back in the spotlight as a feasible approach to artificial intelligence. Since then, numerous variants and improvements to backpropagation (like, notably, the *natural gradient* method [Amari, 1998]) have been proposed, and MLPs have been successfully employed for real-world applications; unfortunately though, again, their practical usefulness has proven to be limited by the tendency of the training procedures to get trapped in local optima [Bengio, 2009; Glorot and Bengio, 2010]. Significant advances on the issue of training multilayer ANNs have been achieved recently with the introduction of *deep belief networks* [Hinton and Salakhutdinov, 2006], and, more generally, with the emergence of a series of techniques known under the collective name of *deep learning* [Schmidhuber, 2015].

backpropagation algorithm

natural gradient

deep learning

In the previous paragraphs, we have seen how the ANN community has found ways of harnessing the high-dimensional dynamics

of recurrent networks and the relative ease of training multilayer networks to design powerful computational devices by adapting the network structure to the desired function. Another technique emerged in the last decade combining these different aspects into a novel architecture: *reservoir computing* [Lukoševičius and Jaeger, 2009], independently proposed as *liquid state machines* by Maass et al. [2002] and *echo state networks* by Jaeger [2001]. As the name implies, this type of ANN relies on a randomly generated *reservoir*, typically composed of recurrently connected units, together with a simple linear (or MLP-like) readout that can be trained by linear regression (or, respectively, backpropagation). This type of system is proven to be universal for continuous-time, continuous-value computing, and has already found real-world application in fields such as speech recognition and time series forecasting [Lukoševičius and Jaeger, 2009].

reservoir computing

1.1.2 *Dynamical systems and complex networks*

In the dynamical systems literature, great attention has traditionally been devoted to studying *synchronisation* phenomena [Pikovsky et al., 2003]. The pioneering work by Winfree [1967] and Kuramoto [1975] showed that the dynamics of an *unstructured* collection of diffusively coupled periodic units can be studied with a mean-field approach, solving for the order parameter related to synchrony as a function of connection strength and heterogeneity in the intrinsic frequencies of the units. Later on, seminal structure-function studies — like that of Watts and Strogatz [1998] on synchronisation of cricket chirping — inaugurated the modern theory of complex networks by popularising concepts such as that of *small world connectivity* (Strogatz [2001]; Newman [2010]; but see Bullmore and Sporns [2009] for a neuroscience-centred review). Following this, a lot of effort has gone towards analysing synchronisation in structured networks [Arenas et al., 2008] for diffusive and pulsed coupling, identical or nonidentical periodic or chaotic elements [Strogatz, 2000; Lago-Fernández et al., 2000; Barahona and Pecora, 2002; Wu, 2003; Pereira et al., 2013].

synchronisation

Kuramoto model

small world connectivity

As will be discussed later (chapter 4), these methods have since found their way to many applications in neuroscience, some of which are very relevant to the topics discussed in this thesis. More generally, the complex systems literature abounds with analyses of structured systems or media in many and diverse contexts such as percolation phenomena, traffic networks, random walks, epidemic spreading, information spreading in social networks, and metabolic networks [Barrat et al., 2008].

1.2 NETWORK STRUCTURE AND FUNCTION IN NEURONAL SYSTEMS

1.2.1 *Cell morphology*

From the biophysical standpoint, neurons themselves can be modelled as electrical circuits, or networks. In particular, the effect of dendritic and axonal morphology on current flow in neurons has been extensively studied since the seminal work by Rall and colleagues [Rall, 1962; Goldstein and Rall, 1974]. Early approaches relied on finding analytical solutions for the cable equation on some classes of branching domains, or on performing numerical analyses of simplified models of membrane excitability. In the following years, increased access to computational resources and the advent of multicompartmental modelling enabled detailed studies of realistic neuronal morphologies (see Stuart et al. [2008] and Segev and Schneidman [1999] for reviews), including investigations into the effect of the geometrical structure of a cell on its firing patterns [Mainen and Sejnowski, 1996] and action potential propagation [Vetter et al., 2001]. Poirazi et al. [2003] studied how spatial effects combine with nonlinear electrical properties in dendrites, and showed that a single pyramidal neuron can indeed be modelled as an abstract multilayer neural network. In some classes of neurons, a more explicit link between structure and function has been established [Agmon-Snir et al., 1998; Gabbiani et al., 2002]. More recently, Torben-Nielsen and Stiefel [2010] proposed a systematic approach to structure-function relationships in dendrites relying on nonlinear optimisation techniques and de Sousa et al. [2014] have investigated how dendritic morphology affects pattern recognition in single neurons. These latter works exemplify the continuing effort of the research community in trying to achieve a principled understanding of the morphological determinants of dendritic and axonal dynamics and, perhaps more importantly, a global picture of which dynamical aspects are relevant for the different functional roles that neurons play in their respective networks.

1.2.2 *Neuronal circuits*

Detailed theories on the interplay between network structure and function have been formulated for several brain regions. Among the most elegant are those on the emergence of orientation selectivity in the visual system [Hubel and Wiesel, 1962; Sompolinsky and Shapley, 1997], the dynamical basis of oscillatory activity in the olfactory bulb [Li and Hopfield, 1989], the coordinate transformation involved in the conversion from sensory to motor signals in parietal cortex [Salinas and Abbott, 1995], and the generation of persistent activity under-

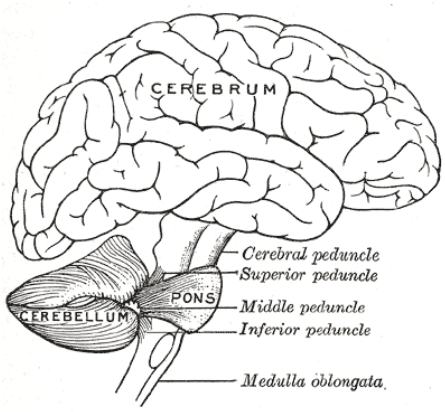


Figure 1: The cerebellum in the human brain. Figure reproduced from Lewis [1918] (public domain).

lying the storage of eye position memory [Seung, 1996] and spatial working memory [Compte et al., 2000].

The cerebellar cortex, the system under study in this thesis, merits a separate mention and will be presented in the next section. In fact, it is particularly well suited to network structure-function analysis due to its regular and relatively simple three layer structure, few neuronal cell types, and its well-established role in motor control (Eccles et al. [1967], and see below).

1.3 THE CEREBELLUM

In this section, I briefly present the salient anatomical and physiological characteristics of the cerebellum, together with an historical overview of some of the most popular theories of cerebellar function. I also provide some additional detail on the aspects that are closer to the focus of this study, such as the physiology of Golgi cells and granule cells, their synaptic inputs, the Marr-Albus functional theories and the issue of synchronisation and oscillatory activity in the granule cell layer.

1.3.1 *Anatomy and physiology*

In humans, cerebellum is located in the posterior cranial fossa; it is separated from the cerebrum by the tentorium cerebelli, and it is attached to the brainstem through the superior, inferior and intermediate peduncles (Figure 1).

Longitudinally, the cerebellum can be divided into a medial (*vermis*), an intermediate (*paravermis*) and a lateral section (*hemispheres*). On the exterior of the organ is a layer of grey matter (the *cerebellar cortex*) folded in three *lobes* (*anterior*, *posterior* and *flocculonodular*) along

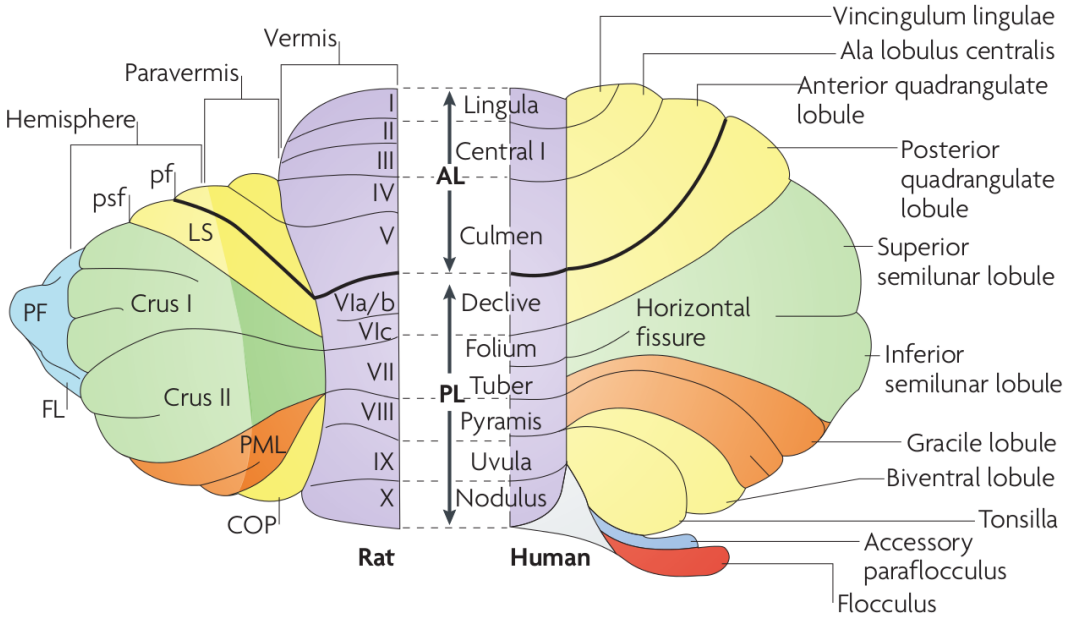


Figure 2: Dorsal view of rat and human cerebellum (resp. left and right), with commonly-used anatomical nomenclature schemes. The *primary fissure* (pf), dividing the anterior and posterior lobes, is shown in bold. AL: *anterior lobe*; COP: *copula pyramidis*; Crus I and Crus II: *ansiform lobule*; FL: *flocculus*; LS: *lobulus simplex*; PF: *paraflocculus*; PL: *posterior lobe*; PML: *paramedian lobule*; psf: *posterior superior fissure*. Figure reproduced with permission from Apps and Hawkes [2009].

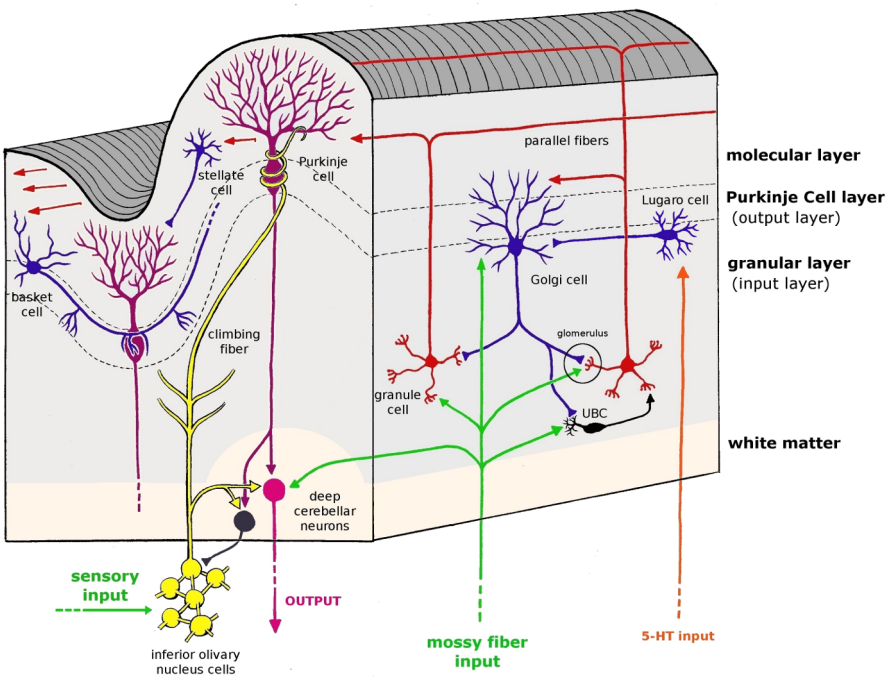


Figure 3: Schematic representation of cerebellar circuitry. Image courtesy of Matteo Farinella.

the anteroposterior axis; each lobe is itself folded in several *lobules*. In total, there are 10 lobules, each of which is further divided in a number of *folia* depending on the species. Different nomenclature schemes exist for indicating lobules and groups of lobules. Figure 2 shows the most commonly used in human and rat. Internally to the cerebellum is a region of white matter, within which four pairs of nuclei called *deep cerebellar nuclei* (DCN) are embedded. The nuclei are called *dentate*, *emboliform*, *globose* and *fastigial*; emboliform and globose nuclei are fused in some animals and termed *interposed*.

deep cerebellar nuclei

1.3.2 *The cerebellar cortex*

The cerebellar cortex is organised in three layers, populated by a small number of cell types. It receives two main types of input, and projects only one output. Its most striking characteristic is the particular symmetry of its network arrangement: in the words of Braitenberg and Atwood [1958],

> Among the well known grey sheets (including retina, telencephalic cortex, cerebellar cortex, tectum of lower vertebrates, inferior olive and certain invertebrate ganglia) the cerebellar cortex deserves a special place because of its particular order of symmetry. The other sheets can be essentially described by giving the organization as it ap-

> pears in a thin slice of tissue cut in a plane perpendicular to the plane of the sheet, the rotation of which, around a vertical axis, will produce the whole spatial organization of the sheet. The cerebellar cortex, on the other hand, can only be represented in two planes perpendicular to each other and having definite relations to the longitudinal and transversal axes of the animal. The whole three dimensional structure, therefore, cannot be obtained by rotation but by translation in two directions, thus producing a *lattice*.

Importantly, such a regular network organisation is strongly conserved across all vertebrates, and "virtually identical" [Sultan and Glickstein, 2007] in mammals and birds.

1.3.2.1 *The granule cell layer*

The input stage of the cortex is the *granule cell layer* (GCL, or *granular layer*). This layer contains *granule cells* (GrC) — after which it is named —, *Golgi cells* (GoC), *Lugaro cells* (LC) and *unipolar brush cells* (UBC) (Figure 3). Granule cells are inhibited by Golgi cells, which form a gap junction-mediated recurrent network and are themselves inhibited by Lugaro cells and possibly other cell types (see below). *Mossy fibres* carry the principal input to the cerebellar cortex from different sources including the spinal cord, the cortex via the pons and the vestibular nerve (see below). MFs branch extensively within the GCL [Wu et al., 1999; Sultan, 2001], forming en-passant presynaptic enlargements known as *rosettes*. They innervate GrCs, GoCs and UBCs. UBCs are excitatory interneurons innervating GoCs and GrCs through mossy fibre-like dendrites called *intrinsic mossy fibres* [Diño et al., 2000]. MF rosettes, GoC axonal terminations and GrC dendrites converge within *glomeruli*, tight synaptic structures where they are enclosed in a glial coat. Glomeruli facilitate spillover of neurotransmitters between neighboring release sites [Rossi and Hamann, 1998; DiGregorio et al., 2002]. Each glomerulus receives dendritic terminals from 15-20 GrCs [Eccles et al., 1967], with each GrC contacting 2 to 7 glomeruli. This, combined with the branching of MFs, ensures that each MFs contacts hundreds of GrCs [Fox et al., 1967], leading to a highly divergent network organisation between MFs and GrCs.

granule cells
Golgi cells
mossy fibres
rosettes
glomeruli

GRANULE CELLS Granule cells (GrCs) are the most abundant cell in the vertebrate brain, their number being estimated to be ~10^{11} in humans [Braitenberg and Atwood, 1958]. They are characterised by a small cell body (5–6 µm diameter in rat [Palay and Chan-Palay, 1974]). Dendritic morphology of GrCs is remarkably well conserved across animal species [Wittenberg and Wang, 2008]. They have a small number of short dendrites: in cats, Palkovits et al. [1972] reports 2 to 7

dendrites, 4 on average, with more than 60% of cells having exactly 4. Dendrites extend on average 13.59 µm from the cell body, but seldom (< 5 %) more than 30 µm and never more than 40 µm. Each dendrite receives synaptic input from a single mossy fibre [Eccles et al., 1967]; only 60 % of dendrites form direct synaptic connections with Golgi cell axon terminals [Jakab and Hámori, 1988]. Granule cells are electrotonically compact [Silver et al., 1992, 1996]; their soma and dendrites are often modelled — as I do in subsection 2.2.1 — as a single electrical compartment.

GOLGI CELLS Golgi cells (GoCs) are inhibitory GABAergic/glycinergic interneurons. Their neurochemical properties have been linked to the existence of distinct genetic subtypes, with mGluR2-positive cells being almost exclusively dual (GABAergic *and* glycinergic) and neurogranine-positive cells being either dual or exclusively GABAergic [Simat et al., 2007]. They present a dendritic arborisation divided in an *apical* part, ascending to the molecular layer, and a *basolateral* part, confined to the granular layer. Their axonal plexus extends in the granular layer, where it innervates numerous glomeruli (estimated to be around 145, for a total of ~1500 contacted GrCs [Kanichay, 2008]). GoCs receive excitatory input from mossy fibres onto their basolateral dendrites [Kanichay and Silver, 2008] and from parallel fibres onto their apical dendrites [Dieudonné, 1998; Palay and Chan-Palay, 1974]. Lugaro cells have been identified as a source of inhibitory input for Golgi cells [Dieudonné and Dumoulin, 2000], together with other populations (including molecular layer interneurons, Purkinje cells via axon collaterals and Golgi cells themselves) where the experimental evidence is less clear [Larramendi and Lemkey-Johnston, 1970; Palay and Chan-Palay, 1974; Dumoulin et al., 2001; Hull and Regehr, 2012]. Golgi cells are autorhythmic, and fire at about 8 Hz in anaesthetised rats [Vos et al., 1999] and in the 2–8 Hz range in slices in absence of excitatory inputs [Dieudonné, 1998; Forti et al., 2006], going up to 50 Hz following sensory stimulation during *in vivo* experiments [Vos et al., 1999; Van Kan et al., 1993]. Golgi cells are interconnected through electrical synapses or *gap junctions* [Dugué et al., 2009].

EXCITATORY SYNAPTIC INPUTS TO THE GRANULE CELLS Excitatory input to granule cells is mediated by activation of AMPARs and NMDARs following glutamate release at intrinsic or extrinsic mossy fibre terminals within the glomerulus [Silver et al., 1992, 1996]. The amplitude and time course of MF→GrC synaptic conductances exhibit wide trial-to-trial variability and heterogeneity across individual connections [Sargent et al., 2005]. *In vivo* recordings show that MF activity presents diverse temporal activity modes, with MFs sig-

nalling rapid discrete sensory events exhibiting high-frequency bursts and relatively quiescent periods [Jörntell and Ekerot, 2006; Rancz et al., 2007], while those that convey slower continuous sensory variables, such as joint angle and head velocity, typically fire continuously at 10–100 Hz [Van Kan et al., 1993; Arenz et al., 2008]. During rate-coded input, the excitatory input is dominated by the buildup of slow spillover-mediated AMPAR and NMDAR conductances [DiGregorio et al., 2002; Nielsen et al., 2004; Schwartz et al., 2012]. The spatial patterns of MF activation are also likely to be highly diverse. Although, as discussed below, MF innervations of the GCL exhibit a large-scale fractured map topology [Shambes et al., 1978], the MFs that innervate an individual GC typically arise from different precerebellar nuclei [Huang et al., 2013]. This suggests that MF activity can be spatially independent at the local scale, and that a key function of the GCL is to combine information from different modalities. However, MF inputs onto individual GCs in vermal areas that encode limb movement carry highly correlated signals [Jörntell and Ekerot, 2006].

INHIBITORY SYNAPTIC INPUTS TO THE GRANULE CELLS Granule cells receive two types of inhibition: a *tonic* inhibitory component, mediated by persistent activation of extrasynaptic $GABA_A$Rs caused by GABA released by glial cells [Farrant and Nusser, 2005; Lee et al., 2010; Duguid et al., 2012], and a *phasic* component, mediated by synaptic release of GABA from Golgi cells [Eccles et al., 1967] through direct [Jakab and Hámori, 1988] or spillover-only [Rossi and Hamann, 1998] connections. Because Golgi cells are innervated through excitatory synapses by both mossy fibres and granule cells, the phasic component provides feedforward [Kanichay and Silver, 2008] and feedback [Cesana et al., 2013] inhibition. In the rate-coding regime, neurotransmitter concentration dynamics at the Golgi to granule cell synapse are dominated by slow buildup of GABA spillover [Rossi and Hamann, 1998; Crowley et al., 2009; Ward, 2012]. On average, 3 direct GoC→GrC synaptic connections and 1 purely spillover-mediated connection contribute to the inhibitory input of a granule cell [Ward, 2012].

1.3.2.2 *The molecular layer*

The *molecular layer* (ML) is located on top of the GCL, and contains two types of interneurons, *stellate cells* (SC) and *basket cells* (BC). ML and BC somata are interspersed between the dendritic trees of *Purkinje cells* (PC), neatly organised in the sagittal plane, and the axonal projections of GrCs, termed parallel fibres (PF), running along the cerebellar folia orthogonally to the PC dendrites (Figure 3). Parallel fibres form excitatory connections with PCs, BCs and SCs and the

Purkinje cells

apical dendrites of GoCs extending in this layer, while BCs and SCs inhibit PCs.

1.3.2.3 *The Purkinje cell layer*

Purkinje cell somata compose the output layer of the cortex. This layer is the thinnest of the three, being only one cell body thick, and is located between the other two (Figure 3). Purkinje cell axons constitute the sole output of the cerebellar cortex. They project to the DCN (where they form inhibitory synapses), except for those originating in the flocculonodular lobe, that can also project to the vestibular nuclei in the brain stem [FitzGerald and Folan-Curran, 2002]. Each PC receives extensive synaptic contacts on its soma and proximal dendrites from a single *climbing fibre* (CF), the axonal projection of a neuron in the *inferior olive* (IO).

climbing fibres

1.3.2.4 *Afferent and efferent connections*

Even though the synaptic organisation detailed above is homogeneous throughout the cortex (with the variable spatial density of UBCs [Mugnaini et al., 2011] being a notable exception), the cerebellum is traditionally divided in three functional regions named after the areas that most contribute to their inputs (see for example Kandel et al. [2012], though this classification is disputed [Apps and Hawkes, 2009]):

- the *vestibulocerebellum* or *archicerebellum*, receiving inputs from the vestibular nerve and projecting directly to the lateral vestibular nuclei, consists in the flocculonodular lobe, and is the most primitive part of the cerebellum. Functionally, it is involved in balance and eye movement control.

- The *spinocerebellum* or *paleocerebellum* consists of the vermis and the paravermis. The former receives visual, auditory and vestibular inputs as well as somatosensory inputs related to the head and other proximal areas; it controls gaze, posture and locomotion, projecting through the fastigial nucleus to cortical and brain stem areas that govern proximal muscles. In an analogous way, the latter is involved in limb control, receiving from and projecting to related areas through the interposed nuclei.

- The *cerebrocerebellum* or *neocerebellum* consists of the lateral part of the hemispheres. It receives inputs from cortical areas of the cerebrum — mostly through the pontine nucleus — and projects to the dentate nucleus. From here, the projections are routed either to motor cortical areas through the thalamus or to the red nucleus. These connections imply involvement in motor planning, motor learning and eye movement; furthermore, the

cerebrocerebellum is thought to be involved in higher cognitive functions.

On a smaller scale, the cerebellar cortex is organised in a *fractured somatotopic map* [Shambes et al., 1978; Voogd and Glickstein, 1998; Apps and Garwicz, 2005] whereby each local area of the cortex is connected in a loop with areas related to the same body region through both its cortical or sensory afferents and its projections [Bostan et al., 2013; Proville et al., 2014]. Beside this, cerebellar circuits also participate in an olivo-cortico-nuclear loop where Purkinje cells innervated by CF originating in a specific region of the IO can inhibit a DCN area innervating that same IO region [Chaumont et al., 2013].

1.3.3 *Functional theories of the cerebellum*

Since the first detailed and extensive anatomical accounts, the structure of the cerebellar cortex has been recognised as being remarkably regular, distinctively symmetrical and well conserved across animal species [Golgi, 1885; Ramón y Cajal, 1889a,b; Ramón y Cajal, 1909]. This provided early inspiration for theories postulating the existence of a general "cerebellar algorithm" implemented consistently throughout the organ to process incoming signals. The detailed electrophysiological description of the cerebellar circuitry given by Eccles et al. [1967], in conjunction with the knowledge accumulated in decades of lesion experiments and observations [Dow and Moruzzi, 1958], spurred a wave of classic papers that built on top of the original idea of the *cerebellum as a neuronal machine* for motor control. Marr [1969] and Albus [1971] proposed two similar schemes for motor learning centered around the idea of climbing fibres providing teaching (error) signals to Purkinje cells. Ito [1970] analysed the cerebellum as a forward model from a control-theoretical standpoint. On an independent track, Braitenberg [1961] popularised the idea of the cerebellum as a timing device by considering the possible role of parallel fibres as delay lines. In the following decades, several schools of thought formed around these proposals, evolving to accommodate an ever-increasing body of experimental evidence including the discovery of long-term synaptic plasticity in the cerebellum [Ito and Kano, 1982; Armano et al., 2000; Jörntell and Ekerot, 2002; Ruediger et al., 2011; Gao et al., 2012] and encoding of motor error signals in CFs [Kitazawa et al., 1998]. The internal model idea grew into a rich theoretical framework encompassing motor control as well as certain aspects of cognition [Wolpert et al., 1998; Ito, 2008]; at the same time, the key role of cerebellar-like structures in reafferent sensory input cancellation in fish was demonstrated [Bell, 1981; Bastian, 1995; Sawtell, 2010; Kennedy et al., 2014]. Other features of the Marr-Albus theory were developed to describe the cerebellum as an adaptive filter [Fujita,

1982; Dean et al., 2010] or a bayesian state estimator [Paulin, 2005]. On a somewhat different direction, Kanerva [1988] recognised in the cerebellar cortical circuitry a possible biological implementation of a distributed random-access memory, while Braitenberg and colleagues [Braitenberg et al., 1997] kept on incorporating new anatomical data into the delay line theory. The timing idea, as a whole, was generalised into a more abstract "space-timing" version in the work by Pellionisz and Llinas [1982], where the cerebellum was seen as providing a metric tensor for a proposed representation, internal to the CNS, of external space-time events. A role for the olivo-cerebellar system as a temporal pattern generator was hypothesised and received some support on theoretical, experimental and modelling grounds [De Zeeuw et al., 2011].

Following the conceptual distinction between *computational* and *mathematical* modelling proposed by Dayan [1994], the models described above can be juxtaposed to the outcomes of a series of efforts enabled, since the end of the nineties, by the increased availability of computing power [Maex and De Schutter, 1998; Medina and Mauk, 2000; Santamaria et al., 2007; Solinas et al., 2010; Vervaeke et al., 2010, 2012; Rössert et al., 2014]. In these "bottom-up" attempts at replicating specific functional properties of the cerebellar circuitry, numerical simulations of detailed biophysical models were performed by incorporating selected structural elements in perfectly controlled (virtual) experimental environments. While very useful, these efforts focused mainly on reproducing low-level signal processing features of the network, and had, at best, only weak links to the theoretical level. As of today, despite the abundance of models — at all levels of abstraction and biological detail — and the wealth of information, the research community is still far from producing a unified vision of the connection between cerebellar anatomy and function.

1.3.4 *Network structure and function in the granule cell layer*

There are two aspects to the structural organisation of the granule cell layer that make it particularly well suited for a study on network determinants of function and dynamics. These are the peculiar, highly divergent, organisation of the MF→GrC synapse at the glomerulus, and the recurrent gap junction-mediated connectivity between Golgi cells.

1.3.4.1 *Expansion recoding in Marr-Albus theories*

In the influential theory first formulated by Marr [1969], the cerebellum is proposed to function as a hetero-associative memory, and in particular as an exhaustive library of motor commands addressable through representations of sensory and proprioceptive contexts (or

situations in Marr's terminology). This is proposed to enable efficient motor learning and control.

More specifically, in this theory MFs are thought to provide PCs with representations of sensory information through the highly divergent MF→GrC→PC pathway. PCs integrate this information with a teaching, or *error*, signal from the CFs. Hebbian learning at the GrC→PC synapse — implemented through long term potentiation (LTP) — enables PCs to conform with their firing to the content of the teaching signal, memorising the desired response to each sensory context they are presented with. Specific patterns of PC firing are then supposed to have a direct effect in eliciting determinate motor responses, and every movement is postulated to be representable as a unique sequence of PC activation patterns. Once a particular movement is memorised, whenever the sensory/proprioceptive situation associated with it is recognised (following for example a conscious command from the cerebral cortex), the cerebellum can autonomously implement the precise sequence of movements needed to complete it successfully, as long as at every instant the organism remains in a recognisable sensory situation.

In Marr's theory, the role of the granule cell layer is to pre-process input patterns to increase their separation and sparseness. This procedure, termed *expansion recoding* in later literature, has the effect of increasing the capacity of the associative memory implemented by the cerebellum (see Albus [1971]; Tyrrell and Willshaw [1992]; but also, more generally, Engel and Van den Broeck [2001] for the mathematical techniques used to derive results on the capacity of abstract neural networks). It is made possible by the highly divergent synaptic organisation of the MF→GrC synapse, providing expansion of the incoming signals into a higher-dimensional code, and by the feed-forward and feed-back inhibition provided by Golgi cells, whose role is to act as dynamic gain modulators keeping the overall activity in the GrC population more *sparse* than that in the afferent mossy fibres. Furthermore, molecular layer interneurons set a dynamic activation threshold on Purkinje cells to prevent them from being activated by subsets of the patterns they're intended to respond to [Marr, 1969; Tyrrell and Willshaw, 1992].

expansion recoding

As mentioned in subsection 1.1.1, Albus [1971] better formalised the theory within the mathematical framework of abstract neural networks by modeling the cerebellum as a perceptron. Importantly, he suggested learning to be mainly based on LTD instead of LTP, and to exists at the synapses that parallel fibres form with molecular layer interneurons as well as at those they form with Purkinje cells. Experimental confirmation of the former prediction [Ito and Kano, 1982] was fundamental in establishing the "Marr-Albus" (or "Marr-Albus-Ito") theory as one of the most popular within the research

community (the latter prediction, too, received experimental support, but only 30 years later by Jörntell and Ekerot [2002]). Even though other aspects of cerebellar physiology or function were studied in depth that do not fit in the Marr-Albus picture (see, for example, Schonewille et al. [2011] and De Zeeuw et al. [2011]), further experimental evidence and theoretical work have kept it relevant in the following decades [Tyrrell and Willshaw, 1992; Kitazawa et al., 1998; Schweighofer et al., 2001; Brunel et al., 2004; Ito, 2006].

1.3.4.2 *Oscillatory activity in the GCL, GoC synchrony and desynchronisation*

In vivo experiments have provided evidence of local field potential oscillation in the granule cell layer in the theta band (8 Hz, in rodents) [Hartmann and Bower, 1998; O'Connor et al., 2002; Roš et al., 2009; Dugué et al., 2009] and beta (13–18 Hz, in monkeys) band [Pellerin and Lamarre, 1997]. No consensus has been reached yet on the functional relevance of the oscillations; however, the presence of an oscillatory rhythm has been proposed to play a role in gating the information flow within the granule cell layer and the olivo-cerebellar loop [D'Angelo et al., 2009] and in establishing communication links with other brain areas, preparing the system for the execution of movements [Hartmann and Bower, 1998; Courtemanche and Lamarre, 2005]. Indeed, GCL oscillations have been shown to be phase locked with LFPs in somatosensory and motor cortices; behaviourally, they are associated with a "quiet wakefulness" state, rapidly disappearing upon initiation of a movement [Pellerin and Lamarre, 1997; Hartmann and Bower, 1998; Courtemanche et al., 2002].

Golgi cells are known to fire in phase with GCL oscillations [Dugué et al., 2009]. In fact, synchrony of GoC activity is thought to be the main dynamical substrate of the oscillations [Dugué et al., 2009; Vervaeke et al., 2010]. Synchronous GoC activity, itself, is thought to be enabled by electrical coupling between GoCs, although oscillatory synaptic input may play a role [Hartmann and Bower, 1998; Vos et al., 1999] and the MF→GrC→GoC feedback loop has been shown to give rise to synchrony in large scale models of the granule cell layer [Maex and De Schutter, 1998]. Indeed, electrical coupling enhances synchrony within neuronal populations by exerting an equalising influence on the membrane potentials of coupled cells, and gap junctions are known to play an important role in establishing synchronous activity in several brain areas (see for example Galarreta and Hestrin [1999]; Beierlein et al. [2000]; Deans et al. [2001]; Traub et al. [2001]). On the other hand, *in vivo* experiments have shown that gap junctions can also enable rapid desynchronisation following a temporally precise mossy fibre input [Vervaeke et al., 2010]; simulations of a biologically detailed network model have suggested

this phenomenon to rely upon features of the network connectivity structure such as the heterogeneity of synaptic strengths [Vervaeke et al., 2010]. Indeed, the interplay between electrical connectivity, cell heterogeneity and the stability of the synchronous activity state has been investigated with analytical methods in a fully connected network model [Ostojic et al., 2009], but a mechanistic explanation of the network determinants of Golgi cell synchrony and desynchronisation in a realistic network model is still lacking.

1.4 AIMS AND OUTLINE OF THIS STUDY

In this study, I investigated the relationship between network structure and function in the granule cell layer by focusing on the two themes presented above.

In chapter 3, I discuss an analysis of how spatial constraints — such as the extent of GrCs and GoC dendrites, the geometry of the GCL and the rules governing the spatial distribution of Golgi cells — affect network statistics within the GCL. Furthermore, I use this analysis to make a prediction — later verified by experiments — on the individual nature of electrical contacts between Golgi cells. chapter 5 shows, using an experimentally constrained spiking GCL network model I built, that the organisation of synaptic connectivity at the glomerulus governs the trade-off between information transmission and sparsification of incoming signals in the GCL. In the light of the Marr-Albus theories of cerebellar function, this suggests a link between the functional role of the network and the strong evolutionary conservation of the anatomy of the cerebellar GCL. Finally, in chapter 4 I investigate the robustness of the synchronous activity state of the gap junction-mediated Golgi cell network. Using simulations of a biophysically detailed model of the network, I dissected the contributions of different network-structural aspects to synchrony robustness; in particular, I showed that a local organisation of connectivity — an element omitted from earlier systematic investigations — is the key to enabling desynchronisation of network activity. Details about the methods and the tools I used and developed to carry out the present study are given in chapter 2, while chapter 6 contains a general discussion of its significance in the context of the existing research literature.

2 METHODS

2.1 STATISTICAL MODELLING OF THE GOLGI CELL GAP-JUNCTION NETWORK

2.1.1 *Exploratory second-order analysis of the spatial distribution of cerebellar Golgi cells*

Raw experimental data consisted in pictures, like the one in Figure 4, taken in immunostaining experiments from thin (30 µm) transverse sections of the cerebellum in 5 adult (P45) mice. Sections were about 5 µm thick. Each section was imaged separately; the resulting pictures were divided in groups of 5, which were then superimposed to give images like that in Figure 4. This procedure yielded about 30 sequential pictures for each animal. Golgi cells were immunolabelled for GlyT2 (green) and neurogranin (blue), and Purkinje cells were labelled with aldolase C (red) to mark the position of zebrin bands [Brochu et al., 1990]. The position and the subtype (GlyT2-positive, neurogranin-positive, or mixed) of all identifiable Golgi cells was manually recorded for each slice by Jean-Luc Dupont (INCI-CNRS, Strasbourg, France). The boundary of the granule cell layer in each cerebellar lobule was defined using the Purkinje cell line and the bottom of the granule cell layer itself as anatomical landmarks. In the following, I assume that the "stacked slices" represented in each picture are thin enough to be considered two-dimensional for the purpose of computing inter-cell distances. In other words, I will study the distribution of inter-cell distances as a purely two-dimensional problem in the transverse plane.

spatial point patterns

To analyse the correlations between the positions of individual Golgi cells I resorted to the statistical theory of *spatial point patterns* [Illian et al., 2008; Diggle, 2013]. I treated every section of the GCL visible in each slice (one for each cerebellar lobule in the picture) as an instance of a *second-order intensity-reweighted stationary* process [Baddeley et al., 2000]. This means that I allowed for the mean density of the cells to vary in space, while expecting the second-order structure to be isotropic and homogeneous after accounting for the spatial variations in intensity (see below).

In the following subsection, I recall briefly a few basic definitions from the theory of spatial point processes; for a detailed explanation, see for example Illian et al. [2008] and Møller and Waagepetersen [2007].

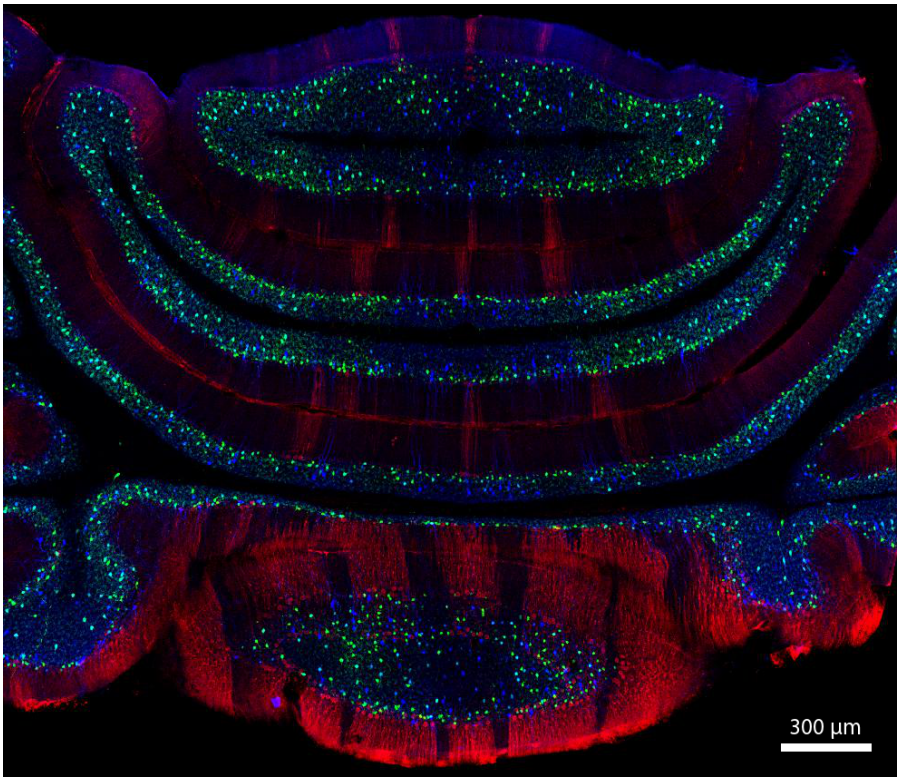


Figure 4: Immunostaining of cerebellar Golgi and Purkinje cells. Green: GlyT2. Blue: neurogranin. Red: aldolase C. Data kindly provided by Jean-Luc Dupont (INCI-CNRS, Strasbourg, France).

2.1.1.1 *Second-order intensity-reweighted stationary spatial point processes*

For a spatial point pattern defined in an unbounded d-dimensional volume, let $p_1(x) = \lambda(x)\,dx$ be the probability that there is a point in the infinitesimal disc (d-dimensional sphere) centred around x. We call $\lambda(x)$ the *intensity function* or *point density function* of the process. In the same way, let $p_2(x,y) = \rho(x,y)\,dx\,dy$ be the probability of finding one point in the infinitesimal volume centred at x and one point in the infinitesimal volume centred at y. Note that this implies that the conditional probability of finding a point in y given that there is a point in x is

$$p(y|x) = p_2(x,y)/p_1(x) = \lambda(x)^{-1}\rho(x,y)\,dy \tag{2.1}$$

Consider now the function

$$g(x,y) \coloneqq \frac{\rho(x,y)}{\lambda(x)\lambda(y)} \tag{2.2}$$

(with $g(x,y) = 0$ if $\lambda(x)\lambda(y) = 0$). If $g(x,y)$ depends only on the distance between x and y, that is $g(x,y) = g(|x-y|) =: g(r)$, then we call g the (inhomogeneous) *pair correlation function* of the process and we

say that the process is *second-order intensity-reweighted stationary* [Baddeley et al., 2000]. For this type of process, we can write Equation 2.1 as

$$p(y|x) = \lambda(y)g(|x-y|)\,dy \tag{2.3}$$

where the dependence on the position now only appears through $\lambda(y)$. Intuitively, we can interpret this by saying that the spatial inhomogeneities in a second-order intensity-reweighted stationary process are entirely "explained" by the spatial dependency of the intensity function $\lambda(x)$.

To get a better understanding of the meaning of g, let's consider a *stationary process*, where $\lambda(x) \equiv \lambda$ for all x. In this case, for any y Equation 2.3 becomes

stationary process

$$p(y|x) = \lambda g(|x-y|)\,dy \tag{2.4}$$

which means that $\lambda g(|x-y|)\,dy$ *is the probability of finding a further point in an infinitesimally small disc of area* dy *at a distance* $|x-y|$ *from any point* x. In the literature, $\lambda g(r)$ is also called *Palm intensity function* [Illian et al., 2008] or *o-ring statistic* [Wiegand and A. Moloney, 2004].

Palm intensity function

o-ring statistic

In the case where the positions of the points are independent,

$$\rho(x,y) = \lambda(x)\lambda(y)$$

and

$$g(r) = 1 \quad \forall r \geqslant 0$$

It is then clear why the pair correlation function is very useful in statistical testing of interactions between the points of an observed process.

2.1.1.2 *Details of statistical analysis*

Estimating $g(r)$ requires knowing or estimating $\lambda(x)$ for the process being analysed. Estimating both functions from a single instance of an inhomogeneous point pattern is problematic [Baddeley et al., 2000]. For each data sample, I therefore computed $\lambda(x)$ with the "leave-one-out" kernel smoother described in Baddeley et al. [2000] to control unwanted biases (but see subsubsection 2.1.1.3 for a discussion of remaining issues and of possible strategies to deal with them). As we assume isotropy, as an estimator for g we use that given in Illian et al. [2008] and implemented in the R package `spatstat`[1] [Baddeley and Turner, 2005] with the isotropic edge correction by Ripley [1979] (incidentally, this means applying methods for the analysis of stationary

1 http://www.spatstat.org/

("infinite") point patterns to a series of *finite* point patterns, as if the Golgi cell population extended indefinitely beyond the lobule boundaries). This yields an estimate $\hat{g}_i(r)$ for each lobule slice i. Following Illian et al. [2008], for each value of r we can then aggregate these estimates by taking the average of the $\hat{g}_i(r)$, weighted by the *isotropised set covariance* $\gamma(W_i)$ (or *rotation average of the set covariance*; see Ohser and Mücklich [2000]) of the observation window W_i defined by the anatomical boundaries of the ith lobule:

isotropised set covariance

$$\hat{g}(r) = \frac{\sum_i \gamma(W_i) g_i(r)}{\sum_i \gamma(W_i)} \tag{2.5}$$

COMPUTATION OF CONFIDENCE BANDS FOR $g(r)$ Following Diggle [2013] and Webster et al. [2006], I estimated the sampling variance of g by the *bootstrap method* [Efron and Tibshirani, 1994] using the `boot` R package. For the bootstrap procedure, I performed sampling with replacement from the pool of estimated values for g, and constructed a confidence interval for $\hat{g}$ with the BCa method [Efron, 1987].

bootstrap method

2.1.1.3 *Future extensions and refinements of the approach*

Even when using the leave-one-out kernel estimator mentioned above, estimating both $\lambda(r)$ and $g(r)$ for each data sample can lead to undesirable statistical biases, especially in the case of a process exhibiting spatial regularity [Baddeley et al., 2000; Illian et al., 2008]. This is a similar problem to that tackled by Webster et al. [2006], one of the few examples in the literature of a second-order analysis of replicated spatial patterns. In that work, a solution was found by registering all data in a common frame of reference through a rigid transformation and estimating λ on the pooled data, using this common global estimate in the calculation of the second order statistics for each data set. In our case, unfortunately, the variability of the shape of the GCL across slices, cerebellar lobules and animals prevents us from following the same approach. On the other hand, the large-scale organisation of the spatial density of Golgi cells is thought to be very consistent throughout the cerebellar cortex if described in the frame of reference provided by zebrin bands [Valera et al., 2012]. This suggests one way of improving the statistical soundness of the analysis above: define a set of "zebrin coordinates", giving the position of each cell as its location relative to the zebrin band it is located in, pool all data in this common coordinate system, perform the estimation for λ on the pooled data and then pull this estimate back into the original coordinate system associated to each data set with the appropriate geometrical transformation. We note that this "anatomical atlas" approach [Wager et al., 2004] can be more technically challenging than the rigid registration procedure used by Webster

et al., as the transformation to be operated between the original and the zebrin coordinates is in general nonlinear [Dupont et al., 2015]. The increased accuracy afforded by such an improved statistical analysis, tough, would allow for a meaningful comparison with a better null model than a completely random spatial process, taking for example into account the finite size of Golgi cell somata [Geurts et al., 2001].

2.1.2 *Estimation of network-structural characteristics from paired recording data*

I detail here the structural Golgi network model used in section 3.2. In an effort to simplify as much as possible what we know from anatomical and functional measurements while still capturing the salient traits of the network, I will initially disregard the effect of the spatial correlations discussed in subsection 2.1.1, considering them separately only at the end of this Subsection.

2.1.2.1 *Connection probability model*

In Vervaeke et al. [2010], the spatial dependence of the cell connection probability is modelled in the following way: if two cells n and m are located at an intersomatic distance r from each other, then the probability that they are connected is

$$p_c^V(r) = \left[-17.45 + \frac{18.36}{\exp\left(\frac{r-267\,\mu m}{39\,\mu m}\right)+1} \right] \cdot \theta(\tilde{r}-r) \tag{2.6}$$

where

$$\tilde{r} = 39\,\mu m \cdot \ln\left(\frac{18.36}{17.45}-1\right) + 267\,\mu m \tag{2.7}$$

and $\theta(x)$ is the Heaviside step function. Unfortunately, this functional form is not very amenable to analytical manipulation, so I fitted the data published in the paper with a better behaved Fermi function-like expression:

$$\begin{aligned} p_c(r) &\doteq p\left(n \text{ and } m \text{ connected } \middle| \|\boldsymbol{r}_n - \boldsymbol{r}_m\| = r\right) \\ &= \frac{a}{e^{(r-r_0)/\Delta}+1} \end{aligned} \tag{2.8}$$

The results of the fit are illustrated in Figure 5, and the corresponding values for the parameters are given in Table 1.

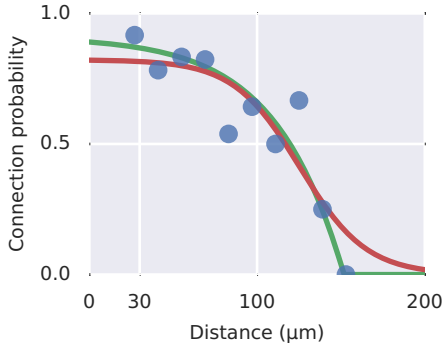


Figure 5: Spatial dependence of pairwise connection probability in Golgi cells. Blue: experimental data from Vervaeke et al. [2010]. Green: $p_C^V(r)$, Equation 2.6. Red: $p_c(r)$, Equation 2.8. Note that no experimental data is available for cell pairs located at a distance smaller than 20 μm.

Parameter	Value
α	0.8223
r_0	125.5 μm
Δ	19.78 μm
L	80 μm
λ	4607 mm^{-3}

Table 1: Parameter values for the the structural Golgi cell network model. See main text, Equation 2.8 and Equation 2.9.

2.1.2.2 *Derivation of the degree distribution for the structural Golgi network model.*

Consider a population of cells within an infinitely extended layer of thickness L (the granule cell layer). To a first approximation, assume that the cell positions are independently distributed with constant spatial density λ. This can be seen as the limit for $N, V \to \infty$ of a finite population of N cells uniformly distributed in a cylindrical region of volume V with radius R and height L, taken in such a way that the spatial cell density λ remains constant:

$$\frac{N}{V} = \frac{N}{\pi R^2 L} \equiv \lambda \tag{2.9}$$

Experimental estimates for L and λ are given in Vervaeke et al. [2010], and are reported in Table 1.

Let $R(N, V; r)$ be the function that, for the finite-size network, gives the number of cells at a distance r from a given point in space, averaged over all possible points taken with uniform probability within the cylindrical volume V. $R(N, V; r)$ is a *radial distribution function*

radial distribution function

(RDF), as defined in the cristallography literature [Kodama et al., 2006]. As we assume that the cell positions are independent of each other, though, averaging over all possible points within the cylinder is equivalent to averaging over all possible cell positions. The number of cells at a distance r from a given cell - be it cell number n - is then on average

$$\frac{N-1}{N} R(N, V; r)$$

and the probability of having a specific cell - be it cell number m - at a distance r from cell n is

$$p\left(\|\boldsymbol{r}_n - \boldsymbol{r}_m\| = r\right) = \frac{1}{N-1}\frac{N-1}{N} R(N, V; r) = \frac{R(N, V; r)}{N}$$

We can then use Equation 2.8 to work out the probability that cell n and cell m are connected, averaged over their positions:

$$\begin{aligned} p(n \leftrightarrow m) &= \int_r p\left(n \leftrightarrow m \;\middle|\; \|\boldsymbol{r}_n - \boldsymbol{r}_m\| = r\right) p\left(\|\boldsymbol{r}_n - \boldsymbol{r}_m\| = r\right) dr \\ &= \frac{1}{N}\int_r p_c(r) R(N, V; r)\, dr \\ &\doteq \frac{\beta}{N} \end{aligned}$$

But this doesn't depend on our choice of n and m, so we can say that β/N *is the probability for any given pair of cells to be connected*. Consequently, the probability for any given cell to be connected to k other cells and not to the remaining $N-1-k$ is

$$\left(\frac{\beta}{N}\right)^k \left(1 - \frac{\beta}{N}\right)^{N-1-k} \quad .$$

But since there are $\binom{N-1}{k}$ ways of choosing those k cells, the probability for any cell to be connected to k other cells (or, in graph-theoretical terminology, the probability for any cell to have *degree* k in the GJ-mediated network) is

$$p(\deg = k) = \binom{N-1}{k} \left(\frac{\beta}{N}\right)^k \left(1 - \frac{\beta}{N}\right)^{N-1-k} \tag{2.10}$$

Now, we don't attempt to explicitly solve β as a function of our model's parameters, but if we define

$$\alpha \doteq \lim_{N,V\to\infty} \beta = \lim_{N,V\to\infty} \int_r p_c(r) R(N, V; r)\, dr$$

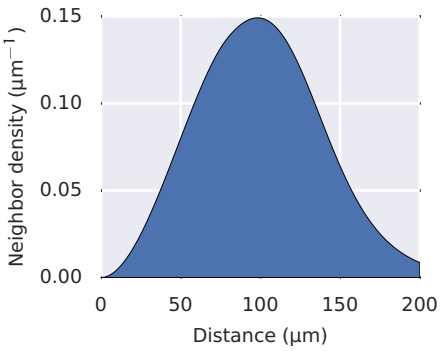


Figure 6: Density of neighbours in the structural Golgi cell network model, i.e. $p_c(r)R_{\text{sheet}}(\lambda, L; r)$ for the experimentally constrained values of the parameters given in Table 1.

where the limit is taken at constant λ, then

$$\alpha = \int_r p_c(r)R_{\text{sheet}}(\lambda, L; r)\, dr$$

where $R_{\text{sheet}}(\lambda, L; r) \doteq \lim_{N,V\to\infty} R(N, V; r)$ is the RDF of an infinite sheet of thickness L and density λ. We can also say that, for the typical cell of the network, $p_c(r)R_{\text{sheet}}(\lambda, L; r)$ is the density of connected neighbours located at a distance r (Figure 6). Under some relatively mild assumptions, α can be calculated analytically (see below, subsubsection 2.1.2.3) to give

$$\alpha = a\lambda(L\pi r_0^2)\left[1 + \frac{\pi^2}{3}\left(\frac{\Delta}{r_0}\right)^2 - \frac{1}{6}\left(\frac{L}{r_0}\right)^2\right] \tag{2.11}$$

and in our case, for the estimates given above for a, r_0 and Δ we have

$$\alpha \simeq 15 \tag{2.12}$$

We can then calculate the degree distribution for an infinite network by noting that (2.10) is a binomial distribution with β/N as the *probability of success* parameter and $N-1$ as the *repetitions* parameter. It follows that, in the infinite limit, $\alpha = \lim_{N,V\to\infty} \beta$ is the *mean degree* and the degree distribution is Poisson:

$$p(\deg = k) = \frac{e^{-\alpha}\alpha^k}{k!}$$

(As a general remark, any spatial network with a connection probability that falls off sufficiently fast with distance and with uncorrelated edges is going to have a Poisson degree distribution in the infinite limit.)

It's worth noting that, apart from the change in the definition of p_c from Equation 2.6 to Equation 2.8, these results follow directly from the same assumptions that have been made for the connectivity model in Vervaeke et al. [2010].

Interestingly, all the parameters in my approach have a natural interpretation in terms of the model: r_0 is a "base" connection radius, a is the pairwise connection probability between cells that are within a distance of r_0, Δ is the width of the region where the connection probability trails off from a to 0, and the two squared factors in Equation 2.11 tell us how much the "softness" of the connection radius and the finite thickness of the layer, respectively, modify the total number of connections from the case where $p_c(r)$ is a step function and the layer is two-dimensional.

2.1.2.3 *Calculation of the average degree* α

As shown above, the average degree for an infinite network can be expressed as

$$\begin{aligned}\alpha &= \int_r p_c(r) R_{\text{sheet}}(\lambda, L; r)\, dr \\ &= \int_0^{+\infty} a R_{\text{sheet}}(\lambda, L; r) \frac{1}{e^{(r-r_0)/\Delta}+1}\, dr\end{aligned} \tag{2.13}$$

where $R_{\text{sheet}}(\lambda, L; r)$ is the radial distribution function of an infinite sheet of density λ and thickness L. The last integral in Equation 2.13 is — unsurprisingly — reminiscent of some expressions found in the Fermi gas theory, and we are going to borrow from there the following integration technique.

Consider, first of all, the functional form of $f(r) \doteq 1/(e^{(r-r_0)/\Delta}+1)$. For $\Delta \to 0$, f tends to a step function, and if $\Delta \ll r_0$ (which would correspond, in the Fermi gas theory, to the "low temperature" $T \ll T_F$ condition) we know that $f'(r) \neq 0$ only where $r \simeq r_0$.

Now, for any integrable function $g = g(r)$, if we define

$$G(r) \doteq \int_0^r g(r')\, dr'$$

we have, keeping in mind that $G(0) = 0$ and $f(+\infty) = 0$,

$$\begin{aligned}I &\doteq \int_0^{+\infty} g(r) f(r)\, dr \\ &= \left[G(r) f(r)\right]_0^{+\infty} - \int_0^{+\infty} G(r) \frac{df}{dr}\, dr \\ &= -\int_0^{+\infty} G(r) \frac{df}{dr}\, dr\end{aligned} \tag{2.14}$$

We now expand $G(r)$ in Taylor series around r_0:

$$G(r) = G(r_0) + (r - r_0)G'(r_0) + \frac{1}{2}(r - r_0)^2 G''(r_0) + \dots \tag{2.15}$$

Since $f'(r) \simeq 0$ only for $r \simeq r_0$, only the first terms of this expansion contribute to the integral in (2.14), but for the time being we keep the full expression. Substituting (2.15) into (2.14) we get

$$I = \sum_{n=0}^{\infty} \Delta^n I_n \left.\frac{d^{(n)}G}{dr^{(n)}}\right|_{r_0} \tag{2.16}$$

where

$$\begin{aligned}
I_n &= \frac{1}{n!}\int_0^{+\infty} \left(\frac{r - r_0}{\Delta}\right)^n f'(r)\,dr \\
&= \frac{1}{n!}\int_0^{+\infty} \left(\frac{r - r_0}{\Delta}\right)^n \frac{e^{\frac{r-r_0}{\Delta}}}{\left(e^{\frac{r-r_0}{\Delta}} + 1\right)^2}\frac{dr}{\Delta} \\
&= \frac{1}{n!}\int_{-r_0/\Delta}^{+\infty} x^n \frac{e^x}{(e^x + 1)^2}\,dx \\
&= \frac{1}{n!}\int_{-r_0/\Delta}^{+\infty} \frac{x^n}{4\cosh^2(x/2)}\,dx \\
&\simeq \frac{1}{n!}\int_{-\infty}^{+\infty} \frac{x^n}{4\cosh^2(x/2)}\,dx
\end{aligned} \tag{2.17}$$

where we have performed the change of variables $x \doteq (r - r_0)/\Delta$, and extended the lower limit of integration to $-\infty$ by using the fact that, if $\Delta \ll r_0$, only the region where $x \simeq 0$ contributes to the integral. The integrand in (2.17) is even when n is even and odd when n is odd, so we can immediately see that $I_n = 0$ when n is odd. On the other hand, when n is even, observing that

$$\int \frac{1}{4\cosh^2(x/2)}\,dx = -\frac{1}{1 + e^x} + C$$

where C is an arbitrary constant, and integrating by parts, we get

$$\begin{aligned}
I_n &\simeq \frac{1}{n!}\int_{-\infty}^{+\infty} \frac{x^n}{4\cosh^2(x/2)}\,dx \\
&= \frac{2}{n!}\int_0^{+\infty} \frac{x^n}{4\cosh^2(x/2)}\,dx \\
&= -\frac{2}{n!}\left[\frac{x^n}{1 + e^x}\right]_0^{+\infty} + \frac{2}{(n-1)!}\int_0^{+\infty} \frac{x^{n-1}}{1 + e^x}\,dx \\
&= \frac{2}{(n-1)!}\int_0^{+\infty} \frac{x^{n-1}}{1 + e^x}\,dx
\end{aligned}$$

but remembering the following integral representation of the Riemann Zeta function (see the DLMF at http://dlmf.nist.gov/25.5.3)

$$\zeta(s) = \frac{1}{\left(1-2^{1-s}\right)\Gamma(s)} \int_0^{+\infty} \frac{x^{s-1}}{1+e^x}\, dx$$

and substituting $\Gamma(n) = (n-1)!$ as n is an integer, we get

$$I_n \simeq \begin{cases} 2\left(1-2^{1-n}\right)\zeta(n) & n \text{ even} \\ 0 & n \text{ odd} \end{cases}$$

In particular, $I_0 = 1$ and $I_2 = \pi^2/6$. As we stated above, it's now possible (again by using $\Delta \ll r_0$) to approximate the Taylor series in Equation 2.16 by arresting it at order $n = 2$. We then get

$$I = \int_0^{+\infty} g(r)f(r)\, dr \simeq G(r_0) + \Delta^2 \frac{\pi^2}{6} G''(r_0) \tag{2.18}$$

Finally, to perform the integral in Equation 2.13 we set

$$g(r) = aR_{\text{sheet}}(\lambda, L; r)$$

An expression for $R_{\text{sheet}}(\lambda, L; r)$ is given in Kodama et al. [2006], and we report it here:

$$R_{\text{sheet}}(\lambda, L; r) = \begin{cases} \lambda \cdot \left(4\pi r^2 - \frac{2\pi}{L} r^3\right) & r \leqslant L \\ \lambda \cdot 2\pi L r & r \geqslant L \end{cases}$$

This implies that

$$\begin{aligned} G(r) &= \int_0^r aR_{\text{sheet}}(\lambda, L; r')\, dr' \\ &= \begin{cases} a\lambda\left(\frac{4}{3}\pi r^3 - \frac{\pi}{2L} r^4\right) & r \leqslant L \\ a\lambda\left[\left(\frac{4}{3}\pi - \frac{\pi}{2}\right)L^3 + \pi L r^2 - \pi L^3\right] & r \geqslant L \end{cases} \\ &= \begin{cases} a\lambda\left(\frac{4}{3}\pi r^3 - \frac{\pi}{2L} r^4\right) & r \leqslant L \\ a\lambda\left(\pi L r^2 - \frac{\pi}{6} L^3\right) & r \geqslant L \end{cases} \end{aligned}$$

so that

$$G(r_0) = \begin{cases} a\lambda\left(\frac{4}{3}\pi r_0^3 - \frac{\pi}{2L} r_0^4\right) & r_0 \leqslant L \\ a\lambda\left(\pi L r_0^2 - \frac{\pi}{6} L^3\right) & r_0 \geqslant L \end{cases}$$

and

$$G''(r_0) = \begin{cases} a\lambda\left(8\pi r_0 - \frac{6\pi}{L}r_0^2\right) & r_0 \leqslant L \\ a\lambda 2\pi L & r_0 \geqslant L \end{cases}$$

Note that $G(r)$ is smooth up to the second derivative, but $\frac{d^3G}{dr^3}$ is discontinuous in L. This, strictly speaking, means that we don't expect the Taylor expansion to hold for all r, but we are still fine as long as the region that dominates the integral in Equation 2.14 does not contain L.

By plugging the values for $G(r_0)$ and $G''(r_0)$ in Equation 2.18, we get for $r_0 > L$

$$\begin{aligned} \alpha &= a\lambda\left[\pi L r_0^2 - \frac{\pi}{6}L^3 + \Delta^2\frac{\pi^3}{3}L\right] \\ &= a\lambda(L\pi r_0^2)\left[1 + \frac{\pi^2}{3}\left(\frac{\Delta}{r_0}\right)^2 - \frac{1}{6}\left(\frac{L}{r_0}\right)^2\right] \end{aligned} \tag{2.19}$$

A NOTE ON THE APPROXIMATIONS. For the sake of clarity, we sum up the approximations and assumptions we had to made to reach an analytical result for the average degree α:

- we used Equation 2.8 instead of Equation 2.6;
- we assumed that the shoulder of the Fermi function was steep enough to put us in the "low temperature" regime;
- we assumed that the Taylor expansion (2.15) is valid for all r in the region that contributes to the integral in Equation 2.14.

The first assumption means approximating the existing model for $p_c(r)$ with a more convenient one, while the second and the third are choices we had to make to perform analytically the integral in Equation 2.14. Overall, they seem like reasonable things to do, as confirmed by the fact that we can get very similar results (less than 0.3% difference on the estimate for α) by numerical integration, which sidesteps the formal difficulties.

2.1.3 *Effect of small scale regularity of cell positions on average degree*

In subsection 2.1.1 and subsection 3.2.1 I discuss how the spatial organisation of Golgi cells shows regularity below the 30 µm scale when analysed in two dimensions in the transverse plane. In three dimensions, if we assume the positions of the cells to follow an intensity-reweighted stationary process, we can expect the same regularity to

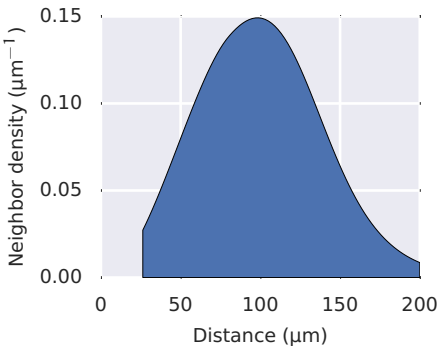


Figure 7: Density of neighbours in the structural Golgi cell network model, modified to take into account small scale spatial regularity below a "hard-core distance" $r_1 = 25\,\mu m$. The average degree corresponds to the area under the curve. Compare with Figure 6.

hold. The simplest possible way of modelling its effect on the connectivity statistics of the 3D network is to approximate the spatial process describing Golgi cell positioning as a three-dimensional *hard-core process*, that is to consider the cells to be placed completely at random as long as their reciprocal distances are always greater than the hard-core distance r_1, which we set at $25\,\mu m$. For simplicity, we imagine the process to extend in the whole three-dimensional space, but we still consider the infinite layer of thickness L as the volume of interest within which we compute the connectivity properties of the network. Under these assumptions, we can make a back-of-the-envelope estimation of the effect of the small-scale regularity in Golgi cell positions by modifying the RDF defined above as follows:

hard-core process

$$R_{sheet}(\lambda, L; r) \longrightarrow R_{sheet}(\lambda, L; r)\theta(r - r_1)$$

where $\theta(r)$ is the Heaviside step function.

The effect of this modification on the average degree of the network can be computed analytically, and turns out to be very small ($< 2\%$). The reason for this is illustrated in Figure 7: for the typical cell, most neighbours are located at distances much larger than $25\,\mu m$, so excluding the $r < 25\,\mu m$ region doesn't significantly alter their number.

2.2 MATHEMATICAL MODELLING OF CELLS AND SYNAPTIC MECHANISMS IN THE GCL

To investigate the MF→GrC and GoC→GoC networks in the granule cell layer I have relied upon numerical simulation of neuronal activity. In this section I will discuss in detail the mathematical models of cell and synaptic dynamics I built and used. All single cell or synaptic models were developed in LEMS/NeuroML [Cannon et al., 2014],

and are available on the Open Source Brain [www.opensourcebrain.org; Gleeson et al., 2012]. Some of the descriptions in this section are adapted from those I wrote for Billings et al. [2014].

2.2.1 *Granule cell*

I built a simple conductance-based integrate-and-fire model of the cerebellar granule cell, based on data published in Rothman et al. [2009], Seja et al. [2012] and Schwartz et al. [2012] and conforming to the `IafRefCell`[2] component type definition in the NeuroML2 schema [Cannon et al., 2014]. During a simulation, the membrane voltage evolves according to

$$-C_m\frac{dV}{dt} = G_m\cdot(V-E_m) + G_{GABAR}\cdot(V-E_{GABAR}) + \\ +\sum_{i=1}^{d} G_{AMPAR}(i;t)\cdot(V-E_{AMPAR}) + b(V)\cdot\sum_{i=1}^{d} G_{NMDAR}(i;t)\cdot(V-E_{NMDAR}) \tag{2.20}$$

where C_m is the membrane capacitance, E_m the reversal potential of the membrane leak conductance G_m, G_{GABAR} and E_{GABAR} the GABAR-mediated conductance and reversal potential, $G_{AMPAR}(i;t)$ and $G_{NMDAR}(i;t)$ the AMPAR and NMDAR-mediated conductances induced by synaptic stimulation from the mossy fibre innervating the i^{th} dendrite, E_{AMPAR} and E_{NMDAR} the AMPAR and NMDAR reversal potentials, and $b(V)$ the NMDAR Mg^{2+} unblock function defined above. Upon V reaching the threshold value V_t a spike is emitted, and V is then clamped to the reset potential V_r for a refractory interval τ_r. All synaptic conductances are discussed in detail below, while values and experimental provenance for all parameters in the model are reported in Table 2.

To give an idea of the input-output transformation properties of the GrC model when coupled to the synaptic conductance models described below, Figure 8 reports its rate-coded input-output curve and output autocorrelation as computed from simulations with a varying number of Poisson synaptic inputs.

2.2.2 *Golgi cell*

Some of the analyses I did required simulating the gap junction network formed by cerebellar Golgi cells [Dugué et al., 2009; Vervaeke et al., 2010]. To achieve this, it is important to model cell morphology to a sufficient level of detail to allow for realistic filtering of synaptic

2 http://www.neuroml.org/NeuroML2CoreTypes/Cells.html#iafRefCell

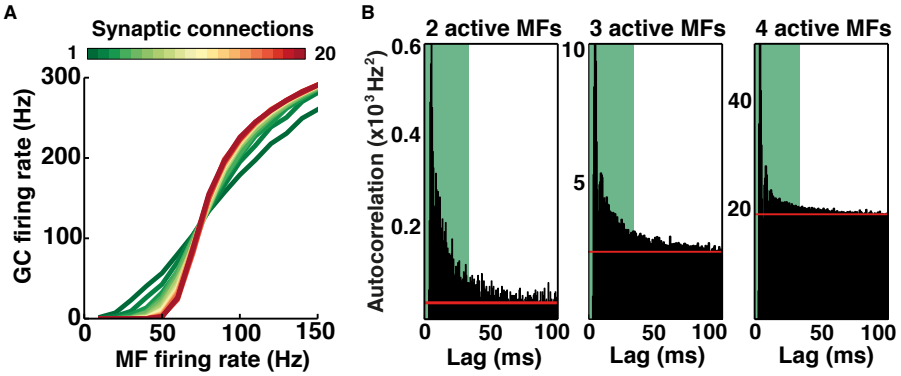


Figure 8: (*Adapted from Billings et al. [2014]*) Input-output properties of an integrate-and-fire GrC model. *A:* rate-coded input-output curve for the granule cell (GrC) model with a variable number of mossy fibre (MF) synaptic inputs per GrC. *B:* autocorrelation of the spiking GC model output (with 4 synaptic inputs) for 2, 3 or 4 active inputs (the case of 1 active input is not shown as the resulting firing rate is very low). *Active* inputs have a mean rate of 80 Hz, *inactive* inputs have a mean rate of 10 Hz. Red line: constant value of the autocorrelation of a Poisson point process with the same mean rate. Green area: 30 ms time window.

currents by the dendritic tree and spatial distribution of gap junctions on the cell [Vervaeke et al., 2012]. This is precisely what the models published in Vervaeke et al. [2012] have been designed for, building upon previous work on modelling the properties of Golgi cell electrophysiology [Solinas et al., 2007] by including highly detailed morphological information obtained through immunofluorescence and standard light microscopy. Unfortunately, the thousands of compartments that make up these models (up from 5 compartments in Solinas et al. [2007]) also make them very demanding from a computational standpoint, and impractical for running large numbers of network simulations. To alleviate this issue, I built and published[3] a *reduced-morphology* version of one of them (called `210710_C1` in the codebase associated with Vervaeke et al. [2012]), using the very minimum number of segments that allowed for realistic dendritic filtering and gap junction placement. Remarkably, the model affords all this in just 10 compartments - the same order of complexity as the original model in Solinas et al. [2007].

The behaviour of the reduced morphology model matches well that of `210710_C1` under several metrics (Figures 9 and 10). In particular, the effect of dendritic filtering is well approximated for subthreshold activity (Figure 9b) as well as for spiking properties (Figures 9d, 10).

3 http://opensourcebrain.org/projects/golgi-cell-reduced-morphology-eugenio-piasini

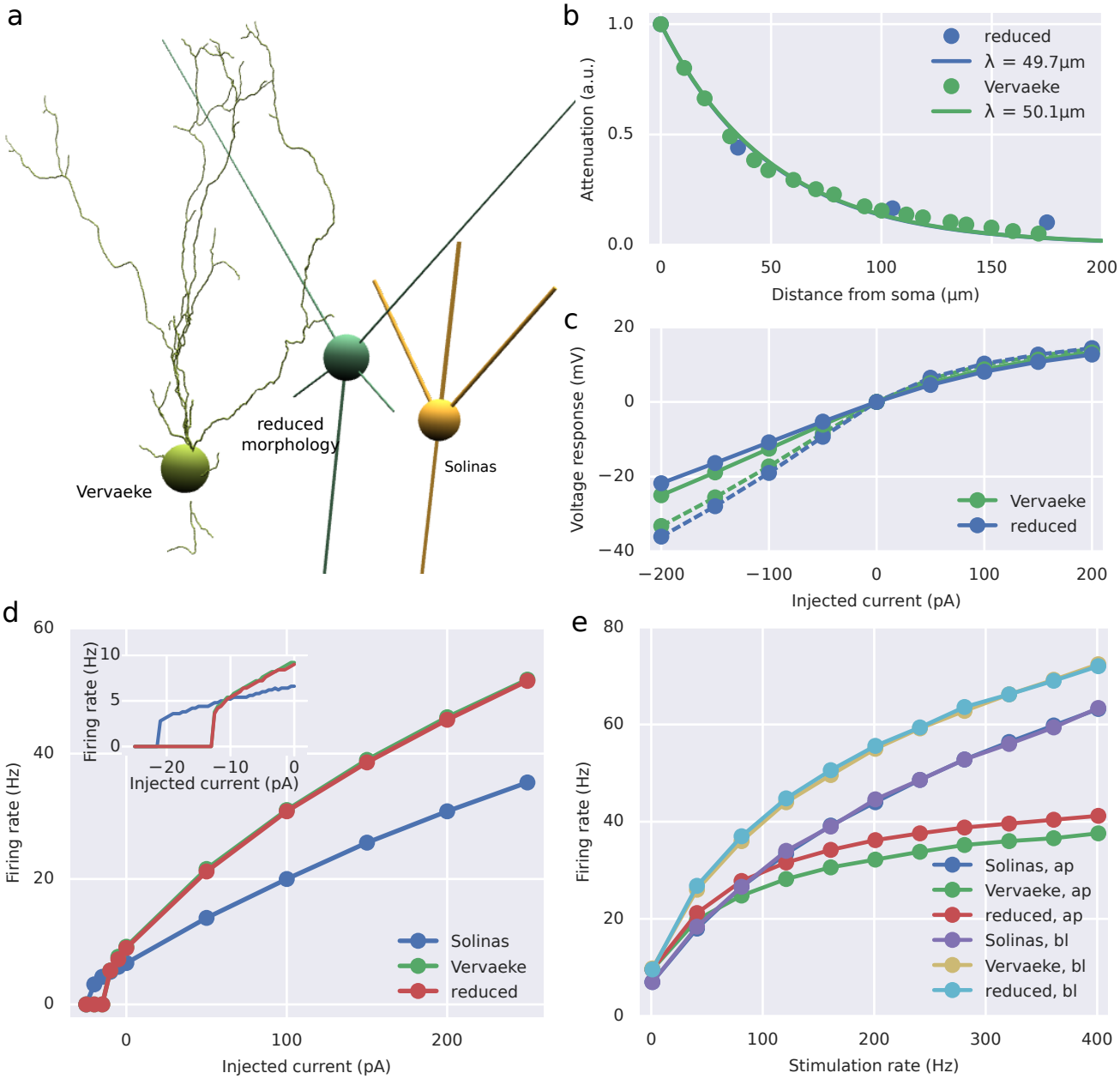


Figure 9: Comparison between the reduced morphology Golgi cell model I realised, the `210710_C1` model in Vervaeke et al. [2012] and the original five-compartment model in Solinas et al. [2007]. *a*: Model morphologies visualised on the Open Source Brain. *b*: Dendritic attenuation of synaptic inputs, measured by simulating a synaptic input on a dendrite at 200 µm from the soma and taking the ratio of the peak current response at the soma and at various points located at different distances along the dendrite. Circles: simulation results. Solid lines: exponential fits for the space dependency of the attenuation, giving space constants λ. Note that the fits are almost identical. *c*: Voltage response to current injection, after blocking Na channels, for cell models embedded in a network with connectivity given by the rules in Vervaeke et al. [2012] (solid lines) or Vervaeke et al. [2010] (dashed lines). *d*: F-I relationships. Inset: detail of the transition from quiescence to firing. *e*: input-output relationships for apical (*ap*) or basolateral (*bl*) Poisson synaptic stimulation.

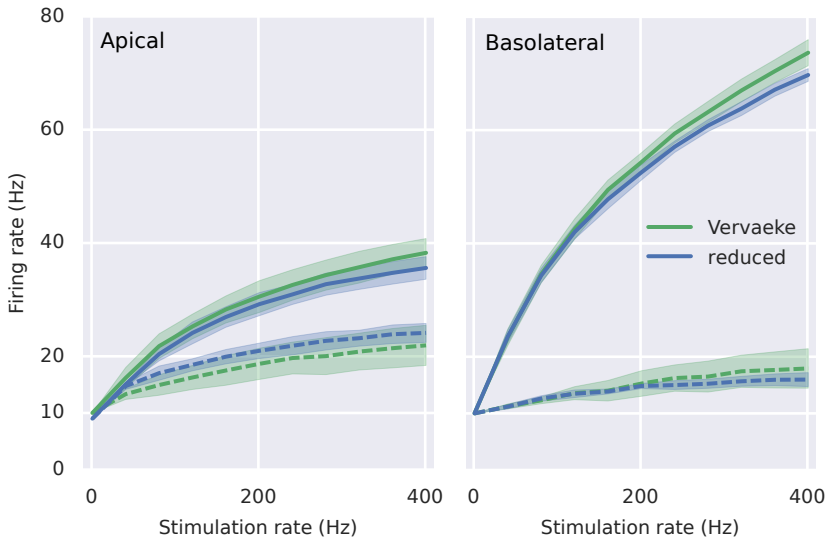


Figure 10: Rate-based input-output comparison of Golgi cell models embedded in a network. Following the analysis presented in figure 3E in Vervaeke et al. [2012], I simulated a network of 45 cells, 15 of which are stimulated with 30 independent Poisson synaptic inputs each, randomly distributed on the apical (left panel) or basolateral (right panel) dendrites. Solid lines: cells undergoing direct stimulation. Dashed lines: all other cells. The connectivity rules for the network were the same as in Vervaeke et al. [2012]. Shadowed region indicates variability within the network ($\pm$ one standard deviation).

2.2.3 *Mossy fibre to granule cell synapse*

The mossy fibre to granule cell synaptic model was built from published experimental data Sargent et al. [2005]; Rothman et al. [2009]; Schwartz et al. [2012]. The total synaptic conductance G, as a function of time and of the history of synaptic events, is defined as the sum of an AMPAR-mediated component G_{AMPA} and an NMDAR-mediated component G_{NMDA}, defined below.

DERIVATION OF SYNAPTIC MODEL PARAMETERS FROM EXPERIMENTAL DATA Estimates for the parameters for the synaptic waveform shape and plasticity mechanisms were recomputed from the experimental data published in Rothman et al. [2009] with a particle swarm optimization algorithm [Deb and Padhye, 2010] implemented using the `inspyred` framework[4]. Parameters for the Mg^{2+} block mechanism were set to the values published in Schwartz et al. [2012]. An example of the fit of the model to the data is shown in Figure 11.

4 https://pypi.python.org/pypi/inspyred

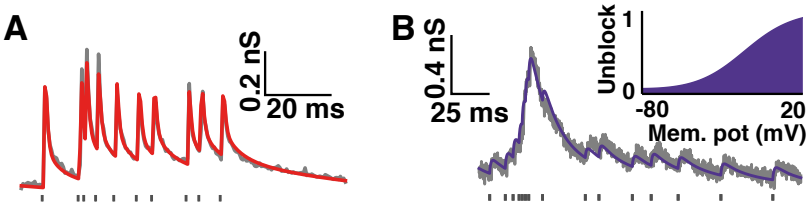


Figure 11: *(Adapted from Billings et al. [2014])* Models of the AMPAR- and NMDAR- mediated components of the MF to GrC synapse. *A:* Fit of the short-term plasticity model (red) of the AMPAR component to a 100 Hz synaptic conductance train obtained from experiments (gray). *B:* Same as for A but for the NMDAR component and an 80 Hz synaptic conductance train. Inset: voltage dependence of NMDAR conductance.

2.2.3.1 *AMPAR-mediated synaptic waveform*

The total AMPAR conductance component arising from a single MF spike through direct quantal release and glutamate spillover is given by

$$G_{AMPA} = p_{AMPA,d} \cdot g_{AMPA,d} + p_{AMPA,s} \cdot g_{AMPA,s}$$

where

$$g_{AMPA,d} = \sum_{i=1}^{2} a_i^{(AMPA,d)} \left(e^{-t/\rho^{(AMPA,d)}} - e^{-t/\delta_i^{(AMPA,d)}} \right)$$

$$g_{AMPA,s} = \sum_{i=1}^{3} a_i^{(AMPA,s)} \left(e^{-t/\rho^{(AMPA,s)}} - e^{-t/\delta_i^{(AMPA,s)}} \right)$$

and $a_i^{(AMPA,d)}$, $\rho^{(AMPA,d)}$, $\delta_i^{(AMPA,d)}$ and $a_i^{(AMPA,s)}$,$\rho^{(AMPA,s)}$,$\delta_i^{(AMPA,s)}$ indicate amplitudes, rise times and decay times for the direct and glutamate spillover-only components of the AMPAR synaptic waveform, respectively. $p_{AMPA,d}$ and $p_{AMPA,s}$ are the independent short-term depression scaling factors for the direct and spillover components, respectively, and are both defined as per the standard NeuroML2 `TsodyksMarkramDepMechanism` component type, as described below.

2.2.3.2 *NMDAR-mediated synaptic waveform*

The total NMDAR conductance component arising from a single MF spike in absence of Mg^{2+} block is given by

$$G_{NMDA} = p_{NMDA} \cdot \sum_{i=1}^{2} a_i^{(NMDA)} \left(e^{-t/\rho^{(NMDA)}} - e^{-t/\delta_i^{(NMDA)}} \right)$$

where $a_i^{(NMDA)}$, $\rho^{(NMDA)}$ and $\delta_i^{(NMDA)}$ indicate conductance amplitudes, rise time and decay times for the two components of the waveform. The p_{NMDA} scaling factor implements short term depression and facil-

itation of the NMDAR component and is defined as per the standard NeuroML2 `TsodyksMarkramDepFacMechanism` component type, as described below.

2.2.3.3 *Mg^{2+} block mechanism*

A Mg^{2+} block mechanism for the NMDAR-mediated synaptic conductance component was modeled as a custom LEMS component type by defining an unblock function $b(V)$ with the Woodhull formalism [Rothman and Silver, 2014]:

$$b(V) = \frac{C_1 e^{\delta_{bind}\theta V} + C_2 e^{-\delta_{perm}\theta V}}{C_1 e^{\delta_{bind}\theta V} + C_2 e^{-\delta_{perm}\theta V} + [Mg^{2+}]_{out} e^{-\delta_{bind}\theta V}}$$

where V is the postsynaptic membrane potential, and $\theta = zF/RT$, with z as the Mg^{2+} ionic charge, F the Faraday constant, R the ideal gas constant, and T the absolute temperature.

2.2.3.4 *Synaptic plasticity mechanisms*

The short term plasticity models in the simulations make use of NeuroML2's standard `TsodyksMarkramDepMechanism`[5] and `TsodyksMarkramDepFacMechanism`[6] component types [Cannon et al., 2014]. These, in turn, are defined as the models formulated in, respectively, Tsodyks and Markram [1997] and Tsodyks et al. [1998], simplified to ignore inactivation. For a synaptic conductance g a plasticity mechanism defines a factor p which multiplies the contribution of a single synaptic event before it gets added to g.

As defined in the NeuroML2 documentation, `TsodyksMarkramDepFacMechanism` is parametrized by an *initial release probability* r, a *depression recovery time* Δ and a *potentiation recovery time* Π. p is defined as

$$p(t) = U(t) \cdot R(t)$$

where $U(t)$ and $R(t)$ are the internal dynamical variables of the model, which are initialized as

$$U(0) = r$$
$$R(0) = 1$$

5 http://www.neuroml.org/NeuroML2CoreTypes/Synapses.html#tsodyksMarkramDepMechanism

6 http://www.neuroml.org/NeuroML2CoreTypes/Synapses.html#tsodyksMarkramDepFacMechanism

Following a synaptic event, U and R get updated according to

$$U \to U + r \cdot (1 - U)$$
$$R \to R \cdot (1 - U)$$

while, in absence of synaptic events, they decay exponentially to their initial values:

$$\frac{dU}{dt} = -\frac{U - r}{\Pi}$$
$$\frac{dR}{dt} = -\frac{R - 1}{\Delta}$$

Finally, `TsodyksMarkramDepMechanism` is a simpler version of `TsodyksMarkramDepFacMechanism` where U is fixed to its initial value.

2.2.4 *Granule cell inhibition*

2.2.4.1 *Golgi cell to granule cell synapse*

Building on experimental data from Golgi-granule paired recording experiments [Ward, 2012], I developed a simple model for both the direct and the spillover-only Golgi-granule connections. In the experiments, Golgi cells were stimulated to fire at a range of rates between 8 and 50 Hz while granule cells were held at a fixed membrane potential; the evoked current pulses in granule cells were then recorded and converted to conductance by dividing for the value of the holding potential. Isolated evoked pulses were averaged to provide an estimate of the synaptic response to a single spike.

Similarly to the excitatory synapse models discussed above, the GoC→GrC synapse models are composed of a multiexponential kernel, representing the response of the synapse to a single spike isolated in time, scaled by a synaptic plasticity term. More precisely, indicating the conductance of the direct connection with $G_{(GABA,d)}$ and that of the spillover-only connection with $G_{(GABA,s)}$,

$$G_{(GABA,d)} = p_{GABA,d} \cdot \sum_{i=1}^{2} a_i^{(GABA,d)} \left(e^{-t/\rho^{(GABA,d)}} - e^{-t/\delta_i^{(GABA,d)}} \right)$$
$$G_{(GABA,s)} = p_{GABA,s} \cdot \sum_{i=1}^{2} a_i^{(GABA,s)} \left(e^{-t/\rho^{(GABA,s)}} - e^{-t/\delta_i^{(GABA,s)}} \right)$$

where the a, ρ and δ parameters indicate respectively amplitudes, rise times and decay times of the components of each model. $p_{GABA,d}$ and $p_{GABA,s}$ are plasticity scaling factors; they are defined as per the standard NeuroML2 `TsodyksMarkramDepMechanism` component type, discussed above, and the parameters controlling their evolution were

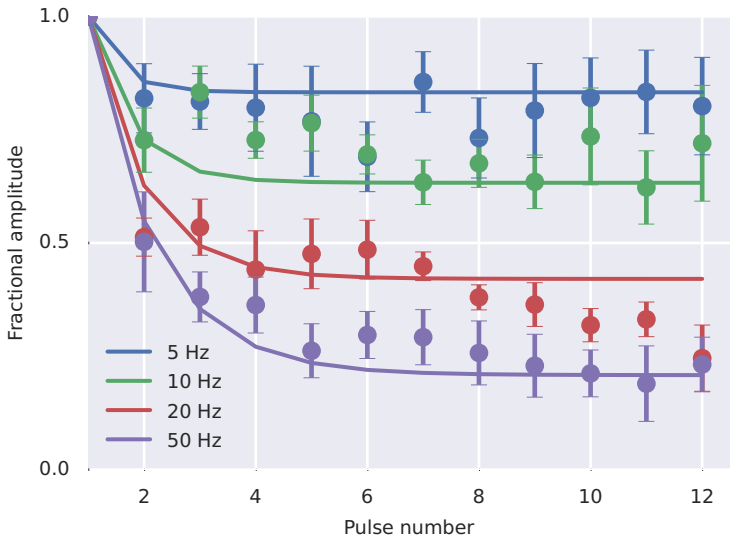


Figure 12: Short term plasticity for the Golgi cell to granule cell synapse: peak conductance amplitude of successive pulses in trains at 5, 10, 20 and 50 Hz. Dots and error bars (SEM): experimental data from Ward [2012]. Solid lines: NeuroML model.

estimated by fitting the short term depression model by Tsodyks and Markram [1997] to the relative change in maximum amplitude of successive pulses in trains delivered at different rates (see figure 12). The other parameters of the model were obtained by fixing p and fitting the shape of the "base waveform" obtained by recording the response of the synapse to a single spike. All fits were performed using a particle swarm optimisation technique [Deb and Padhye, 2010].

2.2.5 *Mossy fibre to Golgi cell synapse*

The MF→GoC synapse model was the same as in Vervaeke et al. [2010], built from experimental data published in Kanichay and Silver [2008]. At a given time t after a synaptic event, the current injected into the cell is

$$I_{MF\to GoC}(t) = G_{MF\to GoC}(t) \cdot \left(V(t) - E_{MF\to GoC}\right)$$

where the base conductance waveform is

$$\begin{aligned} G_{MF\to GoC}(t) = & a_1^{(MF\to GoC)} \left(e^{-t/\delta_1^{(MF\to GoC)}} - e^{-t/\rho^{(MF\to GoC)}} \right) \\ & + a_2^{(MF\to GoC)} \left(e^{-t/\delta_2^{(MF\to GoC)}} - e^{-t/\rho^{(MF\to GoC)}} \right) \end{aligned}$$

2.2.6 *Granule cell (parallel fibre) to Golgi cell synapse*

The PF→GoC synapse model was the same as in Vervaeke et al. [2010], built from experimental data published in Dieudonné [1998]. At a given time t after a synaptic event, the current injected into the cell is

$$I_{PF\to GoC}(t) = G_{PF\to GoC}(t) \cdot \left(V(t) - E_{PF\to GoC}\right)$$

where the base conductance waveform is

$$G_{PF\to GoC}(t) = a^{(PF\to GoC)} \left(e^{-t/\delta^{(PF\to GoC)}} - e^{-t/\rho^{(PF\to GoC)}}\right)$$

2.2.6.1 *A simplified model for steady-state inhibition in the granule cell*

Following the general picture given in section 1.3, we can give a simple approximate description of the steady-state inhibition received by granule cells in the rate-coding regime. In absence of external stimuli, we can consider granule cells as being subjected to a baseline inhibition given by the sum of the tonic inhibitory conductance and the time-averaged 8 Hz phasic component from 3 direct and 1 spillover-only connections [Ward, 2012]. When an external stimulus is introduced, a network-activity-dependent inhibitory conductance is added, proportional to the variation of the time-averaged phasic component with the firing rate of Golgi cells. This results in modelling the inhibition received by granule cells as a constant conductance whose amplitude depends on the level of activity of the network.

By running numerical simulations of the Golgi to granule synapse models discussed above, I computed how the time-averaged inhibitory conductance amplitude changes as a function of the Golgi cell firing rate, for a direct as well as for a spillover-only connection. I then calculated the total average GABAR-mediated conductance on a granule cell (shown in Figure 13) by summing the contributions of 3 direct and 1 spillover-only inhibitory connections to the experimental estimate published in Rothman et al. [2009] for the tonic component of the inhibition.

2.2.7 *Model parameters*

The parameters for the single cell and synaptic models defined in this section can be found in Table 2. Note that, because of how the STP mechanisms are defined, the maximum amplitude of a conductance pulse isolated in time will be of the order of $r^{(AMPA)} \cdot a^{(AMPA)}$ for AMPA and $b(V) \cdot r^{(NMDA)} \cdot a^{(NMDA)}$ for NMDA. This, in practice, means a maximum peak amplitude of 630 pS for both AMPA [Sargent et al., 2005], and unblocked NMDA [Schwartz et al., 2012].

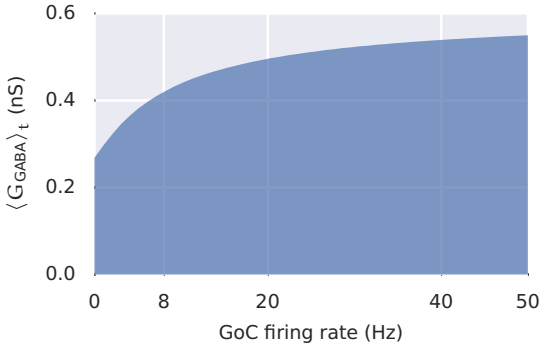


Figure 13: Total time-averaged GABAR-mediated conductance on a granule cell innervated by four Golgi cells (3 direct plus 1 spillover-only connections) as a function of Golgi cell firing rate.

Parameter	Value	Source
E_m	−79.9 mV	Schwartz et al. [2012]
G_m	1.06 nS	— " —
C_m	3.22 pF	— " —
V_t	−40 mV	— " —
V_r	−63 mV	— " —
τ_r	2 ms	— " —
E_{GABAR}	−79.1 mV	Seja et al. [2012]
G_{GABAR}	0.438 nS	Rothman et al. [2009]
E_{AMPAR}	0 mV	— " —
$a_1^{(AMPA,d)}$	3.724 nS	— " —
$a_2^{(AMPA,d)}$	0.3033 nS	— " —
$\rho^{(AMPA,d)}$	0.3274 ms	— " —
$\delta_1^{(AMPA,d)}$	0.3351 ms	— " —
$\delta_2^{(AMPA,d)}$	1.651 ms	— " —
$r^{(AMPA,d)}$	0.1249	— " —
$\Delta^{(AMPA,d)}$	131 ms	— " —
$a_1^{(AMPA,s)}$	0.2487 nS	— " —
$a_2^{(AMPA,s)}$	0.2799 nS	— " —
$a_3^{(AMPA,s)}$	0.1268 nS	— " —
$\rho^{(AMPA,s)}$	0.5548 ms	— " —
$\delta_1^{(AMPA,s)}$	0.4 ms	— " —
$\delta_2^{(AMPA,s)}$	4.899 ms	— " —
$\delta_3^{(AMPA,s)}$	43.1 ms	— " —

Table 2: Parameters used for the models described in this section, together with the provenance of their values or of the experimental data used to estimate them.

Continued from previous page

Parameter	Value	Source
$r^{(AMPA,s)}$	0.2792	— " —
$\Delta^{(AMPA,s)}$	14.85 ms	— " —
E_{NMDAR}	0 mV	Rothman et al. [2009]
$a_1^{(NMDA)}$	17 nS	Schwartz et al. [2012]
$a_2^{(NMDA)}$	2.645 nS	— " —
$\rho^{(NMDA)}$	0.8647 ms	Rothman et al. [2009]
$\delta_1^{(NMDA)}$	13.52 ms	— " —
$\delta_2^{(NMDA)}$	121.9 ms	— " —
$r^{(NMDA)}$	0.0322	— " —
$\Delta^{(NMDA)}$	236.1 ms	— " —
$\Pi^{(NMDA)}$	6.394 ms	— " —
z	2	Schwartz et al. [2012]
T	308.15 K	— " —
$[Mg^{2+}]_{out}$	1 mM	— " —
δ_{bind}	0.35	— " —
δ_{perm}	0.53	— " —
C_1	2.07 mM	— " —
C_2	0.015 mM	— " —
$a_1^{(GABA,d)}$	0.5793 nS	Ward [2012]
$a_2^{(GABA,d)}$	0.1209 nS	— " —
$\rho^{(GABA,d)}$	0.3278 ms	— " —
$\delta_1^{(GABA,d)}$	3.400 ms	— " —
$\delta_2^{(GABA,d)}$	94.94 ms	— " —
$r^{(GABA,d)}$	0.5125	— " —
$\Delta^{(GABA,d)}$	157.9 ms	— " —
$a_1^{(GABA,s)}$	0.003 540 nS	— " —
$a_2^{(GABA,s)}$	0.075 65 nS	— " —
$\rho^{(GABA,s)}$	0.8388 ms	— " —
$\delta_1^{(GABA,s)}$	10.01 ms	— " —
$\delta_2^{(GABA,s)}$	126.6 ms	— " —
$r^{(GABA,s)}$	0.5125	— " —
$\Delta^{(GABA,s)}$	157.9 ms	— " —
$E_{MF\to GoC}$	0 mV	Kanichay and Silver [2008]
$a_1^{(MF\to GoC)}$	0.7 nS	— " —

Table 2: Parameters used for the models described in this section, together with the provenance of their values or of the experimental data used to estimate them.

Continued from previous page

Parameter	Value	Source
$a_2^{(MF\to GoC)}$	0.2 nS	— " —
$\rho^{(MF\to GoC)}$	0.1 ms	— " —
$\delta_1^{(MF\to GoC)}$	0.7 ms	— " —
$\delta_2^{(MF\to GoC)}$	2.5 ms	— " —
$E_{PF\to GoC}$	0 mV	Dieudonné [1998]
$a^{(PF\to GoC)}$	0.67 nS	— " —
$\rho^{(PF\to GoC)}$	0.1 ms	— " —
$\delta^{(PF\to GoC)}$	1.06 ms	— " —

Table 2: Parameters used for the models described in this section, together with the provenance of their values or of the experimental data used to estimate them.

2.2.8 *Granule cell population modelling*

With the goal of including cell heterogeneity in future simulations and generalising the description given in subsection 2.2.1, I analysed the distribution of the five main intrinsic electrophysiological parameters in the integrate-and-fire description (C_m, G_m, E_m, V_t, and V_r) across a population of granule cells. The raw experimental data, kindly provided by Daniel Ward, consisted of an estimate of all these parameters for each cell in a set of 38. The data, shown in Table 3, was estimated by measuring the voltage response to somatic injection of current steps.

Cell #	C_m (pF)	G_m (nS)	E_m (mV)	V_t (mV)	V_r (mV)
1	3.6	1.0	-50	-28	-52
2	3.2	0.48	-75	-22	-53
3	3.5	0.76	-90	-43	-73
4	3.9	0.69	-60	-35	-63
5	4.1	0.62	-70	-30	-55
6	2.6	0.64	-75	-34	-58
7	2.2	0.32	-78	-46	-75
8	4.4	0.88	-60	-29	-59
9	2.4	1.1	-60	-34	-63
10	5.1	1.5	-66	-27	-55

Table 3: Experimental parameter estimates for a population of granule cells. Note that in the original dataset some cells had multiple estimates for their electrophysiological parameters; in those cases, the value presented here is the average of the estimates. C_m: membrane capacitance; G_m: leak conductance; E_m: leak reversal potential; V_t: spike threshold potential; V_r: reset potential.

Continued from previous page

Cell #	C_m (pF)	G_m (nS)	E_m (mV)	V_t (mV)	V_r (mV)
11	4.2	0.80	-90	-46	-73
12	3.9	0.75	-55	-31	-57
13	3.5	0.32	-75	-42	-70
14	3.3	0.65	-60	-32	-60
15	3.9	0.95	-52	-30	-60
16	3.5	0.31	-72	-27	-58
17	2.8	0.82	-65	-32	-61
18	2.8	0.59	-55	-42	-81
19	3.2	0.54	-70	-47	-68
20	3.2	0.47	-56	-32	-62
21	3.0	0.74	-72	-37	-71
22	3.4	1.4	-50	-26	-49
23	2.9	0.65	-58	-26	-54
24	3.9	0.48	-80	-42	-81
25	4.4	1.1	-50	-26	-48
26	3.4	0.79	-66	-42	-66
27	3.9	1.0	-60	-30	-60
28	2.6	0.82	-50	-36	-63
29	3.4	0.83	-90	-42	-66
30	4.4	1.3	-50	-27	-55
31	3.6	0.53	-80	-30	-53
32	4.1	0.68	-90	-37	-68
33	3.2	0.68	-56	-31	-58
34	2.3	0.41	-65	-28	-55
35	3.8	1.2	-50	-37	-58
36	5.2	1.1	-76	-41	-66
37	5.2	0.60	-80	-31	-59
38	3.4	0.77	-67	-29	-53

Table 3: Experimental parameter estimates for a population of granule cells. Note that in the original dataset some cells had multiple estimates for their electrophysiological parameters; in those cases, the value presented here is the average of the estimates. C_m: membrane capacitance; G_m: leak conductance; E_m: leak reversal potential; V_t: spike threshold potential; V_r: reset potential.

To find a representative but concise description of the distribution that could be easily used as a generative model for the population, I decided to fit a Gaussian mixture model [Barber, 2012] to the data. I performed model selection on the number of components of the mixture by applying the Bayesian Information Criterion [Schwarz, 1978] to the results of an expectation-maximisation fit of models with up to 20 components. Conveniently, the preferred number of components was 1. This result was corroborated by fitting a Dirichlet process Gaussian mixture model using the variational approach described in Blei and Jordan [2006], for 100 different values of the hyperparameter of the model logarithmically spaced in the interval $[0.001, 10]$; this procedure confirmed the goodness of our choice by consistently selecting one component only. In conclusion, my granule cell population model (detailed in section 5.3) is a simple Gaussian distribution in the five-dimensional space of model parameters.

2.3 ANATOMICALLY CONSTRAINED LOCAL GRANULE CELL LAYER MODEL

To analyse the information transmission properties of the granule cell layer, I built a biologically detailed network model of the local granule cell layer network, using the GrC model detailed in subsection 2.2.1, the MF-to-GrC synapse model of subsection 2.2.3 and the connectivity structure developed in the context of the work published in Billings et al. [2014], from which the detailed description below was taken. The code I used for generating instances of the network's structure according to the connectivity rules below is available online[7].

2.3.1 *Modelling of the network structure from anatomical data*

Model GrCs were placed at random in a sphere at the measured anatomical density of $1.9 \cdot 10^6 \mathrm{mm}^{-3}$[Billings et al., 2014] within a central subfield of MFs to minimise edge effects (Figure 14; see below). Each model GC made synaptic connections to randomly selected MF rosettes, with the constraint that the dendritic length should be close to 15 µm. In practice, they rarely exceeded 20 µm (Figure 15B), as observed experimentally [Eccles et al., 1967; Palkovits et al., 1972]. Since the maximum distance between two GrCs that could share the same MF input was 40 µm (see also section 3.1), the largest ball of tissue that could be expected to have independent inputs was 80 µm in diameter, which is comparable to the thickness of the GC layer in rodents. The anatomically constrained model of this "local GCL network" contained 176 MF synaptic rosettes and 509 GrCs. For an average of four dendrites per GC [Eccles et al., 1967], a single MF rosette made syn-

7 https://github.com/GuyBillings/GCL-Mathematica-toolbox

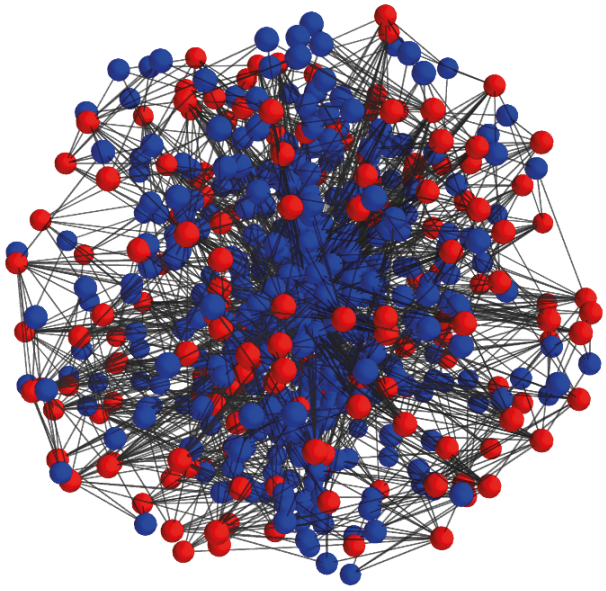


Figure 14: Three-dimensional view of the local granule cell layer network model. Red balls are MF rosettes (glomeruli), blue balls are GrCs.

aptic connections with an average of 12 different GrCs (Figure 15C), consistently with estimates of 15–20 in monkey and cat [Eccles et al., 1967].

Given the measured anatomical ratio of 1:2.9 between glomeruli and GrCs [Billings et al., 2014] and the number of MF synaptic connections per GrC, then the mean number of synaptic connections per MF rosette is fixed. However, in practice the MF rosette degree distribution (i.e. the distribution of the number of synaptic connections per rosette) depends upon the shape of the 3D volume used to construct the model. This occurs because GrC soma within the defined volume of the model can connect to glomeruli located outside by virtue of the finite length of GrC dendrites. This changes the effective glomerular to GrC ratio and also causes the glomeruli outside the GrC cell body volume to be sampled less frequently than those within it, affecting the connectivity statistics of the local network. Confining the whole network, including all dendrites, within a sphere minimized these effects (Figure 14 and 15). Another problem with constructing networks was that randomly connecting MF rosettes and GrCs resulted in some GrCs sampling the same rosette twice or the same MF more than once via differing glomeruli (Figure 15A). For a network having only 4 dendrites this effect was small (~4%) and consistent with experimental observations, but this problem grew as the number of synaptic connections per GrC was increased, causing the number of independent MF inputs to depend upon the number of dendrites. These problems were overcame by imposing a constraint on the

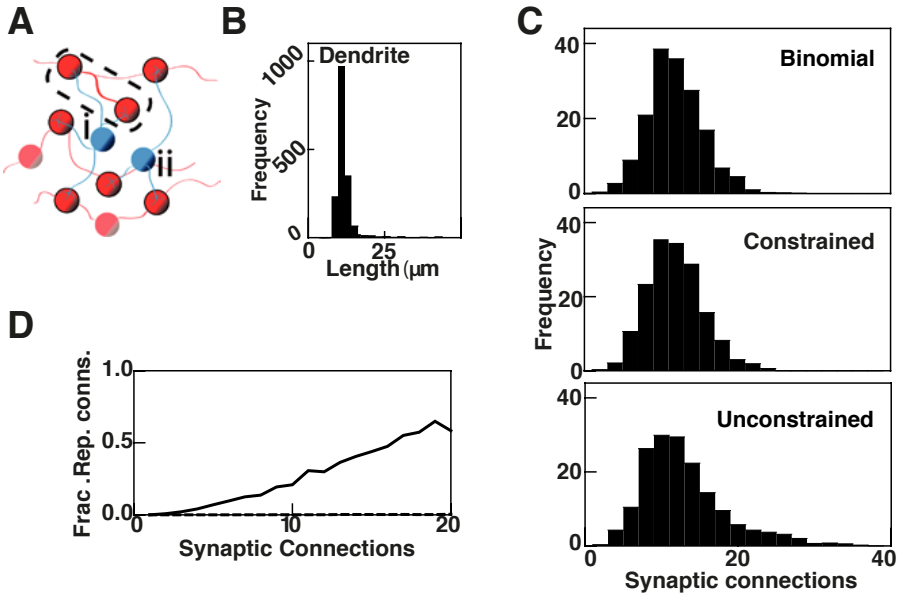


Figure 15: *(adapted from Billings et al. [2014]. See ibid. for details)* Connectivity constraints in the anatomical model produce a binomial distribution of the number of synaptic connections per mossy fibre rosette and biologically plausible dendrite lengths. A: schematic diagram illustrating how multiple synaptic rosettes (glomeruli; red balls) are associated with each mossy fibers (MF; red lines) and distributed in 3D space. These rosettes/glomeruli may in principle be on the same fiber (eg. enclosed with dashed box) or on independent fibers. GrCs (blue balls) connect to a subset of these (highlighted balls) within range of their dendrites (blue lines, known to be ~ 15 µm long on average). For each GrC, there are several ways in which connectivity can give rise to non-uniform numbers of connections with distinct MFs including the case where two dendrites contact a single MF synaptic rosette. Here I illustrate i) A GrC (blue) with d = 4 synaptic inputs, each made onto a different dendrite, but connects to a single MF via two separate synaptic rosettes and ii) Given some connectivity constraint (e.g. dendrite length limit) GrCs can fail to connect all dendrites to d glomeruli. B: Distribution of GC dendrite length in the local GCL network model. C: binomial distribution with a mean of 12, which matches the expected number of synaptic connections per MF in the anatomical model (top), distribution of number of synaptic connections per MF synaptic rosette in anatomical model with constrained random connections (middle), distribution of number of synaptic connections per MF in anatomical model with unconstrained random connections (bottom). D: Fraction of the GrCs that do not have d independent inputs as a function of the number of synaptic connections per GrC in a model with random connectivity (solid black line) and the same fraction in the model having connectivity constraints so as to avoid these effects (dashed black line).

GrC dendrite lengths which, while allowed in principle to assume any value, are as close as possible to 15 µm, and by explicitly forcing GrCs to have all dendrites connecting to independent MFs (Figure 15D). This allowed us to generate a spherical network model with a MF rosette degree distribution that was binomial in shape, rather than one with a tail (Figure 15C) as well as a dendrite length distribution having mean 17 µm and mode 15 µm, Figure 15B. Indeed, most dendrite lengths were smaller than 20 µm for 4 dendrites per GrC (Figure 15B), as found experimentally [Eccles et al., 1967; Palkovits et al., 1972].

2.3.2 *Numerical simulations of network activity*

In the simulations, input signals were represented in each MF by the rate of an independent Poisson spike train, and each spike triggered excitatory synaptic responses in connected GrCs. GrCs received a fixed number d of MF synaptic inputs (depending on the network configuration) in the form of trains of AMPAR- and NMDAR-mediated conductances, generated with the models described in subsection 2.2.3 (Figure 16A and 16B). Furthermore, excitatory synaptic conductances were scaled by a factor of $4/d$ to keep the total time-averaged conductance at the same value across network configurations. GrCs integrated the synaptic inputs, producing trains of EPSPs and spikes (Figure 16C).

Each input pattern was generated by randomly selecting a subset of the 176 MF inputs and designating them to be active (black lines in barcode, Figure 16D), while the remainder were inactive. During the simulation, active MFs fired random Poisson trains with a mean rate of 80 Hz and inactive MFs fired at a "background" rate of 10 Hz (red raster plot in Figure 16D), reflecting the properties of real MF rate-coded inputs [Arenz et al., 2008; Van Kan et al., 1993]. This resulted in individual GrCs receiving both high- and low-frequency trains of synaptic input conductances (e.g., top two and bottom two traces in Figure 16A, respectively).

The output spiking of the 509 GrCs in the network (blue spheres Figure 16E) was recorded over a 30 ms time window, corresponding to the synaptic integration time of GCs (Figure 8; Schwartz et al. [2012]). For each MF input pattern, the number of spikes was calculated for each GrC, and this was expressed as a vector for the GrC population (blue barcode in Figure 16F). The code I wrote to run the simulations and manage the associated data is available online[8].

8 https://github.com/epiasini/BillingsEtAl2014_GCL_SpikingSimulationAndAnalysis

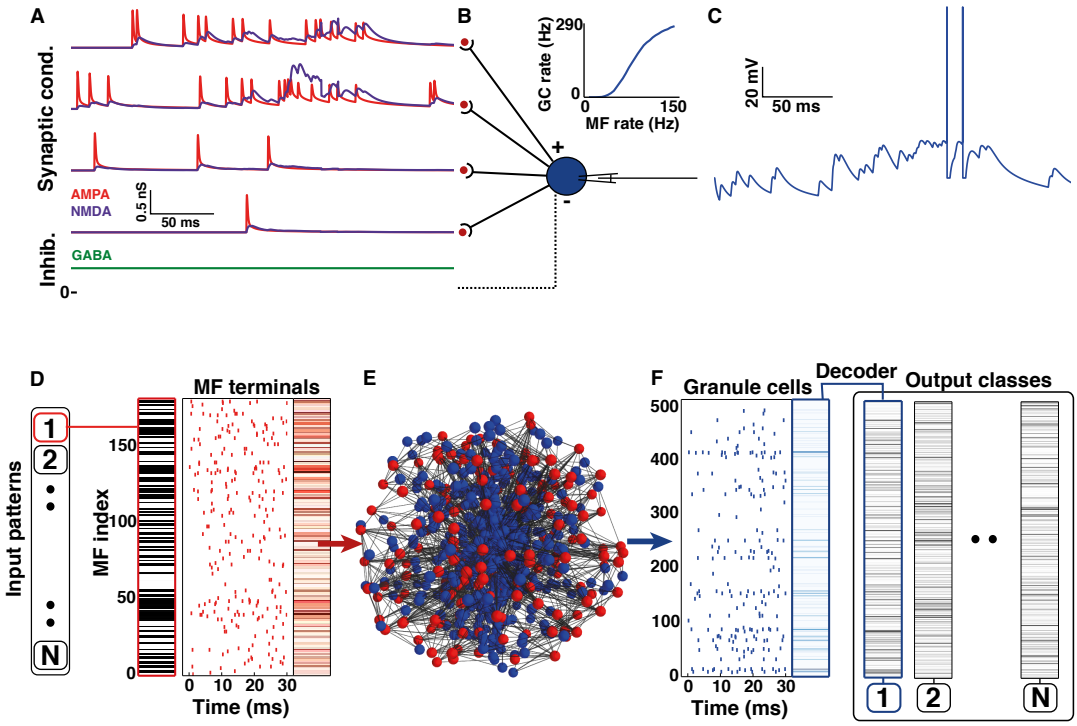


Figure 16: (*From Billings et al. [2014]*) Construction and analysis of an experimentally constrained spiking model of the local granule cell layer network incorporating synaptic mechanisms and tonic inhibition. *A:* Excitatory AMPAR (red) and NMDAR (purple) synaptic conductances for four independent mossy fibre (MF) inputs injected into a model granule cell (GrC). Top two traces: active MFs with excitatory conductance driven by independent Poisson spike trains firing at 80 Hz. Lower two traces: inactive MF firing at 10 Hz. Bottom trace: tonic inhibitory GABAA R conductance (green). *B:* Model GrC with action potential firing rate-coded input-output relationship (above) for four synaptic inputs. *C:* Membrane potential of model Grc during synaptic input in A. *D:* A binary stimulus pattern was randomly selected from a set of N patterns (black active and white inactive on barcode). A Poisson spike train was generated for each MF input (80 Hz active, 10 Hz inactive; red raster plot), thereby setting the timing of synaptic conductances (as in A). Red barcode indicates spike counts for the given realization of the spike trains. *E:* 3D view of the anatomically constrained local GCL network model with 176 MFs in red and 509 Grcs in blue. *F:* Raster plot of Grc firing activity in response to the input. Blue barcode indicates Grc spike count vector (measured over a 30 ms window), which was assigned to one of N output classes (black bar codes) defined using the k-means algorithm on a separate data set.

2.4 ANALYSIS OF NETWORK MODEL ACTIVITY

2.4.1 *Numerical estimation of MI for large feedforward spiking networks*

To quantify information processing in the GCL network, I used sets of N = 1024 MF input patterns, which was the largest number achievable with the computational resources available: the network simulations and analyses presented here required about 2 million 2 GHz compute hours. Since direct calculation of mutual information between input and output spike trains is computationally intractable due to the huge number of possible inputs and outputs (a manifestation of the *curse of dimensionality* [Bishop, 2006; Quian Quiroga and Panzeri, 2009]), I reduced the dimensionality of the output space by cascading the network with a classifier (called the *decoder*; see Figure 16F), which labelled the output spike count vectors as belonging to one of N classes. The network and decoder constituted a communication channel that mapped N input patterns to N output classes (Figure 16D–F), for which I calculated the mutual information (MI) assuming a flat prior over inputs. This fixed the maximum MI achievable as $\log_2 1024 = 10$ bits.

curse of dimensionality

The partitioning of the network outputs in N classes was obtained by performing a Voronoi (nearest-neighbour) tessellation of the output space. N was the smallest number of output classes that allowed full recovery of information, given that I assumed a uniform prior over the inputs. The N seed points for the tessellation were the centroids of the clusters obtained by running the k-means algorithm [Lloyd, 1982] as implemented in `scikit-learn` [Pedregosa et al., 2011] on a training dataset of 30 repetitions per pattern — this size of the training dataset was found to be sufficient to give satisfying performance for the decoder; see Figure 17A. To avoid being trapped in local minima, the algorithm was re-initialized 10 times using `k_means++` [Arthur and Vassilvitskii, 2007], and the best clustering solution was used in the computation of the centroids. Training data was not reused for the computation of MI. Information, which had an upper bound equal to the input entropy ($\log_2 1024 = 10$ bits for 1024 patterns), was calculated with the `pyentropy` package [Ince et al., 2009]. Undersampling bias in the MI estimate [Panzeri et al., 2007] was accounted for with the quadratic extrapolation procedure [Strong et al., 1998] (Figure 17, right). Other bias correction methods [Nemenman et al., 2002; Panzeri and Treves, 1996] were considered, but did not to provide a significantly better performance (data not shown).

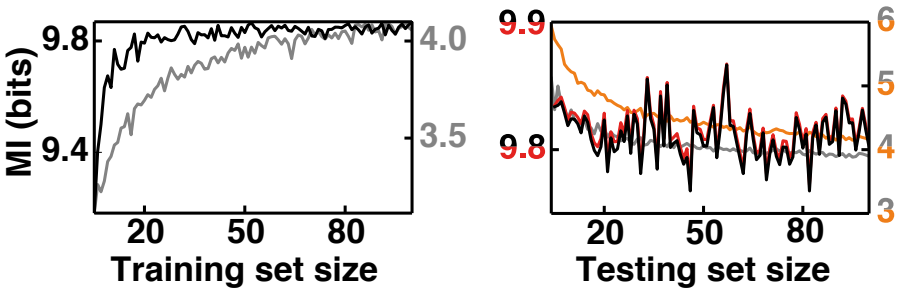


Figure 17: (*Adapted from Billings et al. [2014]*) Properties of the MI estimation procedure for the GCL model. *Left:* MI as a function of the number of repetitions per pattern used to train the decoder, for 1024 MF patterns, 4 synaptic connections and $p(\text{MF}) = 0.1$ (gray) or $p(\text{MF}) = 0.5$ (black). Note the different scales. The network performs poorly for $p(\text{MF}) = 0.1$, transmitting only a fraction of the information present in the input, so this is a particularly stringent test for the performance of the decoder. *Right:* MI as a function of the number of repetitions per pattern used to estimate it, showing the magnitude of undersampling bias and the effect of our chosen bias correction technique. Grey and orange: 4 synaptic connections, $p(\text{MF}) = 0.1$, quadratic extrapolation and no undersampling bias correction, respectively. Black and red: 4 synaptic connections, $p(\text{MF}) = 0.5$, quadratic extrapolation and no undersampling bias correction, respectively.

2.4.2 *Numerical estimation of population sparseness*

After considering a number of sparseness measures [Olshausen and Field, 2004; Hurley and Rickard, 2009], population sparseness was computed using the definition introduced by Vinje and Gallant [2000]:

$$s = \left[C - \frac{\left(\sum_{i=1}^{C} r_i\right)^2}{\sum_{i=1}^{C} r_i^2} \right] \cdot (C-1)^{-1}$$

where C is the number of cells and r_i is the spike count of cell i. This measure is a normalised version of the metric proposed by Treves and Rolls [1991], and it is closely related to the quantity defined by Hoyer [2004].

2.4.3 *Efficient generation of statistically independent samples of network activity*

An initial 150 ms transient was discarded from all simulations to allow the system to go from resting to steady state for the pattern being simulated. Independent recordings of the response to a pattern were extracted from a single simulation by slicing it in 30 ms-long time frames and discarding every other frame, to allow the system state to decorrelate between the frames we kept for analysis. This procedure

allowed me to avoid re-initialising the system from resting conditions for every repetition of the trial being simulated. 30 ms was chosen as decorrelation time as it corresponds to the characteristic time of the autocorrelation function of the GrC spiking output (Figure 8B).

2.5 SIMULATION MANAGEMENT AND NEUROINFORMATICS TOOLS

2.5.1 *LEMS and NeuroML2*

NeuroML [Gleeson et al., 2010; Vella et al., 2014; Cannon et al., 2014] is an XML-based declarative model description language for computational neuroscience. Providing a well-defined, simulator-independent and machine readable open standard for model specification, NeuroML promotes reproducibility, accessibility and portability in computational experiments. NeuroML is widely adopted as an interchange format by existing software (more than 30 third-party simulators, tools and libraries integrate NeuroML support) and web resources like NeuroMorpho.org [Ascoli et al., 2007] and Channelpedia [Ranjan et al., 2011].

As detailed in Cannon et al. [2014], NeuroML2 — the current version of NeuroML — is built on top of LEMS (Low Entropy Model Specification language). LEMS is a domain-independent language for expressing hierarchical mathematical models of physical entities, and NeuroML2 consists in a library of LEMS definitions of mathematical models relevant for computational neuroscience. LEMS and NeuroML models can be interpreted, manipulated and translated to and from other languages via a set of open source libraries and tools, like jLEMS[9], pylems[10], jNeuroML[11], libNeuroML[12], and neuroConstruct[13].

Throughout my PhD, I have contributed to the design of LEMS and NeuroML and to the implementation of the associated tools. In particular, I was responsible for the redesign of the high-level NeuroML constructs related to synaptic models. This contribution in exemplified in Figure 18, which shows the internal LEMS architecture of some of the NeuroML2 components I designed, alongside sample code describing a synaptic model. Furthermore, I maintain a Relax NG [Clark and Murata, 2001] port of the LEMS and NeuroML schemas[14]. This programmer's tool enables advanced features (like code completion and on-the-fly validity checking) on some text editors when interacting with LEMS or NeuroML files.

9 https://github.com/LEMS/jLEMS
10 https://github.com/LEMS/pylems
11 https://github.com/NeuroML/jNeuroML
12 https://pypi.python.org/pypi/libNeuroML
13 http://www.neuroconstruct.org
14 https://github.com/epiasini/LEMS_NeuroML_RelaxNG

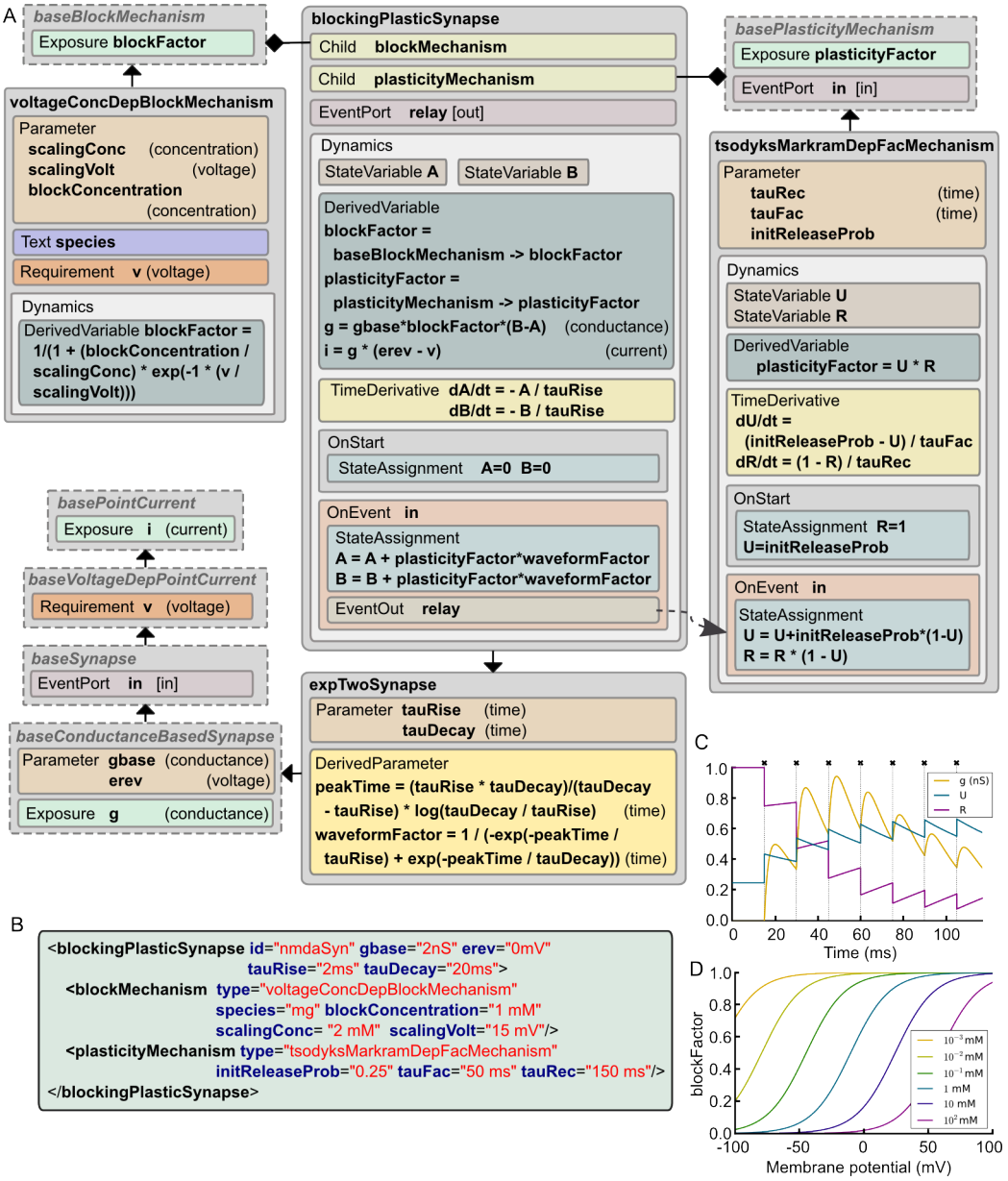


Figure 18: *(From Cannon et al. [2014]. See ibid. for details.)* Example of a complex synapse model expressed in LEMS. *A:* the `Dynamics` block in `blockingPlasticSynapse` specifies the evolution of the synaptic conductance g and how the synaptic current i depends on it and on the postsynaptic membrane potential v. In this example, a `tsodyksMarkramDepFacMechanism` component type provides a short term plasticity model based on Tsodyks et al. [1998], and a `voltageConcDepBlockMechanism` component type provides a simple model of Mg^{2+} block. Note how spike timing information is relayed from the synapse model to the plasticity mechanism through a parent-to-child `EventConnection` declared in the `Structure` element in `tsodyksMarkramDepFacMechanism`. *B:* XML code describing an NMDA receptor mediated synapse with plasticity (parameter values chosen to illustrate behavior in C,D). *C:* Behavior of the synaptic conductance g and the state variables defining plasticity, U and R, during synaptic stimulation (crosses show input events), with the postsynaptic cell clamped to 100 mV to ensure complete Mg^{2+} unblock. *D:* Changes in `blockFactor` with varying membrane potential for different values of Mg^{2+} concentration.

2.5.2 *Open Source Brain*

Open Source Brain[15] (OSB; see Gleeson et al. [2012, 2013]) is a web resource aimed at facilitating sharing, reuse and collaborative development of computational models of neuronal systems. It leverages machine-readable formats like NeuroML and LEMS to automatically extract information from the models. It offers features like interactive in-browser 3D visualisation of morphologies and networks and cross-linking with existing neuroscience databases and ontologies like ModelDB [Hines et al., 2004], Neurolex.org [Larson and Martone, 2013] and NeuroElectro [Tripathy et al., 2014]. It is not meant as a static archival database of models, but it is integrated with GitHub to provide version control, issue tracking, documentation and advanced collaboration facilities, encouraging modellers to adopt best practices regarded as standard in open source software development. Features under development for OSB include in-browser and cloud-based simulation with Geppetto[16] and automated model validation[17].

I participated to the design and implementation of OSB since its inception in 2011 [Gleeson et al., 2012], by providing intellectual contributions on the high-level scientific and community goals for the initiative, and on the best strategies to achieve them. On the technical side, my contributions included setting up the initial prototype for the web resource, designing and implementing the integration mechanism between OSB and GitHub, providing test cases, writing documentation and administering the main OSB server.

2.5.3 *Network simulations*

Single cell and synaptic models were serialized as LEMS/NeuroML2 files. Instantiations of the anatomically detailed network model were generated with Mathematica (Wolfram Research, Champaign, USA-IL) and exported in the GraphML format [Brandes et al., 2002]. These were loaded, respectively, through neuroConstruct's Jython scripting interface [Gleeson et al., 2007] and NetworkX [Hagberg et al., 2008] into custom Python software that generated NEURON simulations [Carnevale and Hines, 2006] through neuroConstruct, distributing the computational load from code generation and simulation across two HPC platforms (the SilverLab's own cluster and UCL's Legion cluster) using the Sun Grid Engine job queuing system. Spike time data was stored in compressed hdf5 archives using `h5py` [Collette, 2013]. The NEURON code generated by neuroConstruct used NEURON's built-in `NetStim` and `mcell_ran4` for pseudorandom event generation.

15 http://opensourcebrain.org/
16 http://www.geppetto.org
17 https://github.com/OpenSourceBrain/osb-model-validation

2.6 TOWARDS AN EXTENSION OF THE NETWORK ACTIVITY ANALYSES TO THE TIME DOMAIN

In the analyses presented above, in the rate-coding regime the activity of the network in a given time slice is represented by a vector $\boldsymbol{v} \in \mathbb{R}^C$ containing the spike count for each of the C cells. Conveniently, this allows us to endow the space of simulated network activity patterns with a simple Euclidean geometrical structure, which we can use as a basis for running a standard dimensionality reduction algorithm. In principle, in a spirit similar to Vargas-Irwin et al. [2014], we could apply the same type of analysis (dimensionality reduction + MI estimation) to more general coding schemes by finding an appropriate representation of the space of network activity patterns, and by equipping it with a suitable geometric structure.

An effective way of doing this is to choose an inner product space for the realisations of network activity, considering that many dimensionality reduction algorithms are (or have variants which are) *kernel based*, i.e. they can operate on a set of data points for which an inner product is defined, while an explicit coordinate representation of any point is not needed [Bishop, 2006; Filippone et al., 2008]. This is normally used in machine learning to implement the so-called *kernel trick*, whereby nonlinearities intrinsic to the data are made more manageable by embedding in a high-dimensional feature space. In my case, the embedding is necessary to endow the data with a geometrical structure, as the spike trains constituting the raw output of the simulations are simple (ordered) sets of spike times.

Several ways of defining inner products between single-cell spike trains have been proposed [Schrauwen and Van Campenhout, 2007; Paiva et al., 2010], as a way of providing a unified kernel-based picture of existing spike train metrics [Victor and Purpura, 1996; Van Rossum, 2001; Schreiber et al., 2003; Kreuz et al., 2011]. For multiple cells, Houghton and Kreuz [2012] developed a kernel-based algorithm for computing the *multi-unit Van Rossum metric* [Houghton and Sen, 2008], which can be adapted to the computation of the associated spike train inner product. The multi-unit Van Rossum metric is a generalisation of the measure defined by Van Rossum [2001] to the case of multiple cells. It is particularly appealing as it allows to choose freely the resolution on the temporal as well as the spatial dimension. This happens by interpolating smoothly not only between rate and time coding "mode" but also between *summed-population code*, where the spikes from all cells are pooled together as if seen by a common downstream integrator, and *labelled-line code*, where each spike is "labelled" with the identity of the cell emitting it, and these labels are taken into account when computing the dissimilarity between two instances of network activity. In the following section, I generalise the approach in

multi-unit Van Rossum metric

summed-population code

labelled-line code

Houghton and Kreuz [2012] to arbitrary kernel-based spike train metrics and I show that the natural geometrical structure for reasoning about this type of metrics is that of a tensor product space.

2.6.1 *A simple formalism for kernel-based multiunit spike train metrics*

Here I define a compact formalism for multiunit kernel-based spike-train metrics by opportunely characterising the space of multiunit feature vectors as a tensor product. Previous results from Houghton and Kreuz [2012] on a formula for efficient computation of multiunit Van Rossum metrics are then re-derived within this framework, while fixing some errors in the original calculations.

Consider a network with C cells. Let

$$\mathcal{U} = \left\{ \boldsymbol{u}^1, \boldsymbol{u}^2, \ldots, \boldsymbol{u}^C \right\}$$

be an *observation of network activity*, where

$$\boldsymbol{u}^i = \left\{ u_1^i, u_2^i, \ldots, u_{N_{\boldsymbol{u}^i}}^i \right\}$$

is the (ordered) set of times of the spikes emitted by cell i. Let $\mathcal{V} = \left\{ \boldsymbol{v}^1, \boldsymbol{v}^2, \ldots, \boldsymbol{v}^C \right\}$ be another observation, different in general from $\mathcal{U}$.

To compute a kernel based multiunit distance between $\mathcal{U}$ and $\mathcal{V}$, we map them to the tensor product space $\mathcal{S} \doteq \mathbb{R}^C \bigotimes L_2(\mathbb{R} \to \mathbb{R})$ by defining

$$|\mathcal{U}\rangle = \sum_{i=1}^{C} |i\rangle \otimes |\boldsymbol{u}^i\rangle$$

where we consider $\mathbb{R}^C$ and $L_2(\mathbb{R} \to \mathbb{R})$ to be equipped with the usual euclidean distances, consequently inducing an euclidean metric structure on $\mathcal{S}$ too.

Conceptually, the set of vectors $\{|i\rangle\}_{i=1}^{C} \subset \mathbb{R}^C$ represents the different cells, while each $|\boldsymbol{u}^i\rangle \in L_2(\mathbb{R} \to \mathbb{R})$ represents the convolution of a spike train of cell i with a real-valued feature function $\phi : \mathbb{R} \to \mathbb{R}$,

$$\langle t|\boldsymbol{u}\rangle = \sum_{n=1}^{N_{\boldsymbol{u}}} \phi(t - u_n)$$

In practice, we will never use the feature functions directly, but we will be only interested in the inner products of the $|i\rangle$ and $|\boldsymbol{u}\rangle$ vectors.

We call $c_{ij} \doteq \langle i|j\rangle = \langle i|j\rangle_{\mathbb{R}^C} = c_{ji}$ the *multiunit mixing coefficient* for cells i and j, and $\langle \boldsymbol{u}|\boldsymbol{v}\rangle = \langle \boldsymbol{u}|\boldsymbol{v}\rangle_{L_2}$ the *single-unit inner product*,

$$\begin{aligned}\langle \boldsymbol{u}|\boldsymbol{v}\rangle &= \langle \{u_1, u_2, \ldots, u_N\} | \{v_1, v_2, \ldots, v_M\}\rangle = \\ &= \int dt \langle \boldsymbol{u}|t\rangle\langle t|\boldsymbol{v}\rangle = \int dt \sum_{n=1}^{N}\sum_{m=1}^{M} \phi(t-u_n)\,\phi(t-v_m) \\ &\doteq \sum_{n=1}^{N}\sum_{m=1}^{M} \mathcal{K}(u_n, v_m)\end{aligned} \tag{2.21}$$

where $\mathcal{K}(t_1, t_2) \doteq \int dt \left[\phi(t-t_1)\,\phi(t-t_2)\right]$ is the *single-unit metric kernel*, and where we have used the fact that the feature function ϕ is real-valued. It follows immediately from the definition above that $\langle \boldsymbol{u}|\boldsymbol{v}\rangle = \langle \boldsymbol{v}|\boldsymbol{u}\rangle$.

Note that, given a cell pair (i, j) or a spike train pair $(\boldsymbol{u}, \boldsymbol{v})$, c_{ij} does not depend on spike times and $\langle \boldsymbol{u}|\boldsymbol{v}\rangle$ does not depend on cell labeling.

With this notation, we can define the *multi-unit spike train distance* as

$$\left\| |\mathcal{U}\rangle - |\mathcal{V}\rangle \right\|^2 = \langle \mathcal{U}|\mathcal{U}\rangle + \langle \mathcal{V}|\mathcal{V}\rangle - 2\langle \mathcal{U}|\mathcal{V}\rangle \tag{2.22}$$

where the *multi-unit spike train inner product* $\langle \mathcal{V}|\mathcal{U}\rangle$ between $\mathcal{U}$ and $\mathcal{V}$ is just the natural bilinear operation induced on $\mathcal{S}$ by the tensor product structure:

$$\begin{aligned}\langle \mathcal{V}|\mathcal{U}\rangle &= \sum_{i,j=1}^{C} \langle i|j\rangle \langle \boldsymbol{v}^i|\boldsymbol{u}^j\rangle = \sum_{i,j=1}^{C} c_{ij}\langle \boldsymbol{v}^i|\boldsymbol{u}^i\rangle \\ &= \sum_{i=1}^{C}\left[c_{ii}\langle \boldsymbol{v}^i|\boldsymbol{u}^i\rangle + c_{ij}\left(\sum_{j<i}\langle \boldsymbol{v}^i|\boldsymbol{u}^j\rangle + \sum_{j>i}\langle \boldsymbol{v}^i|\boldsymbol{u}^j\rangle\right)\right]\end{aligned} \tag{2.23}$$

But $c_{ij} = c_{ji}$ and $\langle \boldsymbol{v}|\boldsymbol{u}\rangle = \langle \boldsymbol{u}|\boldsymbol{v}\rangle$, so

$$\begin{aligned}\sum_{i=1}^{C}\sum_{j<i} c_{ij}\langle \boldsymbol{v}^i|\boldsymbol{u}^j\rangle = \sum_{j=1}^{C}\sum_{i<j} c_{ji}\langle \boldsymbol{v}^j|\boldsymbol{u}^i\rangle = \sum_{i=1}^{C}\sum_{j>i} c_{ji}\langle \boldsymbol{v}^j|\boldsymbol{u}^i\rangle = \sum_{i=1}^{C}\sum_{j>i} c_{ij}\langle \boldsymbol{v}^j|\boldsymbol{u}^i\rangle \\ = \sum_{i=1}^{C}\sum_{j>i} c_{ij}\langle \boldsymbol{u}^i|\boldsymbol{v}^j\rangle\end{aligned}$$

and

$$\langle \mathcal{V}|\mathcal{U}\rangle = \sum_{i=1}^{C}\left[c_{ii}\langle \boldsymbol{v}^i|\boldsymbol{u}^i\rangle + c_{ij}\sum_{j>i}\left(\langle \boldsymbol{v}^i|\boldsymbol{u}^j\rangle + \langle \boldsymbol{u}^i|\boldsymbol{v}^j\rangle\right)\right] \tag{2.24}$$

Now, normally we are interested in the particular case where c_{ij} is the same for all pair of distinct cells:

$$c_{ij} = \begin{cases} 1 & \text{if } i = j \\ c & \text{if } i \neq j \end{cases}$$

Under this assumption, (2.24) reduces to

$$\langle \mathcal{V}|\mathcal{U}\rangle = \sum_{i=1}^{C} \left[\langle \boldsymbol{v}^i|\boldsymbol{u}^i\rangle + c \sum_{j>i} \left(\langle \boldsymbol{v}^i|\boldsymbol{u}^j\rangle + \langle \boldsymbol{u}^i|\boldsymbol{v}^j\rangle \right) \right] \tag{2.25}$$

and substituting (2.25) into (2.22) we get

$$\begin{aligned} \left\| |\mathcal{U}\rangle - |\mathcal{V}\rangle \right\|^2 = \sum_{i=1}^{C} \Bigg\{ & \langle \boldsymbol{u}^i|\boldsymbol{u}^i\rangle + c \sum_{j>i} \left(\langle \boldsymbol{u}^i|\boldsymbol{u}^j\rangle + \langle \boldsymbol{u}^i|\boldsymbol{u}^j\rangle \right) + \\ & + \langle \boldsymbol{v}^i|\boldsymbol{v}^i\rangle + c \sum_{j>i} \left(\langle \boldsymbol{v}^i|\boldsymbol{v}^j\rangle + \langle \boldsymbol{v}^i|\boldsymbol{v}^j\rangle \right) + \\ & - 2 \left[\langle \boldsymbol{v}^i|\boldsymbol{u}^i\rangle + c \sum_{j>i} \left(\langle \boldsymbol{v}^i|\boldsymbol{u}^j\rangle + \langle \boldsymbol{u}^i|\boldsymbol{v}^j\rangle \right) \right] \Bigg\} \end{aligned}$$

Finally, rearranging the terms

$$\begin{aligned} \left\| |\mathcal{U}\rangle - |\mathcal{V}\rangle \right\|^2 = \sum_{i=1}^{C} \Bigg[& \langle \boldsymbol{u}^i|\boldsymbol{u}^i\rangle + \langle \boldsymbol{v}^i|\boldsymbol{v}^i\rangle - 2\langle \boldsymbol{v}^i|\boldsymbol{u}^i\rangle + \\ & + 2c \sum_{j>i} \left(\langle \boldsymbol{u}^i|\boldsymbol{u}^j\rangle + \langle \boldsymbol{v}^i|\boldsymbol{v}^j\rangle - \langle \boldsymbol{v}^i|\boldsymbol{u}^j\rangle - \langle \boldsymbol{v}^j|\boldsymbol{u}^i\rangle \right) \Bigg] \end{aligned} \tag{2.26}$$

2.6.1.1 *Van Rossum-like metrics*

In Van Rossum-like metrics, the feature function and the single-unit kernel are, for $\tau \neq 0$,

$$\phi_\tau^{VR}(t) = \sqrt{\frac{2}{\tau}} \cdot e^{-t/\tau} \theta(t)$$

$$\mathcal{K}_\tau^{VR}(t_1, t_2) = \begin{cases} 1 & \text{if } t_1 = t_2 \\ e^{-|t_1 - t_2|/\tau} & \text{if } t_1 \neq t_2 \end{cases}$$

where θ is the Heaviside step function (with $\theta(0) = 1$), and we have chosen to normalise ϕ_τ^{VR} so that

$$\left\|\phi_\tau^{VR}\right\|_2 = \sqrt{\int dt \left[\phi_\tau^{VR}(t)\right]^2} = 1 \quad .$$

With this choice of kernel, the single-unit inner product (2.21) now becomes

$$\langle \boldsymbol{u}|\boldsymbol{v}\rangle = \sum_{n=1}^{N}\sum_{m=1}^{M} \mathcal{K}^{VR}(u_n, v_m) = \sum_{n=1}^{N}\sum_{m=1}^{M} e^{-|u_n - v_m|/\tau} \tag{2.27}$$

In the $\tau \to 0$ limit,

$$\phi_0^{VR}(t) = \delta(t)$$

$$\mathcal{K}_0^{VR}(t_1, t_2) = \begin{cases} 1 & \text{if } t_1 = t_2 \\ 0 & \text{if } t_1 \neq t_2 \end{cases}$$

and

$$\langle \boldsymbol{u}|\boldsymbol{v}\rangle = \sum_{n=1}^{N}\sum_{m=1}^{M} \delta(u_n, v_m) \quad , \quad \langle \boldsymbol{u}|\boldsymbol{u}\rangle = N$$

MARKAGE FORMULAS For a spike train $\boldsymbol{u}$ of length N and a time t we define the index $\tilde{N}(\boldsymbol{u}, t)$

$$\tilde{N}(\boldsymbol{u}, t) \doteq \max\{n | u_n < t\}$$

which we can use to re-write (2.27) without the absolute values:

$$\begin{aligned}
\langle \boldsymbol{u}|\boldsymbol{v}\rangle &= \sum_{n=1}^{N}\left(\sum_{m|v_m<u_n} e^{-(u_n - v_m)/\tau} + \sum_{m|v_m>u_n} e^{-(v_m - u_n)/\tau} + \sum_{m=1}^{M}\delta(u_n, v_m)\right) \\
&= \sum_{n=1}^{N}\left(\sum_{m|v_m<u_n} e^{-(u_n - v_m)/\tau} + \sum_{m|u_m<v_n} e^{-(v_n - u_m)/\tau} + \sum_{m=1}^{M}\delta(u_n, v_m)\right) \\
&= \sum_{n=1}^{N}\left[\sum_{m=1}^{\tilde{N}(\boldsymbol{v},u_n)} e^{-(u_n - v_m)/\tau} + \sum_{m=1}^{\tilde{N}(\boldsymbol{u},v_n)} e^{-(v_n - u_m)/\tau} + \delta\left(u_n, v_{\tilde{N}(\boldsymbol{v},u_n)+1}\right)\right]
\end{aligned}$$

$$= \sum_{n=1}^{N} \left[e^{-(u_n - v_{\tilde{N}(\boldsymbol{v},u_n)})/\tau} \sum_{m=1}^{\tilde{N}(\boldsymbol{v},u_n)} e^{-(v_{\tilde{N}(\boldsymbol{v},u_n)} - v_m)/\tau} + \right.$$
$$+ e^{-(v_n - u_{\tilde{N}(\boldsymbol{u},v_n)})/\tau} \sum_{m=1}^{\tilde{N}(\boldsymbol{u},v_n)} e^{-(u_{\tilde{N}(\boldsymbol{u},v_n)} - u_m)/\tau} +$$
$$\left. + \delta\left(u_n, v_{\tilde{N}(\boldsymbol{v},u_n)+1}\right) \right] \tag{2.28}$$

For a spike train $\boldsymbol{u}$ of length N, we also define the the *markage vector* $\boldsymbol{m}$, with the same length as $\boldsymbol{u}$, through the following recursive assignment:

$$m_1(\boldsymbol{u}) \doteq 0 \tag{2.29}$$
$$m_n(\boldsymbol{u}) \doteq (m_{n-1} + 1)\, e^{-(u_n - u_{n-1})/\tau} \quad \forall n \in \{2, \ldots, N\} \tag{2.30}$$

It is easy to see that

$$m_n(\boldsymbol{u}) = \sum_{k=1}^{n-1} e^{-(u_n - u_k)/\tau} = \left(\sum_{k=1}^{n} e^{-(u_n - u_k)/\tau} \right) - e^{-(u_n - u_n)/\tau}$$
$$= \sum_{k=1}^{n} e^{-(u_n - u_k)/\tau} - 1 \tag{2.31}$$

and in particular

$$\sum_{n=1}^{\tilde{N}(\boldsymbol{u},t)} e^{-(u_{\tilde{N}(\boldsymbol{u},t)} - u_n)/\tau} = 1 + m_{\tilde{N}(\boldsymbol{u},t)}(\boldsymbol{u}) \tag{2.32}$$

Substituting (2.32) in (2.28), we get

$$\langle \boldsymbol{u} | \boldsymbol{v} \rangle = \sum_{n=1}^{N} \left[e^{-(u_n - v_{\tilde{N}(\boldsymbol{v},u_n)})/\tau} \left(1 + m_{\tilde{N}(\boldsymbol{v},u_n)}(\boldsymbol{v})\right) + \right.$$
$$+ e^{-(v_n - u_{\tilde{N}(\boldsymbol{u},v_n)})/\tau} \left(1 + m_{\tilde{N}(\boldsymbol{u},v_n)}(\boldsymbol{u})\right) + \tag{2.33}$$
$$\left. + \delta\left(u_n, v_{\tilde{N}(\boldsymbol{v},u_n)+1}\right) \right]$$

Finally, note that because of the definition of the markage vector (2.30)

$$e^{-(u_n - u_{\tilde{N}(\boldsymbol{u},u_n)})/\tau} \left(1 + m_{\tilde{N}(\boldsymbol{u},u_n)}(\boldsymbol{u})\right) = e^{-(u_n - u_{n-1})/\tau} \left(1 + m(\boldsymbol{u})\right) = m_n(\boldsymbol{u})$$

so that in particular

$$\langle \boldsymbol{u}|\boldsymbol{u}\rangle = \sum_{n=1}^{N} \left(1 + 2m_n(\boldsymbol{u})\right) \tag{2.34}$$

A formula for the efficient computation of the multiunit Van Rossum spike train distance (resp. inner product) can then be obtained by opportunely substituting (2.33) and (2.34) into the definition in Equation 2.26 (resp. Equation 2.25). For a pair of spike trains with n spikes each, the complexity of the calculation required to compute a spike train distance with this method scales linearly with n. This should be contrasted with a naive kernel approach, which scales like n^2. See the original paper by Houghton and Kreuz [2012] for an in-depth discussion of this and other computational complexity issues.

2.6.2 *A software package for fast computation of multi-unit Van Rossum metrics*

C++ code for the calculation of the distance formula derived above was published with Houghton and Kreuz [2012]. Unfortunately, this code was packaged more as a demo for the paper than as a library, contained no internal documentation or comments, and presented some bugs and unnecessary limitations on the maximum length of spike trains. With the goal of making it more accessible, starting from the original code by Houghton and Kreuz I developed and published a Python C++ extension called `pymuvr`[18], ready to use as a standalone module or to be incorporated into larger pieces of software. While the original code could only compute *distances* between multi-unit spike trains, `pymuvr` can compute *distances* or *scalar products*, allowing for immediate interaction with machine learning packages like `scikit-learn` [Pedregosa et al., 2011]. Compared to existing software, then, `pymuvr` combines the quantitative speed advantage of hand-optimised low-level code with the qualitative difference of the new possibilities enabled by the direct calculation of inner products and the ease of use of a package for Python, one of the most popular languages in the computational neuroscience community [Ray and Bhalla, 2008; Koetter et al., 2009; Davison et al., 2009; Goodman and Brette, 2009; Muller et al., 2009; Cornelis et al., 2012; Vella et al., 2014; Cannon et al., 2014].

Since I had the opportunity to design `pymuvr` from the ground up, I tried to prevent common problems with scientific software development by adopting a set of industry standard best practices [Wilson, 2006]. `pymuvr` is *free software* [Stallman, 2010], published under the

free software

18 https://pypi.python.org/pypi/pymuvr

GNU General Public Licence. Its interface is well documented[19]; the code is extensively commented and kept under version control in a Git repository on GitHub[20], where issues and bug reports can be created and tracked; a battery of unit tests has been written and, following the *continuous integration* model, is run with several different versions of Python every time a new commit is pushed to the GitHub repository[21]. `pymuvr` runs with Python 2 and 3 under GNU/Linux, Mac OSX and Microsoft Windows.

Since its release, `pymuvr` has been added as an optional dependency to `spykeutils`[22], a spike sorting and analysis toolbox. This happened under the spontaneous initiative of the `spykeutils` lead developer, who I had no previous contact with. The package is also advertised on the home page of one of the authors of the algorithm[23] as an alternative to the code originally published with Houghton and Kreuz [2012].

2.6.2.1 *Comparison with an existing pure Python implementation*

To get an idea of how fast `pymuvr` is, I compared its performance to the Python/NumPy implementation of the same algorithm provided by `spykeutils`. An example of such benchmark is shown in Figure 19, demonstrating a ~25x speedup on the set of tasks considered.

2.6.3 *Potential directions of future development*

Naively extending the network activity analyses presented in this thesis to the time domain would undoubtedly be problematic from the point of view of computational complexity. In particular, analysing information transmission by considering spike trains instead of spike counts may require the use of methods such as *approximate kernel k-means* [Chitta et al., 2011] or the embedding approach proposed by Elgohary et al. [2014] to implement kernel k-means on MapReduce [Dean and Ghemawat, 2008]. In the former case, `pymuvr` can be used to easily prototype a Python implementation of the proposed algorithm. In the latter, the Python package or the clean and well documented interface of its internal C++ code could still be useful in conjunction with tools like Apache Hadoop Streaming[24].

19 http://pymuvr.readthedocs.org
20 https://github.com/epiasini/pymuvr
21 https://travis-ci.org/epiasini/pymuvr
22 https://github.com/rproepp/spykeutils
23 http://wwwold.fi.isc.cnr.it/users/thomas.kreuz/sourcecode.html
24 http://hadoop.apache.org/docs/current/hadoop-mapreduce-client/hadoop-mapreduce-client-core/HadoopStreaming.html

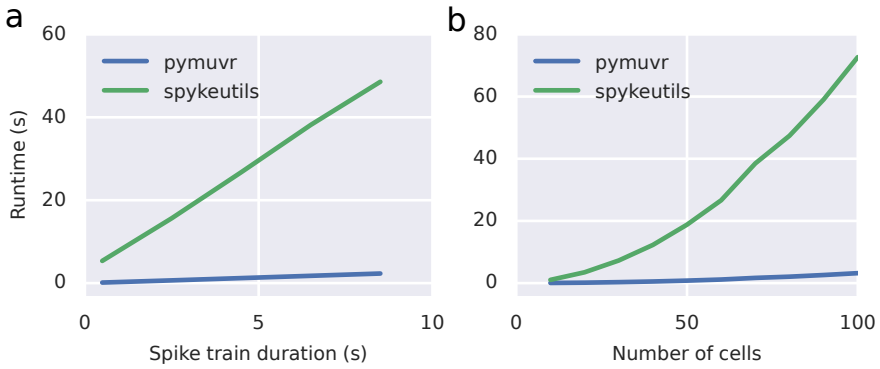


Figure 19: Simple benchmark for the computation of multi-unit Van Rossum distances for `pymuvr` against `spykeutils`. For a group of cells, 5 observations of "network" activity are generated as Poisson spike trains with mean rate 300 Hz, and the full set of $5 \cdot (5-1) = 20$ inter-observation distances are computed with both software packages. The time scale of the exponential convolution for the spike trains is 12 ms and the multiunit mixing parameter of the metric is 0.1. *a*: scaling of the runtime with the spike train length for a fixed number of cells ($C = 50$). Note that, as mentioned in the main text, the complexity of the algorithm implemented by both packages scales linearly with the number of spikes. *b*: scaling of the runtime with the number of cells for a fixed spike train length (3 s).

2.7 NOTES ON SPIKE TRAIN GENERATION

In simulations we often need to generate random event (spike) trains. This could also occur when implementing experimental protocols; indeed, the notes presented in this section follow from some work I did while designing a system for the generation of conductance trains for dynamic clamp experiments.

Most of the time, generating random spike trains means modelling them as *renewal processes* [Gerstner and Kistler, 2002, section 5.2], defined as point processes on the real line where the probability of an event occurring at time t only depends on the time $t - \hat{t}$ that has passed since the last event, occurred at time $\hat{t}$. In other words, in a renewal process the length of successive inter-events intervals is independent, and the probability distribution for inter-event intervals (IEIs, or ISIs in the spike train case) completely defines the process. Choosing one type of renewal process over another means defining the functional form of this probability distribution. For instance, in a *Poisson* process the probability is given by a simple exponential function of $t - \hat{t}$, while in a *gamma* process it is a gamma function.

A good way of letting our assumptions on the temporal statistics of the spike train inform our choice of process type is to choose the process where ISIs are distributed according to the maximum entropy distribution given our constraints (see Pressé et al. [2013] for a general introduction to the maximum entropy principle). For example, the IEI

distribution for the Poisson process with rate r is the maximum entropy distribution defined over $[0, +\infty)$ with mean r^{-1}; and see Park and Bera [2009] for a table of various common distributions together with the constraints under which they are maximum entropy.

2.7.1 *Two examples of "modified Poisson" processes with a refractory period*

The gamma renewal process has long been a popular choice for describing spike trains [Stein, 1965], as it is fairly flexible, and in particular it allows for the representation of a relative refractory period. On the other hand, in cases where the desired temporal structure is completely described by an average event rate τ^{-1} and an absolute refractory period κ, simpler approaches have been sought, often by defining IEIs as the result of some sort of algorithmical manipulation on data generated from the exponential distribution defined by

$$f_\tau(t) \doteq \frac{e^{-t/\tau}}{\tau} \quad , \quad t \in [0, +\infty) \quad .$$

In the examples below, I will illustrate why this is in general not a good idea.

2.7.1.1 *Naive rejection sampling*

In this case, IEIs are generated according to $f_\tau(t)$, but the result is rejected whenever $t < \kappa$. This approach is appealing because it starts from something that satisfies one of the desired properties of the event train (the average rate of the Poisson event train generated with $f_\tau(t)$ is τ^{-1}) and adds the second requirement (the refractory period κ) in what seems to be the simplest possible way. The issue here is that the rejection step skews the rate statistics, so that the resulting average rate is not τ^{-1} anymore. In fact, the IEI distribution we end up sampling from is

$$s_{\tau,\kappa}(t) \doteq \frac{e^{\kappa/\tau}}{\tau} e^{-t/\tau} \quad , \quad t \in [\kappa, +\infty)$$

as can be easily found by renormalising $f_\tau(t)$ over $[\kappa, +\infty)$. Now,

$$\begin{aligned}
\langle t \rangle_{s_{\tau,\kappa}} &= \int_\kappa^{+\infty} t s_{\tau,\kappa}(t)\, dt = \frac{e^{\kappa/\tau}}{\tau} \int_\kappa^{+\infty} t e^{-t/\tau}\, dt \\
&= \frac{e^{\kappa/\tau}}{\tau} \left(\tau t e^{-t/\tau} \Big|_{+\infty}^{\kappa} + \tau \int_\kappa^{+\infty} e^{-t/\tau}\, dt \right) \\
&= e^{\kappa/\tau} \left(\kappa e^{-\kappa/\tau} + \tau e^{-\kappa/\tau} \right) \\
&= \kappa + \tau
\end{aligned}$$

which means that simply rejecting all samples with $t < \kappa$ shifts the average rate from τ^{-1} to $(\kappa+\tau)^{-1}$. So if $r \doteq \tau^{-1}$ is the *nominal* rate of the event train, $R \doteq \kappa^{-1}$ is the *maximum* allowed rate and $\tilde{r}\left[s_{\tau,\kappa}\right] \doteq \left(\langle t \rangle_{s_{\tau,\kappa}}\right)^{-1}$ is the *actual* average rate when sampling from $s_{\tau,\kappa}$, then

$$\tilde{r}\left[s_{\tau,\kappa}\right] = \frac{1}{\frac{1}{R}+\frac{1}{r}} = \frac{R}{1+\frac{R}{r}}$$

and

$$\frac{\tilde{r}\left[s_{\tau,\kappa}\right]}{R} = \frac{1}{1+\frac{R}{r}} \tag{2.35}$$

2.7.1.2 *Rejection sampling with substitution*

This is a variant of the previous case, where whenever we sample a value $t < \kappa$ from $f_\tau(t)$ we consider it to be equal to κ instead of rejecting it. The corresponding IEI distribution is

$$s'_{\tau,\kappa}(t) = \left(1-e^{-t/\tau}\right)\delta(t-\kappa) + \frac{e^{-t/\tau}}{\tau} \quad , \quad t \in [\kappa, +\infty)$$

and the resulting average IEI is

$$\begin{aligned}\langle t \rangle_{s'_{\tau,\kappa}} &= \int_\kappa^{+\infty} t s'_{\tau,\kappa}(t)\,dt = \kappa\left(1-e^{-\kappa/\tau}\right) + \int_\kappa^{+\infty} \frac{te^{-t/\tau}}{\tau}\,dt \\ &= \kappa\left(1-e^{-\kappa/\tau}\right) + e^{-\kappa/\tau}(\kappa+\tau) \\ &= \kappa + \tau e^{-\kappa/\tau}\end{aligned}$$

which, again, entails a distortion between the actual and the nominal average rate: similarly to Equation 2.35,

$$\frac{\tilde{r}[s'_{\tau,\kappa}]}{R} = \frac{1}{1+\frac{R}{r}e^{-r/R}} \tag{2.36}$$

Figure 20 shows a comparison between the results in (2.35) and (2.36), and the desired situation where $\tilde{r} = r$.

2.7.2 *Maximum entropy renewal process with given rate and absolute refractory period*

As we mentioned above, a more principled choice for a renewal process with a given rate and an absolute refractory period would be that whose ISI distribution has maximum entropy under the required constraints; in the following discussion, I will determine this distribution.

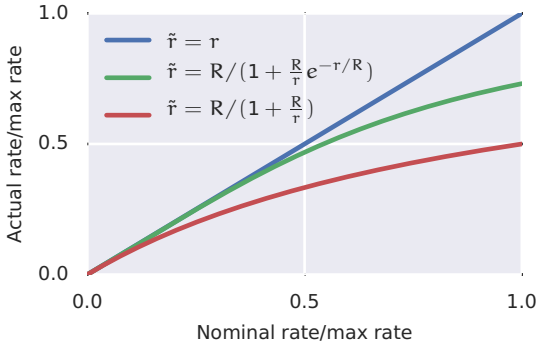


Figure 20: Actual rate $\tilde{r}$ versus nominal rate r, normalised by maximum rate R. Red is rejection sampling (2.35), green is rejection with substitution (2.36).

Note that I will indicate the differential entropy of a probability density function p with support $\mathcal{S}$ by the symbol $h[p]$:

$$h[p] \doteq -\int_{\mathcal{S}} p(x) \log p(x) \, dx$$

Lemma 1. *Let p and q be continuous probability density functions on an interval $I \subset \mathbb{R}$, such that $h[p], h[q] < \infty$. Assume $q(t) > 0$ for all $t \in I$. If*

$$-\int_I p(t) \log q(t) \, dt \leqslant h[q]$$

then

$$h[p] \leqslant h[q]$$

with equality iff $p = q$ almost everywhere.

Proof.

$$\begin{aligned} h[p] &= -\int_I p(t) \log p(t) \, dt \\ &= -\int_I p(t) \log \frac{p(t)}{q(t)} \, dt - \int_I p(t) \log q(t) \\ &= -D_{KL}\left(p \parallel q\right) - \int_I p(t) \log q(t) \, dt \\ &\leqslant -\int_I p(t) \log q(t) \, dt \\ &\leqslant h[q] \end{aligned}$$

where $D_{KL}\left(p \parallel q\right)$ is the Kullback-Leibler divergence [Cover and Thomas, 2006] between p and q and we have used the fact that $D_{KL}\left(p \parallel q\right) \geqslant 0$ for any choice of p and q. The case for equality follows immediately considering that $D_{KL}\left(p \parallel q\right) = 0$ iff $p = q$ almost everywhere. □

We can now write an ansatz for the maximum entropy distribution over the interval $[\kappa, +\infty)$ with mean τ

$$q_{\tau,\kappa}(t) \doteq \frac{e^{-\frac{t-\kappa}{\tau-\kappa}}}{\tau-\kappa} \quad , \quad t \in [\kappa, +\infty) \tag{2.37}$$

Note that $q_{\tau,\kappa}(t)$ is normalised:

$$\int_\kappa^{+\infty} q(t)\,dt = \frac{e^{\frac{\kappa}{\tau-\kappa}}}{\tau-\kappa}\int_\kappa^{+\infty} e^{-\frac{t}{\tau-\kappa}} = e^{\frac{\kappa}{\tau-\kappa}}\left[e^{-\frac{t}{\tau-\kappa}}\right]_{+\infty}^{\kappa} = 1$$

To show that this is indeed the unique solution to our problem, we start by computing its differential entropy.

$$\begin{aligned}
h\left[q_{\tau,\kappa}\right] &= -\int_\tau^{+\infty} q_{\tau,\kappa}(t)\log q_{\tau,\kappa}(t)\,dt \\
&= \frac{1}{\tau-\kappa}\int_\tau^{+\infty} e^{-\frac{t-\kappa}{\tau-\kappa}}\left[\log(\tau-\kappa) + \frac{t-\kappa}{\tau-\kappa}\right]dt \\
&= \log(\tau-\kappa) + \frac{1}{\tau-\kappa}\int_\tau^{+\infty}\frac{t-\kappa}{\tau-\kappa}e^{-\frac{t-\kappa}{\tau-\kappa}}\,dt \\
&= \log(\tau-\kappa) + \int_0^{+\infty} x e^x\,dx \\
&= \log(\tau-\kappa) + 1
\end{aligned}$$

where we have substituted $x \doteq (t-\kappa)/(\tau-\kappa)$. But now, if we consider any other distribution p defined over $[\kappa, +\infty)$ with mean τ, i.e. such that

$$\int_\tau^{+\infty} p(t)\,dt = 1 \quad , \quad \int_\tau^{+\infty} t p(t)\,dt = \tau$$

then

$$\begin{aligned}
-\int_\tau^{+\infty} p(t)\log q(t)\,dt &= \int_\tau^{+\infty} p(t)\left[\log(\tau-\kappa) + \frac{t-\tau}{\tau-\kappa}\right]dt \\
&= \log(\tau-\kappa) + 1 \\
&= h\left[q_{\tau,\kappa}\right]
\end{aligned}$$

Finally, by Lemma 1 this implies

$$h[p] \leqslant h\left[q_{\tau,\kappa}\right]$$

for any p, and the bound is achieved only if $p = q_{\tau,\kappa}$. This proves that $q_{\tau,\kappa}$ is the (unique) maximum entropy distribution on $[\kappa, +\infty)$ with mean τ.

2.7.2.1 *Sampling from the maximum entropy distribution*

If we want to generate a spike train with $q_{\tau,\kappa}$ for inter-spike interval distribution, we are faced with the problem of sampling from it starting from, say, a pseudo-random number generator that can sample from $\mathcal{U}_{[0,1)}$, the uniform distribution over the interval $[0, 1)$. Let's start by defining

$$s_\tau \doteq \frac{e^{-t/\tau}}{\tau}$$

This is a simple exponential distribution, and we can sample from it by computing

$$t_i = -\ln(u_i) \cdot \tau$$

where $u_i \sim \mathcal{U}_{[0,1)}$ (see, for example, Rothman and Silver [2014]). Now observe that, for any $t \in [0, +\infty)$,

$$q_{\tau,\kappa}(t + \tau) = s_{\tau-\kappa}(t)$$

Hence, an easy way of sampling from $q_{\tau,\kappa}$ is to sample from $s_{\tau-\kappa}(t)$ and then add κ to the sampled number.

3

SPATIAL CONSTRAINTS AFFECT NETWORK STRUCTURE IN THE GRANULE CELL LAYER

3.1 A MODEL OF MF TO GRC CONNECTIVITY SHOWS SIGNIFICANT STRUCTURAL PROPERTIES EMERGING FROM SPATIAL CONSTRAINTS

The space-dependent connectivity rules described in subsection 2.3.1 for the local GCL network induce some simple structural characteristics that are easy to measure with standard network-theoretical tools.

weighted one-mode projection

Figure 21 shows the network formed by a *weighted one-mode projection* [Newman, 2010] of the local GCL model with $d = 4$ synaptic connections onto the set of nodes representing GrCs. The nodes of this network coincide with the GrCs in the model; two nodes are connected by an edge if they share at least one neighbour (i.e. they are connected to at least one common MF) in the original network, and the weight of the edge is

$$w = d - s$$

where d is the number synaptic connections per GrC and s is the number of shared neighbours. In Figure 21, one GrC is fixed and the length of the shortest (weighted) path between it and all other nodes of the projected network is shown with a colour code. Considering that the spatial extent of the projected network is the same of the original network (i.e. an 80 µm diameter ball), this shows how the constraint on GrC dendrite length sets the scale for the spatial structure of the network, limiting input sharing to GrCs within a reciprocal distance of about 40 µm. By contrast, Figure 22 shows the same measure for the corresponding randomised bipartite graph, obtained from the local GCL model by randomly reassigning the MF ends of all edges, so that d is conserved. It is evident that in this case the spatial structure of the network is lost, and that the amount of input sharing is lower overall.

bipartite clustering coefficient

We can further expand on the ideas presented above by comparing the *bipartite clustering coefficient* [Robins and Alexander, 2004] of the local GCL model and its randomised counterpart. For a bipartite graph, with nodes of - say, in our case - type "MF rosette" and type "GrC", the bipartite clustering coefficient is the number of 4-cycles formed by a pair of "GrC" nodes being both connected to the same pair of "MF rosette" nodes, normalised by the maximum number of such cycles possible in a network with the given number of nodes. In

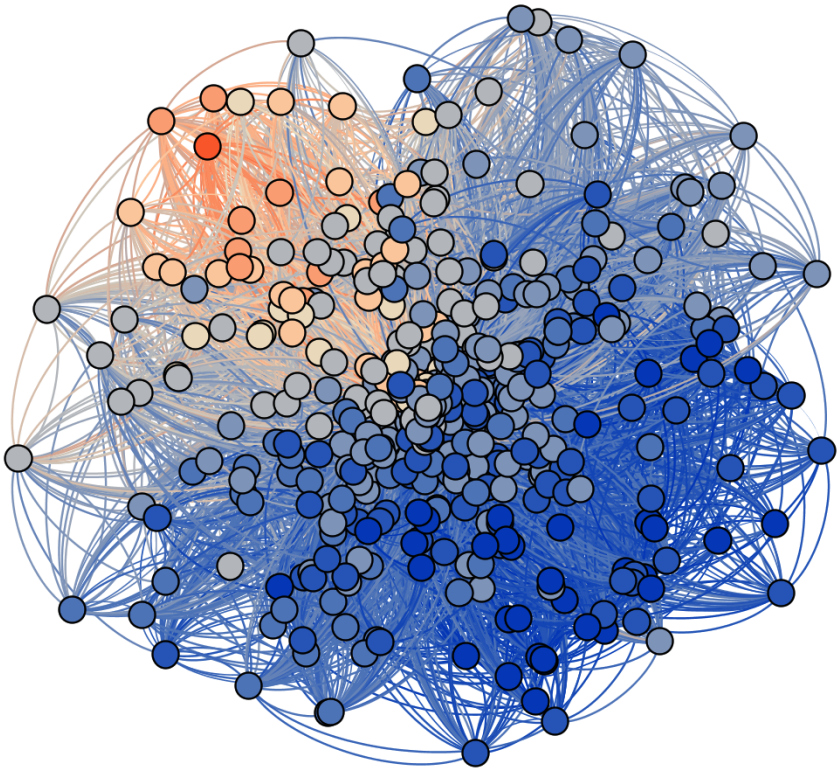


Figure 21: Weighted one-mode projection of the local granule cell layer model with four synaptic connections per GrC onto the set of graph nodes representing granule cells. Cell positions are given by a projection of the 3D coordinates onto the (x,y) plane. Colour code represents the length of the shortest weighted path separating a fixed GrC (the one in bright orange) and every other GrC in the one-mode network. Orange: full input sharing (four MF terminals in common). Light or dark blue: no shared input. Dark blue: no shared input *and* little presence of common neighbours in the projected network. Figure realised with NetworkX [Hagberg et al., 2008] and Gephi [Bastian et al., 2009].

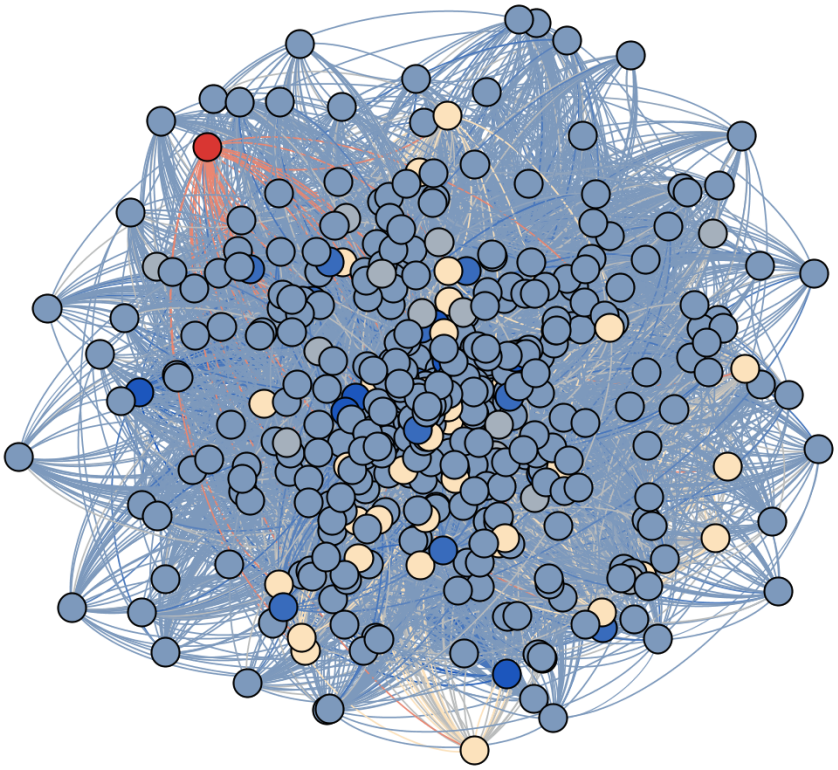


Figure 22: As Figure 21, but for a random bipartite graph with the same nodes and number of synaptic input per GrC as the local GCL model.

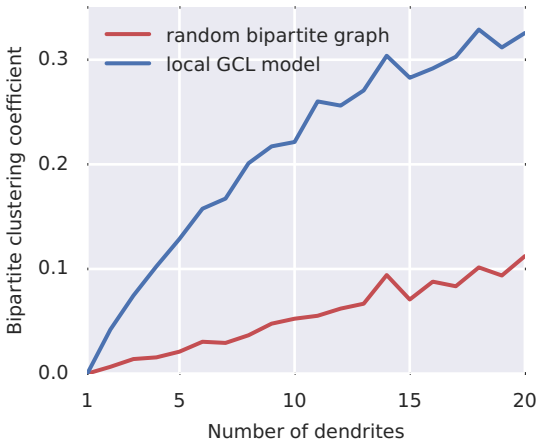


Figure 23: Blue: bipartite clustering coefficient [Robins and Alexander, 2004] for the local GCL network model, as a function of the number of synaptic connections per granule cell. Red: same quantity for a random bipartite network with the same (fixed) number of connections per GrC.

other words, for our model the bipartite clustering coefficient measures the average number of shared input pairs between pairs of granule cells. Indeed, this measure is useful in quantifying the local effect of the spatial structure: the local GCL network model turns out to be more clustered than the equivalent random bipartite network, and for both models clustering increases with the number of GrC synaptic inputs (Figure 23).

3.2 A MODEL OF GOLGI CELL PLACEMENT AND CONNECTIVITY PREDICTS STRUCTURAL STATISTICS FOR THE GOLGI CELL NETWORK

3.2.1 *The spatial distribution of Golgi cells shows regularity below the 30 μm scale*

While the average spatial density of cerebellar Golgi cells has been reported in the literature [Vervaeke et al., 2010] and is the subject of ongoing investigations [Valera et al., 2012], no studies have attempted to quantify higher-order properties of the spatial distribution of these interneurons, such as possible correlations between the positions of individual Golgi cells. As will be discussed in subsection 3.2.2, such statistical properties, conceptually related to the tendency of cells to "aggregate" in clusters or to "repel" each other forming regular lattice-like patterns, can have a role in the study of structure-function relationships in the granule cell layer.

To investigate the second-order structure of the spatial distribution of Golgi cells within the granule cell layer, in the context of a collabor-

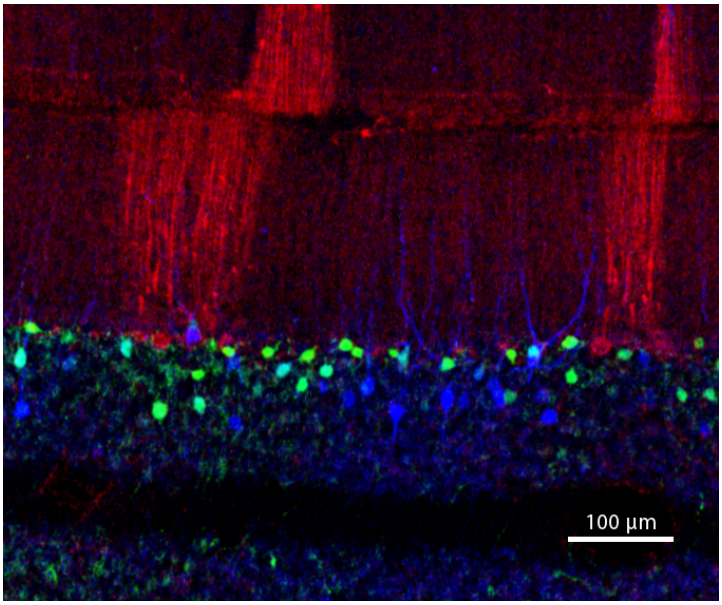


Figure 24: Immunostaining of cerebellar Golgi and Purkinje cells. Detail of the central region in lobule IV. Green: GlyT2-positive Golgi cells. Blue: neurogranine-positive Golgi cells. Red: aldolase-C-positive Purkinje cells. Data kindly provided by Jean-Luc Dupont (INCI-CNRS, Strasbourg, France).

spatial point processes

ation with Antoine Valera I used tools from the theory of *spatial point processes* [Daley and Vere-Jones, 2003; Illian et al., 2008; Diggle, 2013] to analyse experimental data on Golgi cell position kindly provided by Jean-Luc Dupont (INCI-CNRS, Strasbourg, France). Position data was extracted from confocal microscopy images of immunolabelled populations of Golgi cells (Figure 24). The analysis — detailed in subsection 2.1.1 — consisted in estimating the *pair correlation function* $g(r)$ for the experimentally recorded positions of cells within thin transverse sections of lobules II to V. The pair correlation function is defined as follows: given the presence of a cell at a location x, and given another location y, $g(|x-y|)$ is the probability density of finding a cell in y, normalised by the mean spatial density of cells around y (but see subsection 2.1.1 for a more rigorous definition). If the cells were independently distributed according to a spatial density function $\lambda(x)$, $g(r)$ would then be constant and equal to 1 for all values of r. Statistically significant deviations of $g(r)$ from unity represent a tendency of the cells to cluster together (if $g(r) > 1$) or to repel each other (if $g(r) < 1$) at a given distance r. Using this analysis tool I showed that, while Golgi cell density can vary depending on the anatomical location and genetic subtype [Valera et al., 2012], pairs of GoCs are less likely to be found at reciprocal distances smaller than ~30 µm than what would be expected if the cells were indipendently distributed in space (Figure 25).

pair correlation function

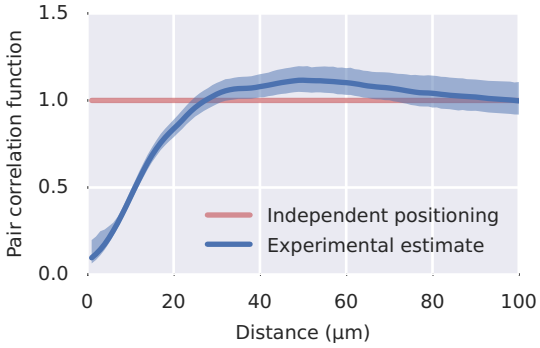


Figure 25: Blue: pair correlation function $g(r)$ for experimentally recorded Golgi cell positions. Shaded region: confidence band for $g(r)$, with 0.95 confidence level. Red: $g(r) \equiv 1$ reference, representing the value $g(r)$ would assume if cell positions were statistically independent.

It should be noted that the pair correlation function of a spatial point process is defined in such a way to be, in a way, independent of the dimensionality of the process, in the sense that $g(r) \equiv 1$ for a completely random process regardless of whether this process is defined on a surface or in a three-dimensional volume. Because of this, even though in my analysis I only considered the two-dimensional positions of Golgi cells in transverse sections, we can expect the pair correlation function of the full three-dimensional distribution of Golgi cells within cerebellar tissue to exhibit a similar behaviour to that presented in Figure 25, if only we assume that, after accounting for possible spatial inhomogeneities (see subsection 2.1.1), the second order structure of the distribution of cell positions is isotropic.

3.2.2 *Estimate of the degree distribution of the Golgi cell network from paired-recording and immunostaining data*

gap junctions

Cerebellar Golgi cells are connected through electrical synapses or *gap junctions* [Dugué et al., 2009]; this connectivity substrate is thought to play a fundamental role in regulating inhibitory synchrony in the granule cell layer [Ostojic et al., 2009; Vervaeke et al., 2010] and in compensating sublinear dendritic integration in Golgi cells [Vervaeke et al., 2012]. To develop further the analysis of synchrony presented in Vervaeke et al. [2010] in the direction of the methods used in the dynamical and complex systems literature [Pikovsky et al., 2003; Arenas et al., 2008], I built a model of the gap junction network that could be seen as a random spatial graph [Barthélemy, 2011] and that is tractable with the tools of network theory [Newman, 2010]. In the model, the nodes of the network (representing Golgi cells) are seen as an instance of a spatial point process [Illian et al., 2008] defined in an infinite layer of finite thickness (representing the granule cell layer),

and the probability of two nodes being connected depends on the distance between them. Similar graph models have been studied in the context of *continuum percolation theory* [Meester, 1996], a branch of stochastic geometry born out of the study of wireless communication networks (see Franceschetti and Meester [2007] or Haenggi et al. [2009] for a review, and [Chiu et al., 2013] for a general introduction to stochastic geometry).

continuum percolation theory

Details of the definition and analysis methods for the GoC network model are given in section 2.1. In brief, the network model is defined in an infinite layer of thickness L. For simplicity, the cells are assumed to be distributed with constant density λ. Cells are connected according to a distance-dependent rule: if two cells are located at a reciprocal distance r, the probability that they are connected is

$$p_c(r) = \frac{a}{e^{(r-r_0)/\Delta} + 1}$$

Parameters L, λ, a, r_0 and Δ are given in Table 1.

Such a structural Golgi network model predicts that the degree distribution of the gap junction network, in the infinite size limit, is Poissonian:

$$p(\deg = k) = \frac{e^{-\alpha}\alpha^k}{k!} \tag{3.38}$$

and that the average degree is

$$\alpha = a\lambda(L\pi r_0^2)\left[1 + \frac{\pi^2}{3}\left(\frac{\Delta}{r_0}\right)^2 - \frac{1}{6}\left(\frac{L}{r_0}\right)^2\right] \tag{3.39}$$

Where λ is the average number of Golgi cells per unit volume, L is the thickness of the granule cell layer, r_0 is a measure of the typical maximum distance between connected GoCs, a is the typical connection probability between two cells that are close together, and Δ is the spatial scale of the transition between what we consider to be "close" and "distant" pairs of cells (for a precise definition of all the terms, see section 2.1). Setting the parameters in Equation 3.39 to values estimated from the experimental data published in Vervaeke et al. [2010] (see subsection 2.1.2 and Table 1), my prediction for the average degree of the network is then

$$\alpha \simeq 15$$

This prediction for the average number of neighbors for the typical cell in the GoC network is significantly higher than the ~10 figure reported in Vervaeke et al. [2010] for a network composed of 45 cells. In Figure 26, I show that the reason for this discrepancy is an edge

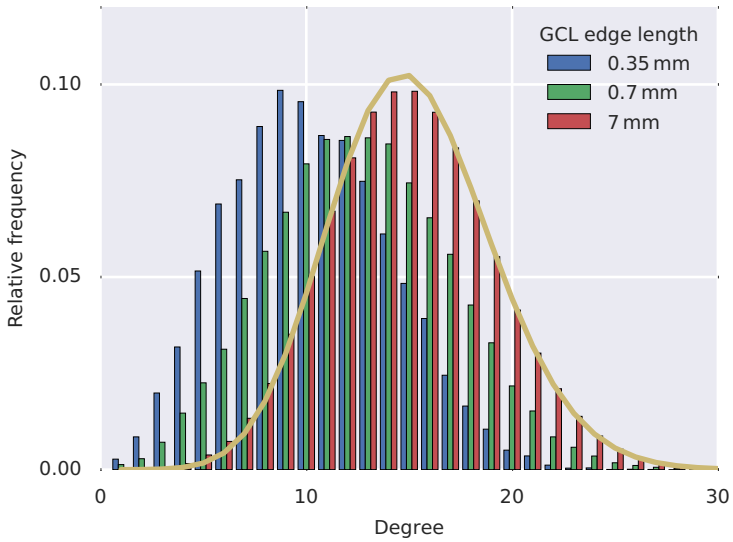


Figure 26: Blue (resp. green, red): average degree distribution for a Golgi network model similar to that in Vervaeke et al. [2010]. Golgi cells are placed at random in a cuboidal box of size 80 µm × 350 µm × 350 µm (resp. 80 µm × 700 µm × 700 µm, 80 µm × 7000 µm × 7000 µm), and are connected with a distance dependent probability given by Equation 2.8. Averages are taken over 400 (resp. 200, 20) intantiations of the model. Yellow: theoretical degree distribution for a network built with the same connection rule in an infinite layer of thickness 80 µm (Equation 3.38 and 3.39).

effect in the original estimate: indeed, on average a network of that size will have a mean degree of about 10. But if we take the same connectivity rules and we apply them to networks of progressively larger sizes, the degree distribution converges to the theoretical result given above for an infinitely large network confined to a sheet of finite thickness (Equation 3.38 and 3.39). Thus, my analytical prediction is corroborated by numerical estimation of the degree distribution on random instantiations of the connectivity model.

3.2.2.1 *Effect of spatial regularity of GoC positions*

As I have shown in subsection 3.2.1, Golgi cells are relatively unlikely to be found at reciprocal distances smaller than 30 µm. In principle, this could clearly have an effect on the network connectivity statistics. In practice, though, for the typical cell in the connectivity model discussed above, most of the connected neighbours end up being located at distances significantly larger than 30 µm, so that their mean number is altered only slightly by taking into account the measured regularity of Golgi cell positions (subsection 2.1.3). Nevertheless, it is still possible that the geometrical disposition of GoCs could significantly impact network dynamics. In fact, the coupling strength between GJ-coupled Golgi cells has been shown to decay with distance on a

shorter spatial scale than the connection probability [Dugué et al., 2009; Vervaeke et al., 2010].

3.3 DISCUSSION

In this chapter, I have shown how the spatial nature of the connectivity rules can give rise to statistical network features in the main electrical and chemical networks of the cerebellar granule cell layer.

The results of section 3.1 provide a possible criterion to find the very minimum spatial scale at which information processing in the GCL should be analysed, by highlighting the fact that neighbouring granule cells necessarily have common sources of synaptic input. The effects of spatial structure on coding in the GCL will be investigated further in chapter 5.

In subsection 3.2.1, the first analysis to be conducted on the higher-order properties of the spatial distribution of Golgi cells is presented. I have found that GoCs are unlikely to be located at reciprocal distances smaller than 30 µm, giving an empirical grounding to the assumptions on spatial regularity in existing models of the GoC network [Dugué et al., 2009] and opening new possibilities for inquiry into the relationship between its structure and its function.

On the other hand, the analytical results in subsection 3.2.2 allow us to estimate the mean value and the variance of one of the most important "connectedness" metrics for the Golgi cell network — namely, the number of other cells that any Golgi cell is coupled to — from simple geometrical considerations. Furthermore, the experimental data that this estimate relies upon is relatively simple and easy to obtain, as only paired recordings for cells at a known distance from each other are needed, making for better statistics if compared to lower-throughput experimental techniques, like morphological reconstructions.

Finally, the results in section 3.2 can also be used as a way of checking the internal consistency of existing experimental data on GoCs with the conceptual model used to interpret it. In fact, if the average number of GJs per Golgi cell is 35 [Vervaeke et al., 2012] and the average number of connections per cell is 15 (subsection 3.2.2), on average there must be about 2.3 gap junctions per connected pair. Indeed, this theoretical prediction matches well with current experimental estimates obtained by direct count of the number of GJs in connected GoC pairs by electron microscopy [Lőrincz et al., 2015].

4

NETWORK STRUCTURE REGULATES SYNCHRONY OF GOLGI CELL FIRING

As described in subsubsection 1.3.4.2, theta band (resp. beta band) LFP oscillations have been observed in the GCL in rat (resp. monkey); behaviourally, they are associated with a state of "quiet wakefulness" of the animal, and they disappear upon initiation of a movement [Hartmann and Bower, 1998; Pellerin and Lamarre, 1997]. Synchronous firing of Golgi cells is thought to be one of the main contributors to the generation of this oscillatory pattern [Dugué et al., 2009; Roš et al., 2009], and to be enabled by the gap junction-mediated recurrent connectivity existing within the cell population [Dugué et al., 2009]; indeed, electrical connectivity has been shown to enable synchronous activity in other neuronal populations [Beierlein et al., 2000; Deans et al., 2001; Traub et al., 2001].

The mechanisms governing the robustness of synchronous activity in the Golgi cell network are not yet completely understood. Theoretical studies have shown that a simplified, fully-connected network model can exhibit bistability, toggling between synchronous and asynchronous activity in response to an external stimulus [Ostojic et al., 2009], and *in vivo* experiments have shown that gap junctions mediate rapid desynchronisation of Golgi cell activity following a temporally precise distributed synaptic input [Vervaeke et al., 2010]. Furthermore, simulations of an experimentally constrained model of a local GoC network have highlighted the importance of *local* and *heterogeneous* GJ-mediated coupling in enabling desynchronisation [Vervaeke et al., 2010]. However, a systematic study of the robustness of the synchronous state in a network with realistic (i.e. space-dependent) connectivity and detailed cell models has never been realised.

To better understand the role of network structure in controlling the robustness of synchronous GoC network activity, I therefore created new experimentally constrained network models that allow the spatial range of GJ connectivity to be manipulated independently from the distribution of synaptic conductances between coupled cells. As I will illustrate below, this allowed me to disentangle the relative contributions of the two aspects by simulating the network and measuring its response to an external stimulus.

4.1 SPATIAL LOCALITY OF THE NETWORK

I modelled a change in locality of the network connectivity rules in a number of ways, detailed below. In general, I found that manipulating the network structure to make it less local gradually enforces a very robust synchronous activity state.

4.1.1 *Variable connectivity range*

In the network model used by Vervaeke et al. [2010], $N = 45$ Golgi cells are scattered at random in a $350\,\mu m \times 350\,\mu m \times 80\,\mu m$ box, to match their experimentally measured average spatial density. Each cell pair, located at a reciprocal distance r, is then connected with a GJ with probability $p_c^V(r)$ (Equation 2.6). The conductance of the GJ is also dependent on r through a relationship $g^V = g^V(r)$, the specific form of which is given in the paper. The endpoints of the synapses are placed at random on the dendritic tree of the connected cells.

For my simulations, I used this model as a starting point. Even though in subsection 3.2.2 I showed that a network this size suffers significant small size and border effects on its structural properties, in most simulations I kept $N = 45$ to allow for a more direct comparison with the results reported by [Vervaeke et al., 2010] (but see below for tests done on larger networks with a different connectivity scheme). To change the connectivity range of the cells and make them more likely to connect to other cells further away I modified the connectivity model in the following way:

$$p_c^V(r) \longrightarrow p_c(r;s) = p_c^V(r/s)$$
$$g^V(r) \longrightarrow g(r;s) = g^V(r/s)$$

where s is a spatial scale factor. This results in an increase in the number of connections; indeed, Equation 3.39 shows that the increase in the average number of connections per cell would be proportional to s^2 for an infinite-sized network if we modelled $p_c(r)$ with Equation 2.8 instead of Equation 2.6. In each random realisation of the network, I compensated for this by scaling down the synaptic conductances to keep their total sum within the network at the same value that would have been obtained by using $p_c^V(r)$ and $g^V(r)$.

I performed numerical simulations using the network model detailed above and the reduced-morphology Golgi cell model described in subsection 2.2.2, following the same protocol as in Vervaeke et al. [2010]. All cells in the network were subject to a realistic MF and PF background noise stimulation [Rancz et al., 2007; Chadderton et al., 2004] distributed over the dendritic tree (details on the synaptic models are given in subsection 2.2.5 and subsection 2.2.6). For each cell, 20

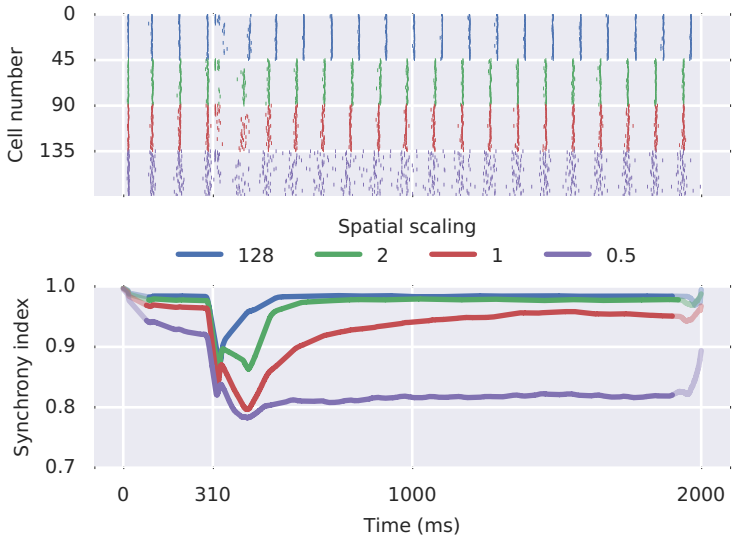


Figure 27: Desynchronisation of Golgi cell network models with varying spatial connectivity scales following a temporally precise synaptic input. *Top:* representative raster plots of network activity for the different models. Cells are numbered progressively through the four network models analysed. In each network, 10 cells receive a brief external synaptic input starting at 310 ms after the beginning of the simulation, following the protocol used by Vervaeke et al. [2010] (see main text). *Bottom:* time-resolved synchrony index for the population (Equation 4.40), averaged over 32 repetitions of the simulation, with a different randomly generated network each time. Values affected by artefacts due to the definition of the synchrony index (which converges to 1 at the beginning and the end of each simulation) are displayed in a muted colour.

(resp. 100) synaptic stimuli were placed at random locations on the basolateral (resp. apical) dendrites, each firing at 2 Hz (resp. 0.5 Hz) with Poisson statistics. Additionally, a random subset of 10 cells received a temporally precise external synaptic input roughly 310 ms after the beginning of the simulation, out of phase with their spontaneous firing pattern. The stimulus consisted in a very brief sequence of MF and PF synaptic events. For each stimulated cell, 8 (resp. 50) synaptic locations randomly located on the basolateral dendrites were stimulated at 200 Hz (resp. 350 Hz) for a period of 10 ms (resp. 15 ms). Apical synaptic events were delayed by an additional 5 ms to reflect synaptic integration time in granule cells.

Spike times were recorded for every cell in the network; an illustrative scatter plot is given in Figure 27, top panel. At any time t, synchrony within the population of N cells was measured as

$$S(t) = 1 - D(t) = 1 - \frac{1}{N(N-1)/2} \sum_{n=1}^{N-1} \sum_{m=n+1}^{N} D^{mn}(t) \qquad (4.40)$$

where $D^{mn}(t)$ is the *bivariate SPIKE-distance* [Kreuz et al., 2013] between the activity of cell m and cell n. The SPIKE-distance is a time-resolved and time scale-adaptive measure of spike train dissimilarity, and $D(t)$ is its averaged extension to a population of size N. Analysis of the simulations shows that networks with long range connectivity immediately fall back in synchrony after the occurrence of the external stimulus, while networks with short range connectivity tend to persist for a longer time — or even indefinitely — in a partially desynchronised state (Figure 27, bottom panel).

SPIKE-distance

4.1.2 *Tunable locality with Watts-Strogatz-like network models*

In the simulations discussed above, changing the spatial scale of the connectivity has the side effect of altering the number of neighbours of each cell in the network. As a minimal way of modelling a change in locality with no such side effect, I built a Golgi network based on the "small world" network model by Watts and Strogatz [1998]. This consists in $N = 720$ cells arranged in a k-ring topology, where each of them is connected to its $k = 14$ nearest neighbours, but where a fraction p of the $720 \cdot 14/2$ edges is "rewired" at random (Figure 28). These connections represent long range shortcuts between areas of the network that would otherwise be spatially segregated; the *rewiring probability* p thus parametrises a smooth transition between a regular, completely local k-ring ($p = 0$) and a Erdős–Rényi random graph of the $G(n, m)$ family [Erdős and Rényi, 1959; Newman, 2010] ($p = 1$). The value of k was chosen to approximate the predicted average for an infinitely large network (subsection 2.1.2); $k = 14$ was selected in-

Watts-Strogatz network model

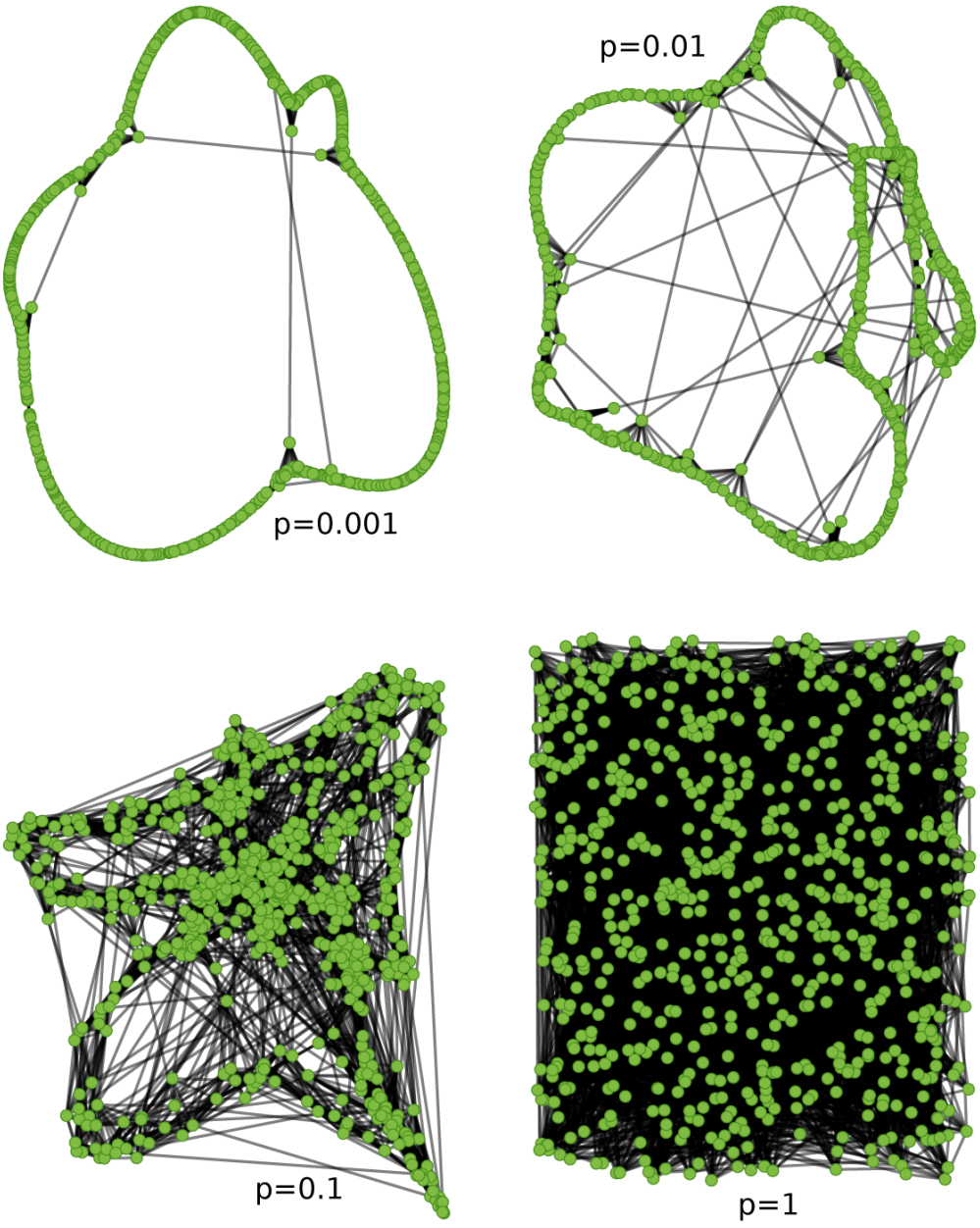


Figure 28: Examples of Watts-Strogatz connectivity substrates for the Golgi cell network model with tunable locality, for different values of the rewiring probability parameter p. In all four cases, there are $N = 720$ cells and $k = 14$ connections per cell. Note that most edges are not visible in the $p = 0.01$ and $p = 0.001$ cases, and that the case $p = 1$ represents a random graph of the $G(n, m)$ family.

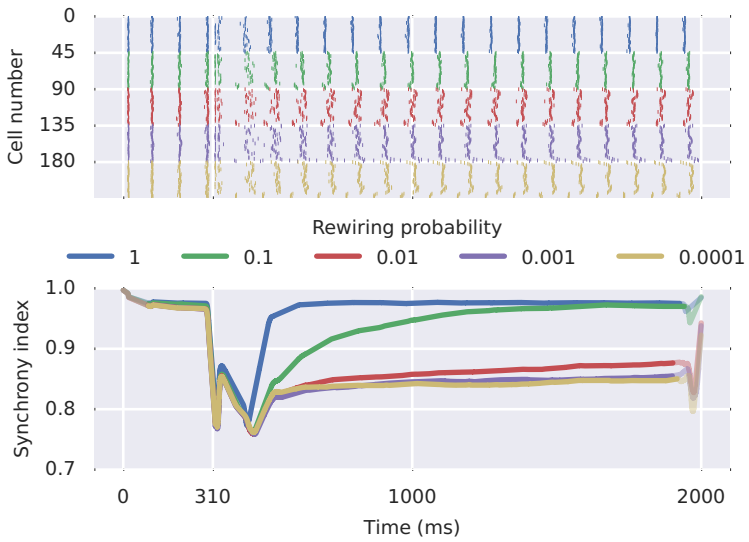


Figure 29: Desynchronisation of Golgi cell network models with tunable locality, built from a Watts-Strogatz-like small world network model. *Top:* representative raster plots of network activity for a random selection of 45 cells for each model tested. Cells shown are numbered progressively through the five network models analysed. Out of the total 720 cells composing the model, 160 cells uniformly distributed around the ring receive a brief external synaptic input starting at 310 ms after the beginning of the simulation, following the protocol used by Vervaeke et al. [2010] (see main text). A video illustrating a representative simulation is available online [http://dx.doi.org/10.6084/m9.figshare.1299068]. *Bottom:* time-resolved synchrony index for the population (Equation 4.40), averaged over 3 repetitions of the simulation, with a different randomly generated network each time. Values affected by artefacts due to the definition of the synchrony index (which converges to 1 at the beginning and the end of each simulation) are displayed in a muted colour.

stead of $k = 15$ to satisfy the k-ring symmetry requirement for k to be even. N was made large enough to ensure a sufficiently smooth small world transition [Watts and Strogatz, 1998].

Simulations for this parametrised model show again that a very robust synchronous activity state can be enforced by making the network less local (see Figure 29 and the accompanying video[1]). Indeed, the transition between a desynchronisable and a non-desynchronisable architecture happens alongside the well known small-world transition for the network structure of the underlying network model [Watts and Strogatz, 1998; Newman, 2010] (Figure 30).

Finally, as an additional control, I have obtained a network model with no spatial structure by taking an instance of the model by [Vervaeke et al., 2010] and randomly shuffling the edges, such that spatial correlations were destroyed but the degree distribution and the mar-

1 http://dx.doi.org/10.6084/m9.figshare.1299068

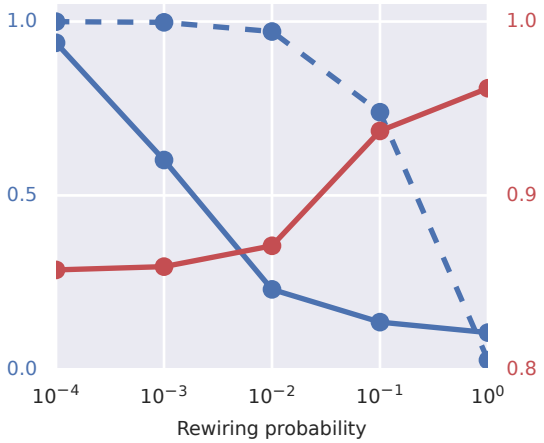


Figure 30: Illustration of the loss of locality during the small-world transition in a Watts-Strogatz-like Golgi cell network model. Dashed blue line: $C(p)/C(0)$, where $C(p)$ is the (unweighted) mean clustering coefficient of the network as a function of the rewiring probability parameter p. Solid blue line: $L(p)/L(0)$, where $L(p)$ is the mean minimum (unweighted) path length between two nodes of the network, as a function of p. Red: temporal average of the synchrony index over the whole duration of the simulations shown in Figure 29 (2000 ms).

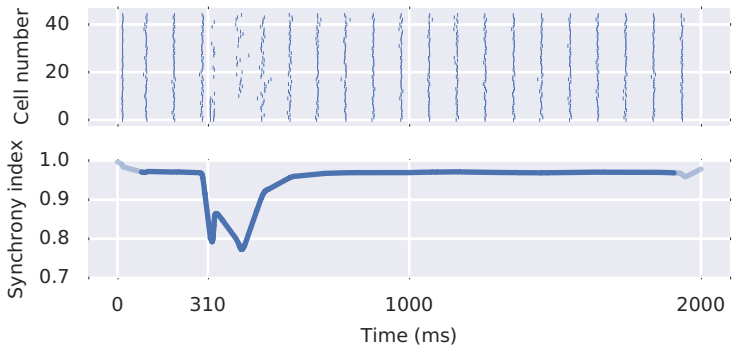


Figure 31: Desynchronisation of a spatially random Golgi cell network model following a temporally precise synaptic input. The network is obtained from that used in Vervaeke et al. [2010] by shuffling all connections between cells. *Top:* representative raster plot of network activity. 10 cells receive a brief external synaptic input starting at 310 ms after the beginning of the simulation, following the protocol used by Vervaeke et al. [2010] (see main text). *Bottom:* time-resolved synchrony index for the population (Equation 4.40), averaged over 32 repetitions of the simulation, with a different randomly generated network each time. Values affected by artefacts due to the definition of the synchrony index (which converges to 1 at the beginning and the end of each simulation) are displayed in a muted colour.

ginal distribution of synaptic weights were conserved (the degree distribution is conserved because, as I showed in section 3.2, the distribution in the spatial case is Poissonian, just like that of an Erdős–Rényi random graph). This model also displays a robust synchronous behaviour, similar to that exhibited by the completely randomised ($p = 1$) Watts-Strogatz network (Figure 31).

4.2 HETEROGENEITY OF SYNAPTIC STRENGTHS

To investigate the effect of the heterogeneity of synaptic strengths within the network, I have created a variant of the model by Vervaeke et al. [2010] where the spatial structure of the connectivity is conserved, but the conductance of the GJs between connected cells is not given by $g^V(r)$ as discussed above. For each synapse, it is instead drawn independently from a Gamma distribution with the same mean of the empirical distribution of synaptic weights in the original model but whose variance is scaled systematically by a factor ν. The Gamma distribution $\Gamma(k, \theta)$ is parametrised by a *shape parameter* k and a *scale parameter* θ, and it is defined by its probability density function in the following way:

$$f(g; k, \theta) = \frac{g^{k-1} e^{-g/\theta}}{\theta^k \Gamma(k)}$$

where $\Gamma(k)$ is the Gamma function evaluated in k. $\Gamma(k, \theta)$ has mean $k\theta$ and variance $k\theta^2$, so performing the following transformation

$$k \longrightarrow k/\nu$$
$$\theta \longrightarrow \theta \cdot \nu$$

achieves the desired scaling of the variance while keeping the mean constant. To choose a reference value for k and θ in the $\nu = 1$ case, I fixed $k\theta$ to the empirical average of the synaptic weights in the original model and fitted the shape parameter k to give the best agreement with the full empirical distribution (Figure 32). This procedure yielded $k = 1.53$ and $\theta = 0.35\,\text{nS}$.

While increasing the heterogeneity of the synaptic weight distribution renders the synchronous activity state more easily disrupted by external synaptic input, the opposite is true only to a certain extent. In fact, there is an upper limit on how synchronous a network we can get by reducing the variability in synaptic strengths between connected cells. This is shown in figure Figure 33, where simulation results for networks with $\nu = 1$, 4 and 9 are compared with the $\nu = 10^{-6}$ case, essentially equivalent to setting all synaptic weights equal to the mean value of the empirical distribution.

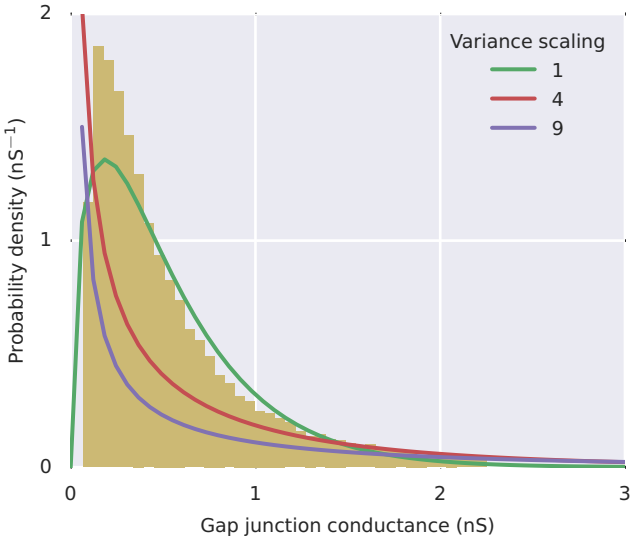


Figure 32: Histogram: empirical distribution of synaptic conductances in the model by Vervaeke et al. [2010], averaged over 100 model realisations. Solid lines: re-scaled Gamma distributions (see main text) for different values of the variance scaling parameter ν.

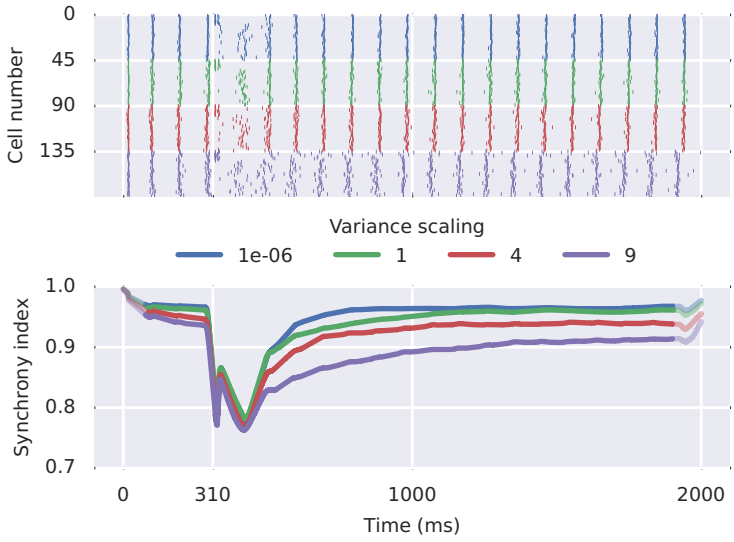


Figure 33: Desynchronisation of Golgi cell network models with varying synaptic weight heterogeneity following a temporally precise synaptic input. *Top:* representative raster plots of network activity for the different models. In each network, 10 cells receive a brief external synaptic input starting at 310 ms after the beginning of the simulation, following the protocol used by Vervaeke et al. [2010] (see main text). *Bottom:* time-resolved synchrony index for the population (see main text), averaged over 32 repetitions of the simulation, with a different randomly generated network each time. Values affected by artefacts due to the definition of the synchrony index (which always converges to 1 at the beginning and the end of each simulation) are displayed in a muted colour.

4.3 DISCUSSION

In this chapter, I have studied the stability of synchronous activity in experimentally constrained models of the network formed by cerebellar Golgi cells through electrical synapses. Following the finding by Vervaeke et al. [2010] that desynchronisation is enabled by a *local* and *heterogeneous* network structure induced by a space-dependent connectivity rule, I analysed the relative contribution of these two aspects to the robustness of synchronous activity to external input. Using numerical simulations, I have quantified the response of the network to a temporally precise synaptic input to ~20 % of the cells, while systematically varying the model parameters governing the locality of the connectivity structure and the heterogeneity of synaptic strengths.

I have found that heterogeneity of the distribution of synaptic conductances within the network has a tangible effect on the ability of the system to desynchronise; however, this effect is dominated by that of the spatial scale of the synaptic organisation. Indeed, networks with long range connections support a very robust synchronous state, and spatial segregation in locally connected networks enables the persistence of a partially desynchronised state. Conceptually, this matches well known results from the theory of complex networks of dynamical systems [Barahona and Pecora, 2002; Arenas et al., 2008].

The results I presented here constitute a significant advancement over the existing literature, as previous theoretical studies on networks of electrically coupled neurons have focused on fully-connected topologies [Ostojic et al., 2009], while in Vervaeke et al. [2010] locality was not considered as an effect of the realistic connectivity structure distinct from heterogeneity, and structural parameters were not investigated systematically. In the future, this work could be extended by formulating a model of network connectivity with tunable locality that would be more realistic than that used in subsection 4.1.2 — for instance by being based on a 2D lattice instead of a k-ring — but at the same time simple enough to be amenable to analytical investigation [Pecora and Carroll, 1998; Barahona and Pecora, 2002; Pereira et al., 2013]. This would allow for the results of the structural analyses in section 3.2 to be directly related to their dynamical and functional implications. A further interesting aspect to study would be the spatial scale at which GJ connectivity can establish and maintain robust synchrony in absence of long-range connections or common oscillatory drive. Finally, another question is the effect that recurrent chemical inhibitory synapses, recently reported to exist within the Golgi cell population [Hull and Regehr, 2012], can have on synchrony.

5

NETWORK STRUCTURE CONTROLS A SPARSIFICATION-TRANSMISSION TRADEOFF IN THE GRANULE CELL LAYER

There is wide consensus that the cerebellar input layer, or granule cell layer (GCL), transforms mossy fibre (MF) inputs, conveying sensory and efferent copy information, into a higher dimensional, sparser code [Marr, 1969]. This type of operation, called *expansion recoding*, increases the separation between input patterns [Olshausen and Field, 2004], thereby enabling downstream cerebellar circuits to perform more effective associative learning [Marr, 1969; Albus, 1971; Tyrrell and Willshaw, 1992; Medina and Mauk, 2000; D'Angelo and De Zeeuw, 2009; Schweighofer et al., 2001], adaptive filtering [Fujita, 1982] and binary addressing [Kanerva, 1988]. Three basic properties are required for divergent feedforward networks to perform effective pattern separation: (1) information must be conserved, (2) the dimensionality of the output coding must be larger than that of the input, and (3) the output code must be sparse. However, the contribution that synaptic connectivity makes to these functions remains poorly understood.

expansion recoding

To understand the role of network structure in determining the effectiveness of the GCL as an expansion recoder, I constructed a biologically detailed network model of a local GCL network where each parameter was constrained by experimental measurements (section 2.3). I then analysed how the expansion recoding performance of the model changed as I altered key structural parameters of the network away from the experimentally measured values.

In my analysis, I assumed cerebellar input signals to be encoded in the firing rates of mossy fibres (MFs), and I modelled GrCs as simple conductance-based integrate-and-fire neurons. These simplifying assumptions are justified by the fact that the soma and dendrites of GrCs form a single electrical compartment, thereby acting as point neurons [Silver et al., 1992, 1996], while much of their excitatory drive is composed of slow spillover-mediated AMPAR and NMDAR conductances that build up over time during rate-coded MF input [DiGregorio et al., 2002; Arenz et al., 2008; Schwartz et al., 2012].

NETWORK MODEL My network model, derived from anatomical experimental data [Billings et al., 2014], represented a ball of tissue with a radius of about 80 µm and contained about 500 granule cells (GrCs), each receiving synaptic inputs from d of the ~180 mossy fibre synaptic rosettes (MF) within the spherical volume. The parameter

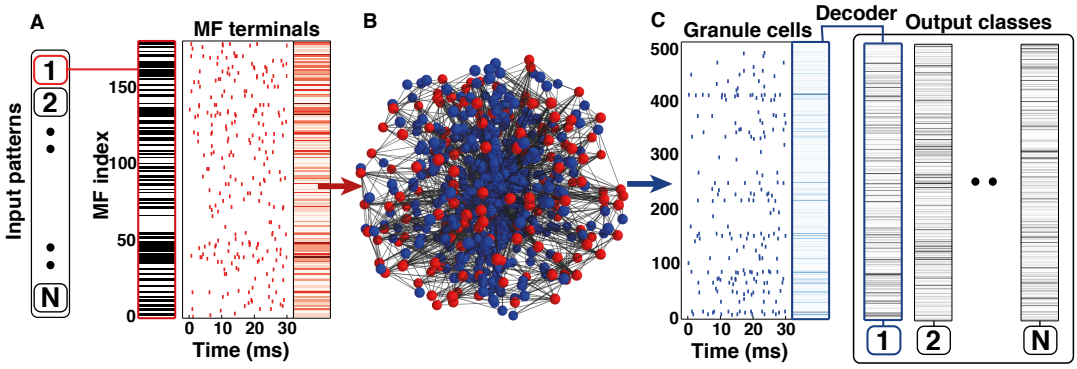


Figure 34: (*Adapted from Billings et al. [2014]*) Analysis of an experimentally constrained spiking model of the local granule cell layer network. *A:* A binary stimulus pattern was randomly selected from a set of N patterns (black active and white inactive on barcode). A Poisson spike train was generated for each MF input (80 Hz active, 10 Hz inactive; red raster plot), thereby setting the timing of synaptic events. Red barcode indicates spike counts for the given realization of the spike trains. *B:* 3D view of the anatomically constrained local GCL network model detailed in subsection 2.3.1, with 176 MFs in red and 509 GrCs in blue. *C:* Raster plot of GrC firing activity in response to the input. Blue barcode indicates GrC spike count vector (measured over a 30 ms window), which was assigned (or *decoded*) to one of N output classes (black bar codes) defined using the k-means algorithm on a separate data set.

d was constant within a network instantiation, but varied across networks. For each network configuration analysed below, I used a fixed instantiation of the connectivity, randomly generated according to the rules given in subsection 2.3.1. See ibid. for details on the network model.

SIMULATION Rate-coded binary input patterns were created by randomly assigning one of two possible firing rates (10 or 80 Hz) to each MF, representing quiescence or activity. The fraction $p(\mathrm{MF})$ of active inputs represented the global activity level of the input to the network. For each of 20 values of $p(\mathrm{MF})$ spanning the $(0, 1)$ range and for each MF, N=1024 independent Poisson trains of synaptic events were then generated, each with the appropriate rate for the given MF and a duration of 30 ms (Figure 34A). GrCs, modelled as conductance-based integrate-and-fire neurons, integrated the synaptic events they received from the d MFs connected to each of them, and responded by firing a certain number of action potentials during the 30 ms time window (Figure 34B and 34C). The spike count for all GrCs was then stored in a vector of the appropriate size. See subsection 2.3.2 for details on the simulations.

ANALYSIS The spike count vectors were the sole output of the simulations I considered for the analysis. The vectors were classified in one of N output classes (Figure 34C), and I analysed the resulting

mapping of N input patterns to N output classes as a communication channel. For a range of values of d and p(MF), I estimated the mutual information (MI, Cover and Thomas [2006]) between input and output assuming a uniform prior over the N input patterns, comparing it with the upper bound of $\log_2 1024 = 10$ bits set by the entropy of the input. Furthermore, for each input pattern I computed the corresponding output population sparseness using the definition given by Vinje and Gallant [2000]. See section 2.4 for details on the analysis of the simulations.

5.1 THE LOCAL ORGANISATION OF THE GCL IS WELL SUITED TO LOSSLESS SPARSE ENCODING

I first examined how effectively spiking network models with different synaptic connectivity transmitted spatially independent MF patterns (Figure 35A). Networks with few synaptic connections per neuron (small d) were most effective at transmitting information across the widest range of MF input activity, defined as the fraction of MF inputs active p(MF) (Figure 35B). Indeed, networks with few inputs recovered almost all of the maximum of 10 bits, which is remarkable, given the noisy nature of the encoding and that GrC spikes were only decoded over 30 ms. However, for networks with larger numbers of synaptic connections per neuron information transmission performance decreased across large regions of p(MF). Reducing or extending the window over which spikes were decoded or altering the firing rates of active MFs shifted the dependence of MI on connectivity, but the overall relationship remained the same (Figure 36 and 37). In contrast to information transmission, GrC population sparseness tended to increase with the number of MF inputs per GrC (Figure 35C).

In agreement with the theoretical predictions made in Billings et al. [2014], these results show that biologically detailed spiking networks with few synaptic connections per neuron are most effective at transmitting information, and that *network connectivity controls a trade-off between information transmission and sparsification*. This trade-off arose because as the number of synaptic connections per GrC increased, the network I-O relationship became highly nonlinear, reducing the average number of spikes per GrC to levels well below the average numbers of spikes per MF (Figure 35G). Although encoding in this region was sparse, so few GrCs were activated that information transmission was compromised. To find the best-performing biological network configurations, I plotted the relationship between the average fraction of information recovered from the GrCs spike counts versus average sparsification performed by each network (Figure 35H). Networks with two to seven synaptic connections per GrC provided the

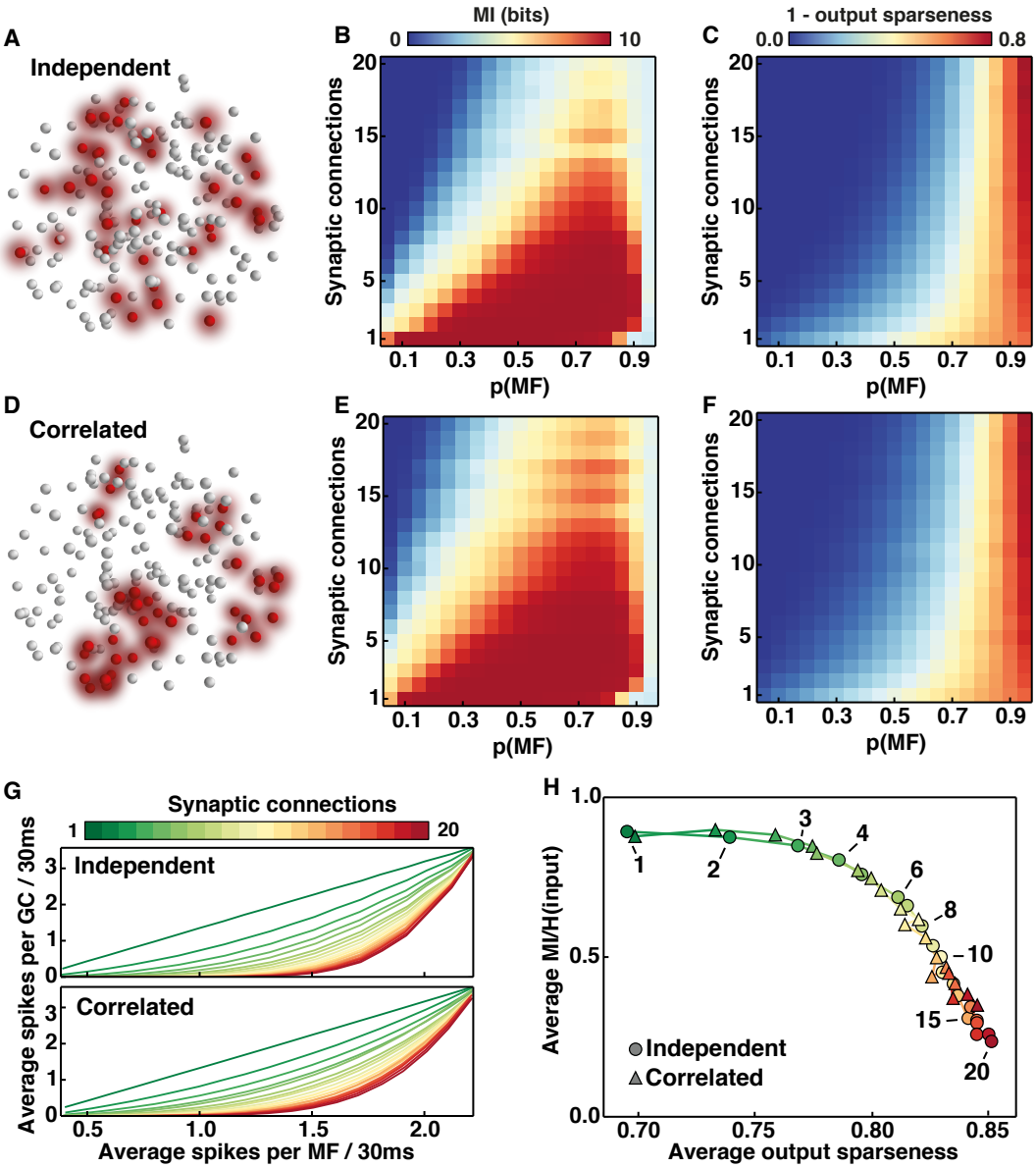


Figure 35: (*Adapted from Billings et al. [2014]*) Sparse encoding in biologically detailed spiking network local models of the GCL with different numbers of synaptic connections per granule cell. *A:* Visualization of independent mossy fiber (MF) inputs in the local granule cell (GrC) layer network model with active MFs in red and inactive MFs in white, for an example random activation pattern. *B:* Mutual information (MI) encoded by the GrC population for 1024 uncorrelated input patterns across the full range of MF activation probability $p(MF)$ in biologically detailed spiking networks with different numbers of synaptic connections per GrC (d). *C:* Same as for B but for 1-average output sparseness. *D, E and F:* Same as for A, B, and C but for a set of 1024 spatially correlated patterns, where neighboring MF inputs were activated in groups of five. *G:* Same as for C and F but visualized as a line plot to show the relationship between average spikes per GC and average spikes per MF in a 30 ms window across all values of $p(MF)$, for networks with different d. *H:* Relationship between average MI (normalized by the MF input entropy) and average output sparseness (across all values of $p(MF)$) for spiking networks with different d (color code) for independent (circles) and spatially correlated (triangles) inputs.

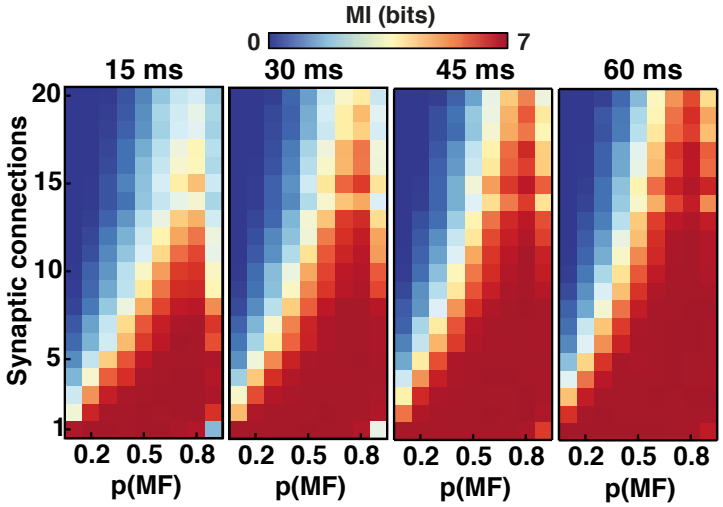


Figure 36: (*Adapted from Billings et al. [2014]*) Effect of the duration of the time window for the integration of network activity on spiking granule cell layer network performance: mutual information (MI) for 128 patterns, estimated after integrating network activity within time windows of different length.

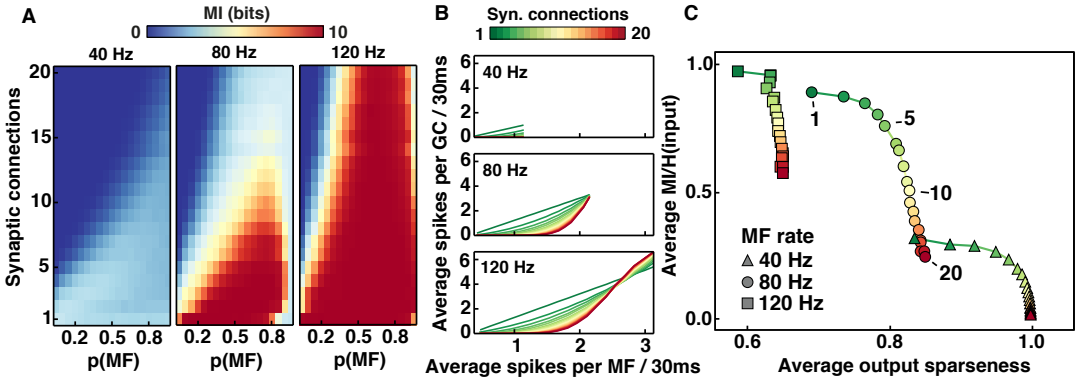


Figure 37: (*Adapted from Billings et al. [2014]*) Effect of different mossy fiber input rates on spiking granule cell layer network performance. *A:* Mutual information (MI) for 1024 input patterns for active mossy fibres (MFs) firing at different rates, with tonic inhibition kept fixed at the average experimental value. *B:* average spikes per granule cell (GrC) versus average spikes per MF within a 30 ms window, across all patterns and all values of the probability of MFs being active $p(MF)$. Different colour lines show relationships for networks with different numbers of synaptic connections. *C:* Average MI, normalized by the input entropy, versus average output sparseness (with averages taken across all values of $p(MF)$) for different active MF rates, parametrised by the number of synaptic connections. Note how, as shown in at the bottom of panel B, there exists a threshold level of MF activity above which the relationship between number of synaptic connections and GC activity is reversed, and the most active networks are those with more connections. This is reflected in panel C, as the networks with many synaptic connections lose the sparsification advantage they have when the inputs are encoded at lower rates, leaving those with fewer connections as an optimal choice for lossless sparse encoding, even though the maximum sparsification attainable is lower.

best solution for performing sparse lossless encoding, with four inputs performing particularly well.

5.1.1 *Dependency of network performance on spatial structure of inputs and network connectivity*

Spatial correlation in MF input activity is likely to occur in real cerebellar networks and may vary from region to region, due to variations in the numbers of MFs arising from different origins [Huang et al., 2013]. I examined how spatial correlations in the MF activity patterns affected information transmission and sparsification in spiking networks by introducing a pronounced spatial correlation in the input, co-activating groups of five neighbouring MFs (Figure 35D). Input correlations marginally increased the information transmission performance over all networks (cf. Figure 35B and 35E). This is due to GrCs in a particular region receiving a larger fraction of active MF inputs and thus having a greater chance of generating a spike. Output sparseness of the GrC population was little affected by spatial correlations in the MF input activity (cf. Figure 35C and 35F). Thus, spatial correlations in MF activity only subtly shift the trade-off between information transmission and sparsification in biologically detailed spiking networks.

I also explored how the spatial structure of the network discussed in section 3.1 affects information transmission by creating a spatially randomised variant of the network model and using it to repeat some of the analyses above (Figure 38). In a local network, the spatial organisation of the connectivity has an adverse, albeit minor, effect on performance. This can be interpreted as a result of the input sharing phenomenon documented in section 3.1, whereby neighbouring GrCs tend to sample from similar sets of MFs.

5.2 A MODEST AMOUNT OF FEEDFORWARD INHIBITION ENHANCES EXPANSION RECODING

Golgi cells provide network-activity-dependent inhibition of GrCs, via feed-forward [Kanichay and Silver, 2008] and feedback inhibition [Cesana et al., 2013], although the impact of this phasic and spillover-mediated component is much weaker than tonic inhibition [Duguid et al., 2012]. Therefore, I tested whether adding network-activity-dependent inhibition to the physiological level of tonic inhibition present in the model improved the performance of my spiking networks. To do this, I scaled the tonic inhibitory conductance as a function of $p(\mathrm{MF})$ from the experimentally measured tonic level. This reduced information transmission in networks with larger numbers of synaptic connections (Figure 39A). However, for networks with few

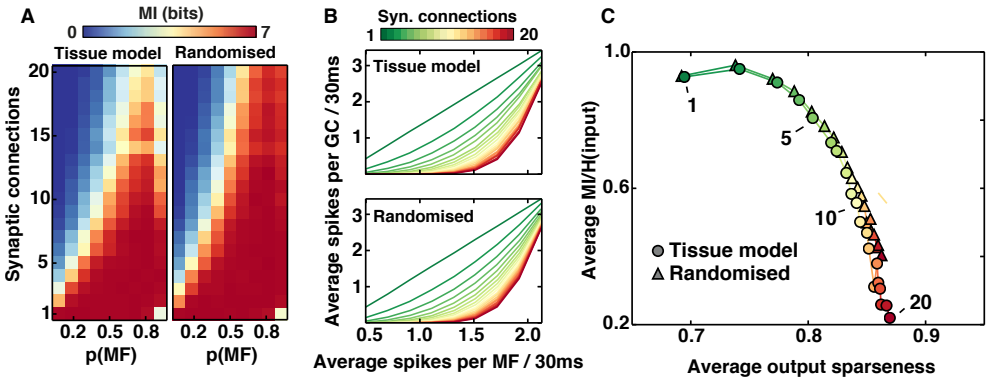


Figure 38: Performance comparison between the anatomical tissue model and a randomised network with the same number of synaptic connections per GrC. *A:* Mutual information (MI) for 128 patterns. *B:* average spikes per GC versus average spikes per MF across all patterns and all values of the probability of MFs being active $p(\mathrm{MF})$, parametrised by the number of synaptic connections. *C:* Average MI, normalised by input entropy, versus average output sparseness (with averages taken across all values of $p(\mathrm{MF})$), parametrised by the number of synaptic connections.

synaptic connections per neuron, information transmission was preserved and GC activity was further sparsened. Biologically detailed networks with three to five synaptic connections, tonic inhibition, and a modest network-activity-dependent inhibition performed lossless sparse encoding better than with tonic inhibition alone (Figure 39C), supporting the theoretical predictions in Billings et al. [2014]. Interestingly, all three of these features are characteristic properties of the cerebellar GCL, suggesting that both the connectivity and the inhibition properties of the system are tuned to enable robust lossless sparse encoding over the widest possible range of MF excitatory drive.

5.3 THE DISTRIBUTION OF ELECTROPHYSIOLOGICAL CHARACTERISTICS ACROSS A GRANULE CELL POPULATION IS LOW-DIMENSIONAL

The biophysical properties of GrCs and of their synaptic inputs from MFs have been shown to be highly heterogeneous [Schwartz et al., 2012; Sargent et al., 2005], and heterogeneity of neuronal properties is known to have an important role in population coding [Shamir and Sompolinsky, 2006]. As a first step towards incorporating cell- and synapse-level heterogeneity in my GCL model, I investigated the distribution of electrophysiological parameters within a rich experimental dataset [Ward, 2012], building a generative model for populations of conductance-based integrate-and-fire models of GrCs.

As I showed in subsection 2.2.8 via formal model selection, it is reasonable to approximate the distribution of the main electrophysiolo-

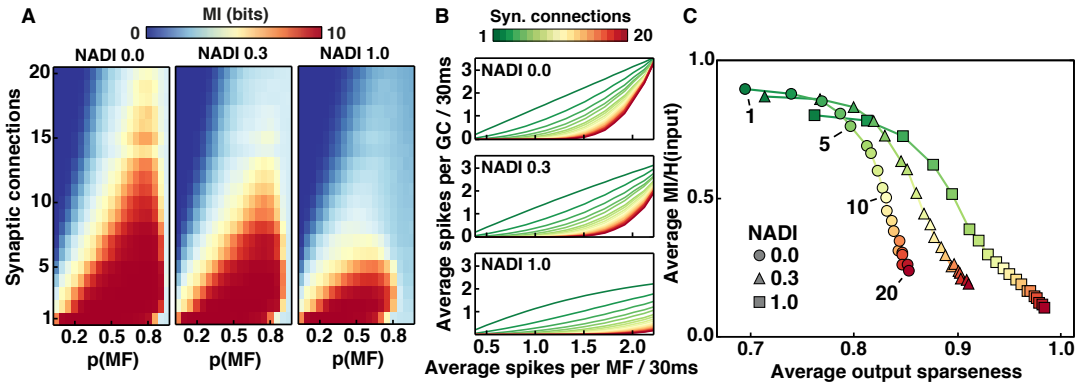


Figure 39: (*Adapted from Billings et al. [2014]*) Effect of network activity-dependent scaling of tonic inhibition (NADI) on granule cell layer network performance. *A:* Mutual information (MI) for 1024 patterns of different active mossy fibers (MFs) with different levels of scaling of tonic inhibition in granule cells (GrCs) with network activity added on top of the physiological baseline; heatmap on the left (NADI=0) is included for comparison and represents the baseline case. In these simulations an extra amount of $GABA_A$R-mediated conductance, proportional to the fraction of active MFs $p(MF)$, was introduced. The NADI parameter is a scaling coefficient for this proportionality relation. For example, for NADI=1 and $p(MF) = 0.6$ the $GABA_A$R-mediated conductance will be increased by 60%, whereas with the same $p(MF)$ but NADI=0.3 the increase will be of 20%. *B:* average spikes per GC versus average spikes per MF across all patterns and all values of the probability of MFs being active $p(MF)$, parametrised by the number of synaptic connections for the different levels of NADI. *C:* Average MI, normalised by input entropy, versus average output sparseness (with averages taken across all values of $p(MF)$) for different values of NADI parametrised by the number of synaptic connections. The value of NADI=0.3, which is my rough estimate of how the time-averaged $GABA_A$R conductance contributed by Golgi cells varies with network activity (subsubsection 2.2.6.1), increased the performance for networks with few synaptic connections.

Parameter	Mean	Std. dev.
Membrane capacitance (pF)	3.6	0.75
Leak conductance (nS)	0.77	0.29
Resting potential (mV)	-66	13
Spike threshold (mV)	-34	6.6
Reset potential (mV)	-62	8.2

Table 4: Granule cell population model: mean and standard deviation for marginal distribution of single cell parameters, estimated from the experimental data in Table 3.

Figure 40: Granule cell population model: Pearson correlation coefficients estimated for all pairs of single cell parameters from the experimental data in Table 3. Stars mark significance level (0.1, 0.05 and 0.01 for one, two and three stars respectively).

gical properties of granule cells — seen as a population of integrate-and-fire (IaF) neurons — with a five-dimensional Gaussian. Equivalently, the distribution of the experimental GrC population in the IaF parameter space is well described by simply giving its mean and covariance matrix, or the mean and standard deviation for the five marginal distributions together with the 5*(5-1)/2=10 correlation coefficients between all possible pairs of single cell parameters. It is this last form that I use to explicitly specify the model, by reporting the values estimated from the experimental data in table Table 4 and Figure 40. Examples of low-dimensional projections of the experimental data, together with some synthetic data generated from the Gaussian model, can be seen in Figure 41.

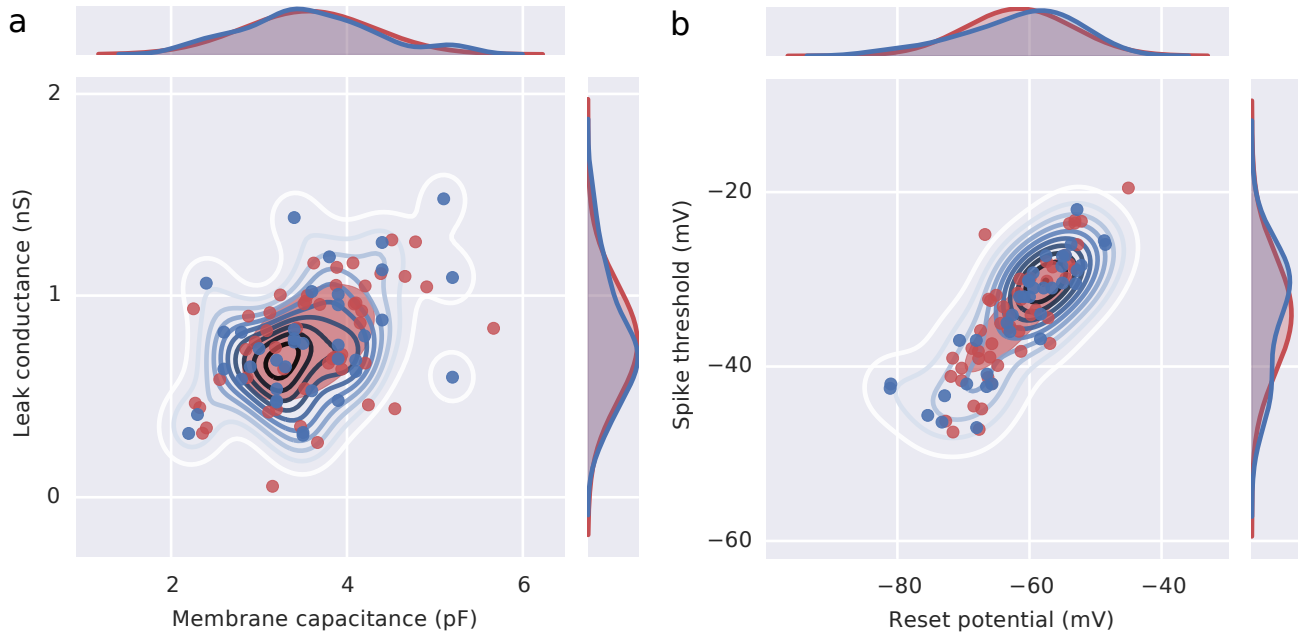


Figure 41: Two of the ten possible bi-dimensional projections of the granule cell population data and model. Blue dots: experimental data points. Blue contour plot: kernel density estimate (KDE) for the experimental data, using Gaussian kernels with bandwidths set according to Scott's normal reference rule [Scott, 2009]. Red dots: a sample of 50 synthetic data points generated from the Gaussian population model. Red ellipses are correlation ellipses (mean $\pm$ one standard deviation along the principal axes) for the model projected on the relevant plane. On the margins: marginal distributions for the experimental data KDE (blue) and the model (red). *a*: leak conductance (G_m) versus membrane capacitance (C_m). *b*: spike threshold (V_t) versus reset potential (V_r).

Eigenvector #	1	2	3	4	5
Membrane capacitance	-0.18	+0.79	-0.28	+0.50	-0.08
Leak conductance	-0.40	+0.47	+0.52	-0.56	+0.17
Resting potential	-0.44	-0.29	+0.59	+0.59	-0.12
Spike threshold	-0.54	-0.22	-0.41	+0.04	+0.70
Reset potential	-0.56	-0.16	-0.36	-0.27	-0.68
Explained variance ratio	60.0%	24.5%	15.9%	6.4%	2.2%

Table 5: Granule cell population model: coordinates and explained variance ratio for the full set of eigenvectors obtained by performing PCA on the correlation matrix for the experimental data in Table 3. Note that, because PCA is done on the correlation matrix, the coordinates in this table are Z-scores, hence dimensionless.

To further investigate the intrinsic complexity of the parameter distribution for the GrC population, I applied Principal Component Analysis (PCA) to the correlation matrix for the experimental data in Table 3 (I used the correlation matrix rather than the covariance matrix as the different parameters have different scales - see Jolliffe [2002] and discussion therein). The results, reported in Table 5, show that about 75% of the variance of the dataset can be explained by a projection on the plane spanned by the eigenvectors associated with the two largest eigenvalues of the covariance matrix.

In this section, I have shown that the granule cell population can be described by a simple distribution over the set of single-cell parameters for integrate-and-fire models and that, furthermore, a projection of the population onto its first two principal components is sufficient to capture most of it variability. Finally, I have used this information to build a generative model of GrC populations that incorporates biologically accurate estimates of heterogeneity.

5.4 DISCUSSION

In this chapter, I have explored the relationship between the structure of feed-forward network models and their ability to transmit information and transform it into a sparse representation, which are both essential for pattern separation [Tyrrell and Willshaw, 1992]. I built a spiking local model of the GCL network with synaptic inputs, neuronal properties, and tonic inhibition constrained to experimentally measured values (section 2.2, section 2.3), and I devised a simple method for analysing it from an information theoretical standpoint (section 2.4). Confirming theoretical predictions [Billings et al., 2014], I showed that that the extent of the synaptic connectivity sets the

trade-off between information transmission and sparse encoding (Figure 35). Furthermore, I have used the flexibility afforded by my numerical approach to test for the effect of spatial structure in the input signals (Figure 35) and in the connectivity of the network (Figure 38), and to introduce a simple model of feed-forward inhibition that can be directly related to experimental findings (Ward [2012], Figure 39). My results suggest that spatial correlations have only a minor effect at the spatial scale I analysed, and that a modest amount of feed-forward-like inhibition — compatible with experimental estimates in the rate-coding regime — enhances the information transmission-sparsification trade-off. Additionally, I have shown that these findings are robust with respect to the choice of the less well constrained parameters like the duration of the time window allowed for synaptic integration and the firing rate of mossy fibres (Figure 36, Figure 37).

Even though the models used in my analysis present a detailed phenomenological description of synaptic integration in granule cells, important aspects of cerebellar cytology are not contemplated, like heterogeneity of synaptic [Sargent et al., 2005] and cell [Schwartz et al., 2012; Ward, 2012] properties. To begin to address this, I have analysed single-cell GrC properties in a rich experimental dataset (section 5.3). I have derived a simple statistical model capturing most of their variability, showing that a GrC population can be characterised with just a few distributional parameters. Together with existing quantitative descriptions of the variability of the MF→GrC synapse [Sargent et al., 2005], this model will allow to include the effect of cell- and synapse-level heterogeneity in future extensions of this study.

Another critical aspect that was only partially explored time to due limitations is the effect of departures of the statistics of the input from the assumption of "uncorrelated binary rate coding" [Billings et al., 2014], like the spatio-temporal correlations that could be introduced by unipolar brush cells [Mugnaini et al., 2011; Kennedy et al., 2014] and the temporal signal structures involved in communicating rapid discrete sensory events [Jörntell and Ekerot, 2006; Rancz et al., 2007]. Temporally structured input signals could be analysed using a similar approach and computational tools like those I built and published (described in section 2.6). Detailed biophysical models of the UBC are available [Subramaniyam et al., 2014], and could be readily included in the network model. Finally, the impact of the stochastic nature of neurotransmitter release on information transmission in granule cells [Arleo et al., 2010] could also be an important element to take into account in an analysis extended to the temporal coding regime.

The main practical obstacle to the extension of this approach is the scaling of its computational cost with the introduction of new elements in the simulations or analyses. This type of analysis is not expected to scale well with the dimensionality of input and output

spaces. Indeed, controlling for the systematic error (or *bias*) that can arise in estimations of mutual information is a hard problem [Treves and Panzeri, 1995; Nemenman et al., 2002; Paninski, 2003; Panzeri et al., 2007; Hausser and Strimmer, 2009; Archer et al., 2013] that gets harder as the number of possible inputs and outputs of our system increases with respect to our sampling capacity. On the other hand, considering a temporal code for inputs and outputs would require the use of more sophisticated dimensionality reduction tools for the output (like the spike train distance-based method discussed in section 2.6), which could come at a significant computational cost. Incorporating stochastic synaptic transmission would, in general, make the system noisier and thus harder to sample, feeding back into the bias problem mentioned above.

In conclusion, while the work presented in this chapter can be extended to investigate the relevance of other features of the synaptic organisation of the GCL and MF input in the *expansion recoding* picture, scaling the approach to larger input/output dimensionalities would be intrinsically hard, even if done with more powerful computational tools.

6 GENERAL DISCUSSION

In this study, I have investigated the relationship between the network structure of neuronal ensembles in the input stage of the cerebellum and their dynamical properties. Using data from imaging and electrophysiology experiments, as well as published anatomical information, and combining computational, statistical and information theoretical approaches, I have determined some of the functional implications of the salient anatomical and physiological aspects of the granule cell layer — the glomerular organisation of the MF→GrC synapse and the electrical connectivity between Golgi cells.

6.1 SPATIAL CONNECTIVITY RULES AND NETWORK STRUCTURE IN THE GCL

A distinguishing feature of the cerebellar cortex is the homogeneous and symmetric arrangement of its cytoarchitecture [Braitenberg and Atwood, 1958]. On a local scale within the granule cell layer, this is enabled by a combination of local spatial distribution rules and morphological properties for the different cell types.

6.1.1 *Local structure and MF input sharing at the glomerulus*

One of the local aspects preeminently reflecting this property within the granule cell layer is how the glomerular organisation of the MF to GrC synapses conspires with the stereotypical morphology of granule cells [Wittenberg and Wang, 2008] to constrain network structure. In section 3.1, I quantified the extent to which the typical length of granule cell dendrites affects the combinatorial properties of MF sampling by GrCs. I showed that the amount of input sharing induced among neighbouring GrCs by a realistic model of glomerular organisation is significant — if compared with the case of spatially random connections — even when confined to a spherical volume of 80 µm diameter. As expected, input sharing was limited to cell pairs within ~40 µm of each other. This has significant implications, as it determines the maximum range of correlations expected to be observed in GCL activity following spatially independent MF input, and as such characterises a natural choice for the minimum size of a local network model. Furthermore, this spatial scale matches the reported extent of activation spots in the GCL following focal MF stimulation in slice [Mapelli et al., 2010; Gandolfi et al., 2014].

6.1.2 *Spatial distribution of Golgi cells*

Eccles et al. [1967] proposed that Golgi cells are arranged in an regular hexagonal pattern, "tiling" the cortex with minimal overlap of their roughly isotropic dendritic arborisations. This picture is now known not to be accurate, as Golgi cells form extensive dendro-dendritic contacts in both the granular layer and molecular layer [Vervaeke et al., 2010, 2012]. Nevertheless, to my knowledge, no studies have ever attempted to quantify the reciprocal positions assumed by Golgi cell bodies within the granule cell layer (on the other hand, Altman and Winfree [1977] provided a careful characterisation of reciprocal positions in Purkinje cells).

In section 3.1 I analysed a large data set from immunohistochemistry experiments, and I have found evidence of spatial regularity below the 30 µm scale in the transverse plane. Golgi cell bodies are typically 5–10 µm (small Golgi) or ~15 µm (large Golgi) in diameter [Palay and Chan-Palay, 1974; Simat et al., 2007], so it is possible that the regularity extends — albeit weakly — beyond the direct effect of cell body exclusion. Nevertheless, before reaching a definitive conclusion it is necessary to better control for systematic errors that can arise in this type of estimate [Baddeley et al., 2000]. Additional research is currently being carried out following the method outlined in subsubsection 2.1.1.3 to address this issue.

Another point to consider is that the second order structure of the spatial distribution of the cells could be anisotropic, presenting a stronger regularity in the directions parallel to the surface of the cortex. With the existing data, this could perhaps be tested by considering only those regions at the folds of the lobules, where the cortex is in fact parallel to the slicing direction (see Figure 4). On the other hand, the curved geometry of the cortex at the lobule folds could introduce systematic distortions in such an estimate, and for a reliable study a new experimental preparation cut parallel to the cortex — or indeed a full 3D representation of the tissue — would be preferable.

As a topic of further research, it could also be interesting to isolate those regions of GCL where all GoC subtypes are sufficiently abundant [Simat et al., 2007; Valera et al., 2012], and test for the existence of cross-subtype position correlations [Diggle, 1986; Møller and Waagepetersen, 2003; Diggle et al., 2006]. This could perhaps shed some light on the possible developmental mechanisms responsible for the establishment of the regularity, clarifying whether they act uniformly across subpopulations or if they can be linked to those responsible for the inhomogeneous anatomical placement of different subtypes [Valera et al., 2012].

Overall, regular placement of Golgi cells in space could have significant functional implications, as connection probability and coupling

strength within the recurrent network formed by GoCs (see below) is known to be strongly distance-dependent [Dugué et al., 2009; Vervaeke et al., 2010].

6.1.3 *Structural aspects of the Golgi cell network*

Cerebellar Golgi cells have been shown to be interconnected through electrical synapses [Dugué et al., 2009]. Electrophysiological experiments and morphological reconstructions have shown that the electrical coupling between pairs of GoCs is distance dependent [Dugué et al., 2009; Vervaeke et al., 2010], and the connexin36-containing gap junctions (GJs) mediating the synapses are distributed nonuniformly on the dendritic trees of the cells [Vervaeke et al., 2012]. Our understanding of the structure of the Golgi cell network rests exclusively on measurements on single cells and more or less randomly selected pairs of cells; no systematic investigation has been attempted into structural aspects such as the degree distribution of the network or the incidence of higher-order connectivity motifs. With the available information, multiple models of gap junction connectivity have been developed for use in numerical simulations of network activity [Dugué et al., 2009; Vervaeke et al., 2010, 2012] and analytical studies [Ostojic et al., 2009]. Nevertheless, these models are either overly abstract (that by [Ostojic et al., 2009] is really a generic model of electrically coupled spiking neurons; furthermore, it does not include any spatial structure) or they are specified as fairly complicated algorithms, making them suitable to generate instantiations of the network but providing very little insight on their own.

In section 2.1 and section 3.2 I have formulated a model of Golgi network connectivity that is constrained by the available experimental evidence while still allowing for some simple structural characteristics to be computed analytically. Inclusion of the spatial regularity effects discussed in subsection 6.1.2 did not seem to contribute significantly to the structural aspects I quantified, although it could still bear some significance from a dynamical point of view. The model predicts that the degree distribution of the Golgi network is Poisson, and that on average each cell is directly connected with about 15 other cells. Combined with available data on the numerosity of gap junctions of Golgi cell bodies [Vervaeke et al., 2012], this immediately leads to a further prediction on the average number of gap junctions mediating an individual connection between GoCs. The prediction (~2.3 GJs per connection) has found supporting evidence from electron microscopy studies [Lőrincz et al., 2015].

Even though my prediction for the average connectivity has received indirect experimental confirmation, it is possible and perhaps likely that, because of its simplicity, the model presented here fails to

capture some relevant structural aspect of the network. Indeed, different genetic subtypes of Golgi cells are known to exist [Simat et al., 2007], and their spatial distribution has been shown to be strongly dependent on anatomical location [Valera et al., 2012]; this raises the possibility of subtype-specific connectivity within the Golgi cell population. Moreover, the distribution of gap junctions on Golgi cell dendrites is spatially nonuniform [Vervaeke et al., 2012], and although GoC dendrites display a roughly isotropic organisation [Palay and Chan-Palay, 1974], the asymmetry between apical and basolateral arborisations could imply that connection probability does not only depend on intersomatic distance, as I have assumed. Finally, interneurons of the molecular layer are known to assume complex connectivity configurations [Rieubland et al., 2014] that can't be explained by a simple network model with distance-dependent but uncorrelated edges, and Golgi cells could share this feature.

Future experiments could test directly the predicted connectivity statistics by dye coupling — a technique that has already seen successful application in another population of cerebellar interneurons [Mann-Metzer and Yarom, 1999]. Another structural aspect worthy of study, and with evident functional implications as I discuss below, is the distribution of the number of gap junctions per connected cell pair; indeed, this is the subject of ongoing investigations [Lőrincz et al., 2015].

6.2 SYNCHRONY AND NETWORK ORGANISATION IN GOLGI CELLS

The granule cell layer exhibits low frequency local field potential (LFP) oscillations [Pellerin and Lamarre, 1997; Hartmann and Bower, 1998]. These oscillations, often synchronised with corresponding activity patterns in the sensorimotor cortex [O'Connor et al., 2002; Courtemanche and Lamarre, 2005], are associated with a behavioural state of "quiet wakefulness" of the animal, and rapidly disappear in coincidence with the initiation of a movement; furthermore, analogous transitions have been linked with attention regulation [Harris and Thiele, 2011]. Golgi cells fire synchronously in phase with the LFP oscillations; indeed, they are electrically interconnected in a network mediated by connexin36-containing gap junctions [Dugué et al., 2009], and electrical coupling has been proposed as a dynamical substrate for the maintenance and active control of synchronous activity. Gap junctions enable the establishment of coherent activity in presence of asynchronous synaptic input [Dugué et al., 2009] while at the same time mediating rapid desynchronisation following a temporally precise distributed stimulation [Vervaeke et al., 2010]. In particular, this last dynamical property has been proposed to play a role in the disappearance of the LFP oscillations [Vervaeke et al., 2010].

6.2.1 *Effect of the spatial structure of the network*

In section 4.2, I have shown how the spatial connectivity range between electrically coupled Golgi cells affect the ability of the GoC network to desynchronise. Critically, my results indicate that, if realistic physiological and synaptic properties for GoCs are taken into account, spatial localisation of the connectivity is necessary to enable the network to leave the synchronous state following a short, distributed synaptic input. This highlights the importance of studying the spatial aspects of the connectivity if we are to understand the mechanisms underlying synchronous activity, and points towards new lines of future enquiry.

6.2.1.1 *Implications of the presence of chemical inhibition between Golgi cells*

Recently, experimental evidence has been provided for the existence of chemical inhibition between Golgi cells [Hull and Regehr, 2012]. The functional implications of this type of connectivity are not clear yet, but it is reasonable to expect it to play a role in the establishment and regulation of synchrony. Intuitively, since in Golgi cells low-pass filtering of action potentials at gap junctions produces predominantly inhibitory spikelets [Dugué et al., 2009], recurrent chemical inhibition should not by itself give rise to qualitatively new effects. However, electrical synapses are exclusively tied to dendritic overlap as there are no axonal gap junctions [Vervaeke et al., 2012], while Golgi cell axons are organised parasagittally on a larger spatial scale than the dendrites [Barmack and Yakhnitsa, 2008]. This means that recurrent GABAergic synapses between Golgi cells could significantly alter the structure of the network both by increasing the spatial extent of its connectivity and by breaking the rotational symmetry assumed — at least in the cortical plane — for the dendritic arborisations of Golgi cells.

6.2.1.2 *Further directions for an analytical study*

Numerical approaches to the systematic study of network dynamics are very flexible, as they allow for models to be tightly constrained by experimental data while at the same time enabling rapid exploration of their behaviour through free manipulation of their properties; on the other hand, analytical studies of simplified models often bear clearer insight. My findings confirm and reinforce the observation made by Ostojic et al. [2009] that further analytical studies of synchrony in Golgi cells will have to accommodate for the complex structure induced by the spatial embedding of the network.

Synchronisation in diffusively coupled complex networks has been treated with the *master stability function* approach and other tech-

niques that reduce the synchronisability of a network to the spectral properties of its graph laplacian (Pecora and Carroll, 1998; Barahona and Pecora, 2002; Wu, 2003; Pereira et al., 2013; see also Arenas et al. [2008]). Presumably, similar methods could be applied to the problem of synchronisation in Golgi cells. For this to be feasible, an appropriately simplified model of a single Golgi cell and of electrical coupling, including dendritic attenuation of synaptic currents, should be formulated. As a starting point, note that for a spatially embedded network, and at least in the case of a deterministic spatially-dependent connectivity model, the graph laplacian is a *euclidean matrix*, i.e. it belongs to a class of random matrix ensembles whose spectral properties have been first studied by Mézard et al. [1999] for the case of independent "cell" positions. How to integrate stochastic connectivity and, possibly, spatially regular cell positioning into this picture remains to be understood.

6.2.2 *Effect of intrinsic physiological heterogeneities*

Heterogeneity in synaptic conductances was found to play a role in enabling desynchronisation (section 4.2). Yet, inducing a significant increase in the instability of the synchronous network state required assuming a distribution of synaptic strengths whose shape was very different than that predicted by the analysis of experimental data [Vervaeke et al., 2010]. Therefore, stability of synchronous activity appeared to vary rather smoothly with the heterogeneity of synaptic conductance.

For the sake of simplicity, in my analysis I did not include variability in physiological properties within the cell population, an element which has been recognised as important, though not fundamental, in facilitating asynchronous network states [Ostojic et al., 2009; Vervaeke et al., 2010]. Furthermore, an independent source of heterogeneity in effective synaptic strength — the random placement of gap junctions on the dendritic arborisations of Golgi cells — was present and constant throughout the analysis, and did not get explored due to time limitations. This would have been particularly interesting in the light of the experimental results by Vervaeke et al. [2012] showing that the spatial distribution of gap junctions on Golgi cell dendrites is indeed nonuniform. Evidently, these are both possible avenues for an immediate extension of the work presented here.

6.3 EXPANSION RECODING IN THE GRANULE CELL LAYER

(Some of the ideas discussed in this section have previously appeared in Billings et al. [2014]). In chapter 5, I investigated how network structure affects information transmission in the granule cell layer, in the

context of the functional hypothesis of *expansion recoding*. In order to do this, I built and simulated an experimentally constrained local model of the GCL, and applied a combination of techniques from machine learning and information theory, appropriately deployed on a suitable HPC infrastructure.

Confirming theoretical predictions [Billings et al., 2014], I found that the properties of the cerebellar input layer are particularly well suited for performing lossless sparse encoding. My results therefore provide a computational explanation of why the most numerous neuron in the brain of vertebrates receives an average of four excitatory synaptic inputs.

6.3.1 *The relationship between feedforward network structure and function*

To understand how synaptic connectivity affects network function, it is first necessary to understand why information transmission and sparse encoding are competing functions. Although a local network of 500 GrCs could potentially encode an astronomical number of MF input patterns (i.e., 2^{500} for a binary rate code like the one used in this work), when GrC activity is reduced to low levels the capacity of the network to encode patterns shrinks considerably. My results show that the balance between information transmission and the sparseness of the encoding is set by the extent of synaptic connectivity. Having few excitatory synaptic connections per neuron and a high level of inhibition provides a highly effective trade-off between information transmission and sparse encoding. The fact that the synaptic connectivity of the best performing networks match that found in the cerebellar input layer suggests that the GCL structure is optimized for transforming MF input patterns into a higher dimensional sparser code, without information loss. This can be advantageous for improving the performance of a downstream associative memory [Tyrrell and Willshaw, 1992] and to preserve energy [Attwell and Laughlin, 2001]. My results extend classical work on the relationship between cerebellar structure and function [Marr, 1969; Albus, 1971; Kanerva, 1988], by showing that the synaptic connectivity between MFs and GrCs is a major determinant of information transmission and sparse encoding in this brain region.

6.3.2 *Relationship between Mossy Fiber Activity and Granule Cell Layer Properties*

Since the cerebellum receives dynamic patterns of sensory-motor inputs via the MF system, the ability to perform lossless sparse encoding over a wide range of MF excitatory drive is likely to be crucial. MFs are known to exhibit diverse activity patterns in time and space:

they can fire continuously in the 10–100 Hz range [Van Kan et al., 1993; Arenz et al., 2008] or in rapid bursts reaching 700 Hz [Jörntell and Ekerot, 2006; Rancz et al., 2007], while a set of MFs impinging on an individual GrC can carry multimodal [Huang et al., 2013] or highly correlated signals [Jörntell and Ekerot, 2006]. My results show that the synaptic connectivity found within the GCL can losslessly transform a wide range of MF excitatory drive into a sparse GrC population code, even when spatial correlations in MF activity are present. Moreover, the broad bandwidth of MF-GrC signaling [Saviane and Silver, 2006] also enables certain sensory stimuli, such as whisker deflection, which generate high-frequency MF bursts, to be relayed through the GCL. Thus, my results suggest that the GCL acts as a general purpose sparse encoder of rate-coded MF inputs that has the flexibility to respond rapidly to urgent stimuli.

6.3.3 *Determinants of network encoding capacity*

My results show that noisy biologically detailed spiking networks can comfortably encode 1024 rate-coded MF input patterns, assuming GrC spikes are integrated over 30 ms (Figure 35). On the other hand, theoretical results show that the encoding capacity of noise-free binary networks is sufficiently large to encode all patterns that an animal could possibly encounter during its lifetime [Billings et al., 2014]. While the number of patterns that a real local GCL network encodes falls between these two extremes, it may vary widely across cerebellar regions because encoding capacity depends on MF firing rates (Figure 37), the properties of inhibition (Figure 39), the time window over which GrC firing is integrated (Figure 36) and — albeit in a minor way on the scales tested — spatial correlations in MF activity (Figure 35D–H). These considerations suggest that the encoding capacity of a local GCL network will depend strongly on the properties of the MF inputs it receives.

Another way to increase the encoding capacity is to increase the number of local networks engaged. *In vivo* recordings from GrCs in mouse vestibular cerebellum indicate that ~400 MF→GrC synapses are required to encode head velocity at the precision observed in man [Arenz et al., 2008], suggesting that multiple local GCL networks are involved. Indeed, MF axons, which form ~20 en passant synaptic rosettes [Eccles et al., 1967; Sultan, 2001], enable GrCs in neighboring local networks to sample the same MF signals. The idea that many GrCs are required for sensory representations is supported by the finding that markedly reducing the number of functional GrCs induces deficits in consolidation of motor learning (but concomitant changes in long-term plasticity could also contribute to these effects; Galliano et al., 2013). These observations are consistent with the no-

tion that multiple local GCL networks are involved in certain sensory-motor tasks and that the large MF to GrC divergence found in the cerebellum is required for efficient encoding.

6.3.4 *The Properties of Inhibition and Encoding Capacity*

My results show that the physiological level of tonic $GABA_AR$-mediated inhibition [Brickley et al., 1996] provides a robust solution for performing lossless sparse encoding. Recent *in vivo* recordings show that tonic inhibition dominates other forms of inhibition in GrCs, accounting for 98% of the inhibitory charge [Duguid et al., 2012]. This sets a relatively high threshold so that simultaneous activity from three or more rate-coded MF inputs are typically required to reach GrC firing threshold [Jörntell and Ekerot, 2006; Schwartz et al., 2012]. The importance of a high GrC spike threshold to cerebellar function is reinforced by the finding that when tonic inhibition was eliminated in $GABA_A\alpha_6$ knockout mice, two-pore K^+ channels were upregulated, thereby maintaining threshold at a high level [Brickley et al., 2001]. However, knocking out transporters has been more effective in modulating GrC threshold. Deletion of the GABA transporter GAT1 increased tonic inhibition in GrCs by 4-fold and was associated with tremor and ataxia [Chiu et al., 2005]. My results suggest that information loss could have co-occurred with these behavioural effects. Lowering GrC spike threshold by selectively deleting the KCC2 chloride transporter in GrCs, which my results would suggest reduces the sparseness of encoding, impairs learning consolidation [Seja et al., 2012]. Thus, my model of GCL function provide insights into how alterations in the level of tonic inhibition could impair cerebellar function and why GrC spike threshold is tightly regulated by homeostatic mechanisms.

When an appropiate amount of network-activity-dependent inhibition was added to tonic inhibition, it further sparsified GrC encoding without loss of information. Furthermore, the temporal structure of GoC activity (such as the synchronicity discussed elsewhere in this study) has been shown to have little effect on GrC input-output gain in the rate coding regime [Ward, 2012]. My results therefore support previous proposals that Golgi cells aid sparse coding by controlling the gain to the GCL [Marr, 1969; Albus, 1971; Schweighofer et al., 2001]. Although weaker, the phasic and spillover components of Golgi cell-mediated inhibition [Rossi et al., 2003] may also contribute to temporal pattering [Ward, 2012], which could perform temporal sparsening of GrC spikes, time slicing [D'Angelo and De Zeeuw, 2009] and introduce delays that are important for learning temporal operations such as eyeblink conditioning [Medina and Mauk, 2000] and signal cancellation [Kennedy et al., 2014]. However,

the longer temporal delays during signal cancellation are mediated by unipolar brush cells [Kennedy et al., 2014], which form short-range intrinsic MFs that could increase spatial correlations in the vestibular cerebellum where they are more numerous. Indeed, regional variations in the origin of MF inputs, the presence of UBCs and synaptic plasticity within the MF-GrC-Golgi cell circuit could tune the spatiotemporal transformation that specific GCL "modules" perform. My results show that few synaptic connections per GrCs provide a robust structural framework that enables a wide range of MF activity patterns to be transmitted and sparsified efficiently.

6.3.5 *Synaptic Connectivity of the Cerebellar Input Layer Is Evolutionarily Conserved*

The cerebellum is an ancient brain structure that arose in the early vertebrates. In terms of numbers, cerebellar GrCs dominate the vertebrate CNS, making up more than half of all the neurons in the human brain [Williams and Herrup, 1988]. Remarkably, the morphology of cerebellar GrCs is conserved across a wide range of species including fish, amphibians, reptiles, and mammals [Llinas, 1969; Wittenberg and Wang, 2008], demonstrating that it has been evolutionarily conserved. In mammals, the observed range is two to seven dendrites (and thus MF inputs) per GrC [Palkovits et al., 1972] with four per cell being the most common configuration. The strikingly similarity between the synaptic connectivity in the cerebellar GCL and the feedforward networks that provide the best trade-off between information transmission and sparsification provides a functional explanation for why the characteristic dendritic morphology of cerebellar GrCs has been conserved for hundreds of millions of years.

6.3.6 *Comparison of the structure of the GCL to other networks*

The cerebellar GCL is not the only example of a network that performs sparsification and has few synaptic connections per neuron. GrCs in the dorsal cochlear nucleus and deep GrC in the electrosensory lobe (ELL) of the electric fish have one to four synaptic inputs, averaging to three in the ELL [Kennedy et al., 2014; Mugnaini et al., 1980; Zhang et al., 2007]. Kenyon cells in the mushroom body of the fly receive an average of seven synaptic inputs from olfactory projection neurons [Caron et al., 2013]. Indeed, expansion recoding in the insect mushroom body has other similarities to the cerebellar input layer [Laurent, 2002], including random connectivity [Caron et al., 2013] and inhibitory interneurons that facilitate sparsification [Papadopoulou et al., 2011], enhancing pattern separation and enabling the discrimination of similar odours [Lin et al., 2014]. These

examples suggest that other brain structures may have converged on a similar feed-forward network structure for performing lossless sparse encoding.

6.4 SUMMARY

The cerebellar cortex is characterised by its remarkable anatomical regularity, and by the distinctive set of symmetries possessed by its architecture. Its structure is, to a great extent, shared across all vertebrates; the relative sizing of its subregions are linked to functional specialisations in different animals. It has been proposed that the cerebellum acts as modular, parallel *neuronal machine*, implementing a set of widely applicable operations on a variety of signals of different nature.

In this work, I have studied the input stage of the cerebellar cortex, analysing the architectural substrates of some of its putative modes of operation. A prediction — which later received experimental support — on the nature of the electrical connectivity between individual Golgi cells was made, on the basis of an analytically tractable model of the gap junction-mediated Golgi cell network. The spatial structure of the GoC network was found to play a key role in the control of synchronous activity. The peculiar and evolutionarily conserved dendritic morphology of granule cells, together with the properties of GABAR-mediated inhibition from Golgi cells, was found to enable the granule cell layer to function effectively as an *expansion recoder*.

BIBLIOGRAPHY

Agmon-Snir, H., Carr, C. E., and Rinzel, J. The role of dendrites in auditory coincidence detection. *Nature*, 393(6682):268–272, May 1998. (Cited on page 19.)

Albus, J. S. A theory of cerebellar function. *Mathematical Biosciences*, 10(1–2):25 – 61, 1971. (Cited on pages 27, 29, 101, 120, and 122.)

Altman, J. and Winfree, A. T. Postnatal development of the cerebellar cortex in the rat: V. spatial organization of purkinje cell perikarya. *The Journal of Comparative Neurology*, 171(1):1–16, 1977. (Cited on page 115.)

Amari, S.-i. Natural gradient works efficiently in learning. *Neural Computation*, 10(2):251–276, February 1998. (Cited on page 17.)

Apps, R. and Garwicz, M. Anatomical and physiological foundations of cerebellar information processing. *Nat Rev Neurosci*, 6(4):297–311, April 2005. (Cited on page 27.)

Apps, R. and Hawkes, R. Cerebellar cortical organization: a one-map hypothesis. *Nat Rev Neurosci*, 10(9):670–681, September 2009. (Cited on pages 21 and 26.)

Archer, E., Park, I. M., and Pillow, J. W. Bayesian and quasi-bayesian estimators for mutual information from discrete data. *Entropy*, 15 (5):1738–1755, 2013. (Cited on page 113.)

Arenas, A., Díaz-Guilera, A., Kurths, J., Moreno, Y., and Zhou, C. Synchronization in complex networks. *Physics Reports*, 469(3):93 – 153, 2008. (Cited on pages 18, 87, 100, and 119.)

Arenz, A., Silver, R. A., Schaefer, A. T., and Margrie, T. W. The contribution of single synapses to sensory representation in vivo. *Science*, 321(5891):977–980, 2008. (Cited on pages 25, 61, 101, and 121.)

Arleo, A., Nieus, T., Bezzi, M., D'Errico, A., D'Angelo, E., and Coenen, O. J.-M. D. How synaptic release probability shapes neuronal transmission: information-theoretic analysis in a cerebellar granule cell. *Neural Comput*, 22(8):2031–2058, Aug 2010. (Cited on page 112.)

Armano, S., Rossi, P., Taglietti, V., and D'Angelo, E. Long-term potentiation of intrinsic excitability at the mossy fiber–granule cell synapse of rat cerebellum. *The Journal of Neuroscience*, 20(14):5208–5216, 2000. (Cited on page 27.)

Arthur, D. and Vassilvitskii, S. K-means++: The advantages of careful seeding. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '07, pages 1027–1035, Philadelphia, PA, USA, 2007. Society for Industrial and Applied Mathematics. ISBN 978-0-898716-24-5. (Cited on page 63.)

Ascoli, G. A., Donohue, D. E., and Halavi, M. Neuromorpho.org: A central resource for neuronal morphologies. *The Journal of Neuroscience*, 27(35):9247–9251, 2007. (Cited on page 65.)

Attwell, D. and Laughlin, S. B. An energy budget for signaling in the grey matter of the brain. *J Cereb Blood Flow Metab*, 21(10):1133–1145, October 2001. (Cited on page 120.)

Baddeley, A. J., Møller, J., and Waagepetersen, R. Non- and semi-parametric estimation of interaction in inhomogeneous point patterns. *Statistica Neerlandica*, 54(3):329–350, 2000. (Cited on pages 32, 34, 35, and 115.)

Baddeley, A. and Turner, R. Spatstat: An r package for analyzing spatial point patterns. *Journal of Statistical Software*, 12:1–42, 2005. (Cited on page 34.)

Barahona, M. and Pecora, L. M. Synchronization in small-world systems. *Physical Review Letters*, 89:054101, Jul 2002. (Cited on pages 18, 100, and 119.)

Barber, D. *Bayesian Reasoning and Machine Learning*. Cambridge University Press, 2012. (Cited on page 58.)

Barmack, N. H. and Yakhnitsa, V. Functions of interneurons in mouse cerebellum. *The Journal of Neuroscience*, 28(5):1140–1152, 2008. (Cited on page 118.)

Barrat, A., Barthélemy, and Vespignani, A. *Dynamical Processes on Complex Networks*. Cambridge University Press, 2008. (Cited on page 18.)

Barthélemy, M. Spatial networks. *Physics Reports*, 499(1–3):1 – 101, 2011. (Cited on page 87.)

Bastian, J. Pyramidal-cell plasticity in weakly electric fish: a mechanism for attenuating responses to reafferent electrosensory inputs. *J Comp Physiol A*, 176(1):63–73, Jan 1995. (Cited on page 27.)

Bastian, M., Heymann, S., and Jacomy, M. Gephi: An open source software for exploring and manipulating networks. In *International AAAI Conference on Weblogs and Social Media*, 2009. (Cited on page 83.)

Beierlein, M., Gibson, J. R., and Connors, B. W. A network of electrically coupled interneurons drives synchronized inhibition in neocortex. *Nat Neurosci*, 3(9):904–910, Sep 2000. (Cited on pages 30 and 91.)

Bell, C. C. An efference copy which is modified by reafferent input. *Science*, 214(4519):pp. 450–453, 1981. (Cited on page 27.)

Bengio, Y. Learning deep architectures for ai. *Found. Trends Mach. Learn.*, 2(1):1–127, Jan 2009. (Cited on page 17.)

Billings, G., Piasini, E., Lőrincz, A., Nusser, Z., and Silver, R. A. Network structure within the cerebellar input layer enables lossless sparse encoding. *Neuron*, 83(4):960–974, 2014. (Cited on pages 45, 46, 49, 58, 59, 60, 62, 64, 101, 102, 103, 104, 105, 107, 108, 111, 112, 119, 120, and 121.)

Bishop, C. M. *Pattern recognition and machine learning*. Springer New York, 2006. (Cited on pages 63 and 68.)

Blei, D. M. and Jordan, M. I. Variational inference for dirichlet process mixtures. *Bayesian Anal.*, 1(1):121–143, 03 2006. (Cited on page 58.)

Bostan, A. C., Dum, R. P., and Strick, P. L. Cerebellar networks with the cerebral cortex and basal ganglia. *Trends in Cognitive Sciences*, 17(5):241–254, May 2013. (Cited on page 27.)

Braitenberg, V. Functional interpretation of cerebellar histology. *Nature*, 190:539–540, May 1961. (Cited on page 27.)

Braitenberg, V., Heck, D., and Sultan, F. The detection and generation of sequences as a key to cerebellar function: experiments and theory. *Behav Brain Sci*, 20(2):229–45; discussion 245–77, Jun 1997. (Cited on page 28.)

Braitenberg, V. and Atwood, R. P. Morphological observations on the cerebellar cortex. *The Journal of Comparative Neurology*, 109(1):1–33, 1958. (Cited on pages 22, 23, and 114.)

Brandes, U., Eiglsperger, M., Herman, I., Himsolt, M., and Marshall, M. GraphML progress report structural layer proposal. In Mutzel, P., Jünger, M., and Leipert, S., editors, *Graph Drawing*, volume 2265 of *LNCS*, pages 501–512. Springer Berlin Heidelberg, 2002. ISBN 978-3-540-43309-5. (Cited on page 67.)

Brickley, S. G., Cull-Candy, S. G., and Farrant, M. Development of a tonic form of synaptic inhibition in rat cerebellar granule cells resulting from persistent activation of gabaa receptors. *The Journal of Physiology*, 497(3):753–759, 1996. (Cited on page 122.)

Brickley, S. G., Revilla, V., Cull-Candy, S. G., Wisden, W., and Farrant, M. Adaptive regulation of neuronal excitability by a voltage- independent potassium conductance. *Nature*, 409(6816):88–92, January 2001. (Cited on page 122.)

Brochu, G., Maler, L., and Hawkes, R. Zebrin II: A polypeptide antigen expressed selectively by purkinje cells reveals compartments in rat and fish cerebellum. *The Journal of Comparative Neurology*, 291 (4):538–552, 1990. (Cited on page 32.)

Brunel, N., Hakim, V., Isope, P., Nadal, J.-P., and Barbour, B. Optimal information storage and the distribution of synaptic weights: Perceptron versus purkinje cell. *Neuron*, 43(5):745 – 757, 2004. (Cited on page 30.)

Bullmore, E. and Sporns, O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci*, 10(3):186–198, March 2009. (Cited on page 18.)

Cannon, R. C., Gleeson, P., Crook, S., Ganapathy, G., Marin, B., Piasini, E., and Silver, R. A. LEMS: A language for expressing complex biological models in concise and hierarchical form and its use in underpinning NeuroML 2. *Frontiers in Neuroinformatics*, 8(79), 2014. (Cited on pages 44, 45, 50, 65, 66, and 74.)

Carnevale, N. and Hines, M. *The NEURON book*. Cambridge University Press, 2006. (Cited on page 67.)

Caron, S. J. C., Ruta, V., Abbott, L. F., and Axel, R. Random convergence of olfactory inputs in the drosophila mushroom body. *Nature*, 497(7447):113–117, May 2013. (Cited on page 123.)

Cesana, E., Pietrajtis, K., Bidoret, C., Isope, P., D'Angelo, E., Dieudonné, S., and Forti, L. Granule cell ascending axon excitatory synapses onto golgi cells implement a potent feedback circuit in the cerebellar granular layer. *The Journal of Neuroscience*, 33(30):12430–12446, 2013. (Cited on pages 25 and 106.)

Chadderton, P., Margrie, T. W., and Hausser, M. Integration of quanta in cerebellar granule cells during sensory processing. *Nature*, 428 (6985):856–860, April 2004. (Cited on page 92.)

Chaumont, J., Guyon, N., Valera, A. M., Dugué, G. P., Popa, D., Marcaggi, P., Gautheron, V., Reibel-Foisset, S., Dieudonné, S., Stephan, A., Barrot, M., Cassel, J.-C., Dupont, J.-L., Doussau, F., Poulain, B., Selimi, F., Léna, C., and Isope, P. Clusters of cerebellar purkinje cells control their afferent climbing fiber discharge. *Proceedings of the National Academy of Sciences*, 110(40):16223–16228, 2013. (Cited on page 27.)

Chitta, R., Jin, R., Havens, T. C., and Jain, A. K. Approximate kernel k-means: Solution to large scale kernel clustering. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '11, pages 895–903, New York, NY, USA, 2011. ACM. ISBN 978-1-4503-0813-7. (Cited on page 75.)

Chiu, C.-S., Brickley, S., Jensen, K., Southwell, A., Mckinney, S., Cull-Candy, S., Mody, I., and Lester, H. A. Gaba transporter deficiency causes tremor, ataxia, nervousness, and increased gaba-induced tonic conductance in cerebellum. *The Journal of Neuroscience*, 25(12): 3234–3245, 2005. (Cited on page 122.)

Chiu, S. N., Stoyan, D., Kendall, W. S., and Mecke, J. *Stochastic Geometry and its Applications*. Wiley, 3rd edition, 2013. (Cited on page 88.)

Clark, J. and Murata, M. Relax NG specification. http://www.oasis-open.org/committees/relax-ng/spec-20011203.html, 2001. (Cited on page 65.)

Collette, A. *Python and HDF5*. O'Reilly Media, 2013. (Cited on page 67.)

Compte, A., Brunel, N., Goldman-Rakic, P. S., and Wang, X.-J. Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cerebral Cortex*, 10(9): 910–923, 2000. (Cited on page 20.)

Cornelis, H., Rodriguez, A. L., Coop, A. D., and Bower, J. M. Python as a federation tool for genesis 3.0. *PLoS ONE*, 7(1):e29018, 01 2012. (Cited on page 74.)

Courtemanche, R. and Lamarre, Y. Local field potential oscillations in primate cerebellar cortex: Synchronization with cerebral cortex during active and passive expectancy. *Journal of Neurophysiology*, 93 (4):2039–2052, 2005. (Cited on pages 30 and 117.)

Courtemanche, R., Pellerin, J.-P., and Lamarre, Y. Local field potential oscillations in primate cerebellar cortex: Modulation during active and passive expectancy. *Journal of Neurophysiology*, 88(2):771–782, 2002. (Cited on page 30.)

Cover, T. and Thomas, J. *Elements of Information Theory*. John Wiley & Sons, New Jersey, 2006. (Cited on pages 79 and 103.)

Cowan, J. D. and Sharp, D. H. Neural nets. *Quarterly Reviews of Biophysics*, 21:365–427, 1988. (Cited on page 16.)

Crowley, J. J., Fioravante, D., and Regehr, W. G. Dynamics of fast and slow inhibition from cerebellar golgi cells allow flexible control

of synaptic integration. *Neuron*, 63(6):843 – 853, 2009. (Cited on page 25.)

Daley, D. and Vere-Jones, D. *An introduction to the theory of point processes. Volume I: Elementary Theory and Methods*, volume XXI of *Springer Series in Statistics*. Springer, 2nd edition, 2003. (Cited on page 86.)

D'Angelo, E., Koekkoek, S., Lombardo, P., Solinas, S., Ros, E., Garrido, J., Schonewille, M., and Zeeuw, C. D. Timing in the cerebellum: oscillations and resonance in the granular layer. *Neuroscience*, 162 (3):805 – 815, 2009. New Insights in Cerebellar Function. (Cited on page 30.)

Davison, A. P., Hines, M., and Muller, E. Trends in programming languages for neuroscience simulations. *Frontiers in Neuroscience*, 3 (36), 2009. (Cited on page 74.)

Dayan, P. Computational modelling. *Current Opinion in Neurobiology*, 4(2):212 – 217, 1994. (Cited on page 28.)

de Sousa, G., Maex, R., Adams, R., Davey, N., and Steuber, V. Dendritic morphology predicts pattern recognition performance in multi-compartmental model neurons with and without active conductances. *Journal of Computational Neuroscience*, pages 1–14, 2014. (Cited on page 19.)

De Zeeuw, C. I., Hoebeek, F. E., Bosman, L. W. J., Schonewille, M., Witter, L., and Koekkoek, S. K. Spatiotemporal firing patterns in the cerebellum. *Nat Rev Neurosci*, 12(6):327–344, Jun 2011. (Cited on pages 28 and 30.)

Dean, J. and Ghemawat, S. MapReduce: Simplified data processing on large clusters. *Commun. ACM*, 51(1):107–113, January 2008. (Cited on page 75.)

Dean, P., Porrill, J., Ekerot, C.-F., and Jörntell, H. The cerebellar microcircuit as an adaptive filter: experimental and computational evidence. *Nat Rev Neurosci*, 11(1):30–43, Jan 2010. (Cited on page 28.)

Deans, M. R., Gibson, J. R., Sellitto, C., Connors, B. W., and Paul, D. L. Synchronous activity of inhibitory networks in neocortex requires electrical synapses containing connexin36. *Neuron*, 31(3):477 – 485, 2001. (Cited on pages 30 and 91.)

Deb, K. and Padhye, N. Development of efficient particle swarm optimizers by using concepts from evolutionary algorithms. In *Proceedings of the 12th annual conference on Genetic and evolutionary computation*, pages 55–62. ACM, 2010. (Cited on pages 48 and 52.)

Dieudonné, S. Submillisecond kinetics and low efficacy of parallel fibre-golgi cell synaptic currents in the rat cerebellum. *The Journal of Physiology*, 510(3):845–866, 1998. (Cited on pages 24, 53, and 56.)

Dieudonné, S. and Dumoulin, A. Serotonin-driven long-range inhibitory connections in the cerebellar cortex. *The Journal of Neuroscience*, 20(5):1837–1848, 2000. (Cited on page 24.)

Diggle, P. J. Displaced amacrine cells in the retina of a rabbit: analysis of a bivariate spatial point pattern. *Journal of Neuroscience Methods*, 18(1–2):115 – 125, 1986. (Cited on page 115.)

Diggle, P. J. *Statistical Analysis of Spatial and Spatio-Temporal Point Patterns*. Monographs on Statistics and Applied Probability. CRC Press, 3rd edition, 2013. (Cited on pages 32, 35, and 86.)

Diggle, P., Eglen, S., and Troy, J. Modelling the bivariate spatial distribution of amacrine cells. In Baddeley, A., Gregori, P., Mateu, J., Stoica, R., and Stoyan, D., editors, *Case Studies in Spatial Point Process Modeling*, volume 185 of *Lecture Notes in Statistics*, pages 215–233. Springer New York, 2006. ISBN 978-0-387-28311-1. (Cited on page 115.)

DiGregorio, D. A., Nusser, Z., and Silver, R. Spillover of glutamate onto synaptic {AMPA} receptors enhances fast transmission at a cerebellar synapse. *Neuron*, 35(3):521 – 533, 2002. (Cited on pages 23, 25, and 101.)

Diño, M. R., Nunzi, M. G., Anelli, R., and Mugnaini, E. Unipolar brush cells of the vestibulocerebellum: afferents and targets. In N.M. Gerrits, C. d. Z. T.J.H. Ruigrok, editor, *Cerebellar modules: Molecules, morphology and function*, volume 124 of *Progress in Brain Research*, pages 123 – 137. Elsevier, 2000. (Cited on page 23.)

DLMF. Nist digital library of mathematical functions. http://dlmf.nist.gov/, Release 1.0.9 of 2014-08-29, 2014. URL `http://dlmf.nist.gov/`. Online companion to Olver et al. [2010]. (Cited on pages 42 and 140.)

Dow, R. and Moruzzi, G. *The physiology and pathology of the cerebellum*. Univ Of Minnesota Press, 1958. (Cited on page 27.)

Duguid, I., Branco, T., London, M., Chadderton, P., and Häusser, M. Tonic inhibition enhances fidelity of sensory information transmission in the cerebellar cortex. *The Journal of Neuroscience*, 32(32): 11132–11143, 2012. (Cited on pages 25, 106, and 122.)

Dugué, G. P., Brunel, N., Hakim, V., Schwartz, E., Chat, M., Lévesque, M., Courtemanche, R., Léna, C., and Dieudonné, S. Electrical coupling mediates tunable low-frequency oscillations and resonance in

the cerebellar golgi cell network. *Neuron*, 61(1):126 – 139, 2009. (Cited on pages 24, 30, 45, 87, 90, 91, 116, 117, and 118.)

Dumoulin, A., Triller, A., and Dieudonné, S. Ipsc kinetics at identified gabaergic and mixed gabaergic and glycinergic synapses onto cerebellar golgi cells. *The Journal of Neuroscience*, 21(16):6045–6057, 2001. (Cited on page 24.)

Dupont, J.-L., Valera, A., Cruveiller, H., Piasini, E., Silver, A., Poulain, B., and Isope, P. Unpublished observations. Unpublished, 2015. (Cited on page 36.)

D'Angelo, E. and De Zeeuw, C. I. Timing and plasticity in the cerebellum: focus on the granular layer. *Trends in Neurosciences*, 32(1): 30 – 40, 2009. (Cited on pages 101 and 122.)

Eccles, J. C., Ito, M., and Szentágothai, J. *The Cerebellum as a Neuronal Machine*. Springer-Verlag, 1967. (Cited on pages 20, 23, 24, 25, 27, 58, 59, 61, 115, and 121.)

Efron, B. Better bootstrap confidence intervals. *Journal of the American Statistical Association*, 82(397):171–185, 1987. (Cited on page 35.)

Efron, B. and Tibshirani, R. J. *An introduction to the bootstrap*, volume 57 of *Monographs on Statistics and Applied Probability*. CRC press, 1994. (Cited on page 35.)

Elgohary, A., Farahat, A. K., Kamel, M. S., and Karray, F. Embed and conquer: Scalable embeddings for kernel k-means on MapReduce. In *Proceedings of the 2014 SIAM International Conference on Data Mining*, pages 425–433. SIAM, 2014. (Cited on page 75.)

Engel, A. and Van den Broeck, C. *Statistical mechanics of learning*. Cambridge University Press, 2001. (Cited on page 29.)

Erdős, P. and Rényi, A. On random graphs. *Publicationes Mathematicae Debrecen*, 6:290–297, 1959. (Cited on page 94.)

Farrant, M. and Nusser, Z. Variations on an inhibitory theme: phasic and tonic activation of gabaa receptors. *Nat Rev Neurosci*, 6(3):215–229, March 2005. (Cited on page 25.)

Filippone, M., Camastra, F., Masulli, F., and Rovetta, S. A survey of kernel and spectral methods for clustering. *Pattern Recognition*, 41 (1):176 – 190, 2008. (Cited on page 68.)

FitzGerald, M. J. T. and Folan-Curran, J. *Clinical neuroanatomy and related neuroscience*. WB Saunders, 2002. (Cited on page 26.)

Forti, L., Cesana, E., Mapelli, J., and D'Angelo, E. Ionic mechanisms of autorhythmic firing in rat cerebellar golgi cells. *The Journal of Physiology*, 574(3):711–729, 2006. (Cited on page 24.)

Fox, C. A., Hillman, D. E., Siegesmund, K. A., and Dutta, C. R. The primate cerebellar cortex: a golgi and electron microscopic study. *Prog Brain Res*, 25:174–225, 1967. (Cited on page 23.)

Franceschetti, M. and Meester, R. *Random Networks for Communication: From Statistical Physics to Information Systems*. Cambridge University Press, Cambridge, 2007. (Cited on page 88.)

Fujita, M. Adaptive filter model of the cerebellum. *Biological Cybernetics*, 45:195–206, 1982. 10.1007/BF00336192. (Cited on pages 27 and 101.)

Gabbiani, F., Krapp, H. G., Koch, C., and Laurent, G. Multiplicative computation in a visual neuron sensitive to looming. *Nature*, 420 (6913):320–324, November 2002. (Cited on page 19.)

Galarreta, M. and Hestrin, S. A network of fast-spiking cells in the neocortex connected by electrical synapses. *Nature*, 402(6757):72–75, November 1999. (Cited on page 30.)

Galliano, E., Gao, Z., Schonewille, M., Todorov, B., Simons, E., Pop, A. S., D'Angelo, E., van den Maagdenberg, A. M. J. M., Hoebeek, F. E., and De Zeeuw, C. I. Silencing the majority of cerebellar granule cells uncovers their essential role in motor learning and consolidation. *Cell Rep*, 3(4):1239–1251, Apr 2013. (Cited on page 121.)

Gandolfi, D., Pozzi, P., Tognolina, M., Chirico, G., Mapelli, J., and D'Angelo, E. The spatiotemporal organization of cerebellar network activity resolved by two-photon imaging of multiple single neurons. *Frontiers in Cellular Neuroscience*, 8(92), 2014. (Cited on page 114.)

Gao, Z., van Beugen, B. J., and De Zeeuw, C. I. Distributed synergistic plasticity and cerebellar learning. *Nat Rev Neurosci*, 13(9):619–635, September 2012. (Cited on page 27.)

Gerstner, W. and Kistler, W. M. *Spiking neuron models: Single neurons, populations, plasticity*. Cambridge university press, 2002. (Cited on page 76.)

Geurts, F., Timmermans, J.-P., Shigemoto, R., and Schutter, E. D. Morphological and neurochemical differentiation of large granular layer interneurons in the adult rat cerebellum. *Neuroscience*, 104 (2):499 – 512, 2001. (Cited on page 36.)

Gleeson, P., Cantarelli, M., Piasini, E., and Silver, R. Advanced 3D visualisation of detailed neuronal models using the Open Source Brain repository and interaction with other neuroinformatics resources. *BMC Neuroscience*, 14(Suppl 1):P363, 2013. (Cited on page 67.)

Gleeson, P., Steuber, V., and Silver, R. A. neuroconstruct: A tool for modeling networks of neurons in 3d space. *Neuron*, 54(2):219 – 235, 2007. (Cited on page 67.)

Gleeson, P., Crook, S., Cannon, R. C., Hines, M. L., Billings, G. O., Farinella, M., Morse, T. M., Davison, A. P., Ray, S., Bhalla, U. S., Barnes, S. R., Dimitrova, Y. D., and Silver, R. A. Neuroml: a language for describing data driven models of neurons and networks with a high degree of biological detail. *PLoS Comput Biol*, 6(6): e1000815, Jun 2010. (Cited on page 65.)

Gleeson, P., Piasini, E., Crook, S., Cannon, R., Steuber, V., Jaeger, D., Solinas, S., D'Angelo, E., and Silver, R. A. The Open Source Brain initiative: enabling collaborative modelling in computational neuroscience. *BMC Neuroscience*, 13(Suppl 1):O7, 2012. (Cited on pages 45 and 67.)

Glorot, X. and Bengio, Y. Understanding the difficulty of training deep feed-forward neural networks. In *Proc. AISTAT*, 2010. (Cited on page 17.)

Goldstein, S. S. and Rall, W. Changes of action potential shape and velocity for changing core conductor geometry. *Biophysical Journal*, 14(10):731 – 757, 1974. (Cited on page 19.)

Golgi, C. *Sulla fina anatomia degli organi centrali del sistema nervoso*. Calderini, Reggio Emilia, 1885. (Cited on page 27.)

Goodman, D. F. M. and Brette, R. The brian simulator. *Frontiers in Neuroscience*, 3(26), 2009. (Cited on page 74.)

Haenggi, M., Andrews, J., Baccelli, F., Dousse, O., and Franceschetti, M. Stochastic geometry and random graphs for the analysis and design of wireless networks. *Selected Areas in Communications, IEEE Journal on*, 27(7):1029–1046, September 2009. (Cited on page 88.)

Hagberg, A. A., Schult, D. A., and Swart, P. J. Exploring network structure, dynamics, and function using NetworkX. In *Proceedings of the 7th Python in Science Conference (SciPy2008)*, pages 11–15, Pasadena, CA USA, August 2008. (Cited on pages 67 and 83.)

Harris, K. D. and Thiele, A. Cortical state and attention. *Nat Rev Neurosci*, 12(9):509–523, September 2011. (Cited on page 117.)

Hartmann, M. J. and Bower, J. M. Oscillatory activity in the cerebellar hemispheres of unrestrained rats. *Journal of Neurophysiology*, 80(3): 1598–1604, 1998. (Cited on pages 30, 91, and 117.)

Hausser, J. and Strimmer, K. Entropy inference and the james-stein estimator, with application to nonlinear gene association networks. *J. Mach. Learn. Res.*, 10:1469–1484, December 2009. (Cited on page 113.)

Hertz, J., Krogh, A., and Palmer, R. *Introduction to the Theory of Neural Computation*, volume 1 of *Santa Fe Institute Studies in the Sciences of Complexity*. Addison-Wesley, 1991. (Cited on page 16.)

Hines, M., Morse, T., Migliore, M., Carnevale, N., and Shepherd, G. Modeldb: A database to support computational neuroscience. *Journal of Computational Neuroscience*, 17(1):7–11, 2004. (Cited on page 67.)

Hinton, G. E. and Salakhutdinov, R. R. Reducing the dimensionality of data with neural networks. *Science*, 313:504–507, 2006. (Cited on page 17.)

Hinton, G. E. and Sejnowski, T. J. Optimal perceptual inference. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1983. (Cited on page 17.)

Hopfield, J. J. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 79(8):2554–2558, 1982. (Cited on page 17.)

Houghton, C. and Kreuz, T. On the efficient calculation of van rossum distances. *Network: Computation in Neural Systems*, 23(1-2):48–58, 2012. PMID: 22568695. (Cited on pages 68, 69, 74, and 75.)

Houghton, C. and Sen, K. A new multineuron spike train metric. *Neural Comput*, 20(6):1495–1511, Jun 2008. (Cited on page 68.)

Hoyer, P. O. Non-negative matrix factorization with sparseness constraints. *J. Mach. Learn. Res.*, 5:1457–1469, December 2004. (Cited on page 64.)

Huang, C.-C., Sugino, K., Shima, Y., Guo, C., Bai, S., Mensh, B. D., Nelson, S. B., and Hantman, A. W. Convergence of pontine and proprioceptive streams onto multimodal cerebellar granule cells. *eLife*, 2, 2013. (Cited on pages 25, 106, and 121.)

Hubel, D. H. and Wiesel, T. N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, 160(1):106–154, 1962. (Cited on page 19.)

Hull, C. and Regehr, W. G. Identification of an inhibitory circuit that regulates cerebellar golgi cell activity. *Neuron*, 73(1):149–158, 2012. (Cited on pages 24, 100, and 118.)

Hurley, N. and Rickard, S. Comparing measures of sparsity. *Information Theory, IEEE Transactions on*, 55(10):4723–4741, 2009. (Cited on page 64.)

Illian, J., Penttinen, A., Stoyan, H., and Stoyan, D. *Statistical analysis and modelling of spatial point patterns*, volume 70. John Wiley & Sons, 2008. (Cited on pages 32, 34, 35, 86, and 87.)

Ince, R. A. A., Petersen, R. S., Swan, D. C., and Panzeri, S. Python for information theoretic analysis of neural data. *Frontiers in Neuroinformatics*, 3:4, 2009. (Cited on page 63.)

Ito, M. Neurophysiological aspects of the cerebellar motor control system. *International Journal of Neurology*, 7(2):162, 1970. (Cited on page 27.)

Ito, M. Cerebellar circuitry as a neuronal machine. *Progress in Neurobiology*, 78(3–5):272 – 303, 2006. <ce:title>The Contributions of John Carew Eccles to Contemporary Neuroscience</ce:title>. (Cited on page 30.)

Ito, M. Control of mental activities by internal models in the cerebellum. *Nat Rev Neurosci*, 9(4):304–313, Apr 2008. (Cited on page 27.)

Ito, M. and Kano, M. Long-lasting depression of parallel fiber-purkinje cell transmission induced by conjunctive stimulation of parallel fibers and climbing fibers in the cerebellar cortex. *Neuroscience Letters*, 33(3):253 – 258, 1982. (Cited on pages 27 and 29.)

Jaeger, H. The "echo state" approach to analysing and training recurrent neural networks. Technical Report GMD Report 148, German National Research Center for Information Technology, 2001. (Cited on page 18.)

Jakab, R. L. and Hámori, J. Quantitative morphology and synaptology of cerebellar glomeruli in the rat. *Anatomy and Embryology*, 179:81–88, 1988. 10.1007/BF00305102. (Cited on pages 24 and 25.)

Jolliffe, I. T. *Principal Component Analysis*. Springer series in statistics. Springer-Verlag, 2nd edition, 2002. (Cited on page 111.)

Jörntell, H. and Ekerot, C.-F. Reciprocal bidirectional plasticity of parallel fiber receptive fields in cerebellar purkinje cells and their afferent interneurons. *Neuron*, 34(5):797–806, 2002. (Cited on pages 27 and 30.)

Jörntell, H. and Ekerot, C.-F. Properties of somatosensory synaptic integration in cerebellar granule cells in vivo. *J Neurosci*, 26(45): 11786–11797, Nov 2006. (Cited on pages 25, 112, 121, and 122.)

Kandel, E. R., Schwartz, J. H., Jessell, T. M., Siegelbaum, S. A., and Hudspeth, A. J. *Principles of Neural Science*. McGraw-Hill, New York, fifth edition, 2012. (Cited on page 26.)

Kanerva, P. *Sparse Distributed Memory*. The MIT Press, 1988. (Cited on pages 28, 101, and 120.)

Kanichay, R. T. and Silver, R. A. Synaptic and cellular properties of the feedforward inhibitory circuit within the input layer of the cerebellar cortex. *The Journal of Neuroscience*, 28(36):8955–8967, 2008. (Cited on pages 24, 25, 52, 55, and 106.)

Kanichay, R. T. *Feed-forward excitation of interneurons in the cerebellar granule cell layer*. PhD thesis, University College London, 2008. (Cited on page 24.)

Kennedy, A., Wayne, G., Kaifosh, P., Alvina, K., Abbott, L. F., and Sawtell, N. B. A temporal basis for predicting the sensory consequences of motor commands in an electric fish. *Nat Neurosci*, 17(3):416–422, March 2014. (Cited on pages 27, 112, 122, and 123.)

Kitazawa, S., Kimura, T., and Yin, P.-B. Cerebellar complex spikes encode both destinations and errors in arm movements. *Nature*, 392(6675):494–497, April 1998. (Cited on pages 27 and 30.)

Kodama, K., Iikubo, S., Taguchi, T., and Shamoto, S.-i. Finite size effects of nanoparticles on the atomic pair distribution functions. *Acta Crystallographica Section A*, 62(6):444–453, Nov 2006. (Cited on pages 38 and 42.)

Koetter, R., Bednar, J., Davison, A., Diesmann, M., Gewaltig, M.-O., Hines, M., and Muller, E., editors. *Python in neuroscience*. Frontiers in Neuroinformatics, 2009. (Cited on page 74.)

Kreuz, T., Chicharro, D., Greschner, M., and Andrzejak, R. G. Time-resolved and time-scale adaptive measures of spike train synchrony. *J Neurosci Methods*, 195(1):92–106, Jan 2011. (Cited on page 68.)

Kreuz, T., Chicharro, D., Houghton, C., Andrzejak, R. G., and Mormann, F. Monitoring spike train synchrony. *Journal of Neurophysiology*, 109(5):1457–1472, 2013. (Cited on page 94.)

Kuramoto, Y. Self-entrainment of a population of coupled nonlinear oscillators. In Araki, H., editor, *International Symposium on Mathematical Problems in Theoretical Physics*, volume 39 of *Lecture Notes in Physics*, pages 420–422, New York, NY, 1975. Springer. (Cited on page 18.)

Lago-Fernández, L. F., Huerta, R., Corbacho, F., and Sigüenza, J. A. Fast response and temporal coherent oscillations in small-world networks. *Phys. Rev. Lett.*, 84:2758–2761, Mar 2000. (Cited on page 18.)

Larramendi, L. M. H. and Lemkey-Johnston, N. The distribution of recurrent purkinje collateral synapses in the mouse cerebellar cortex: An electron microscopic study. *The Journal of Comparative Neurology*, 138(4):451–482, 1970. (Cited on page 24.)

Larson, S. D. and Martone, M. Neurolex.org: An online framework for neuroscience knowledge. *Frontiers in Neuroinformatics*, 7(18), 2013. (Cited on page 67.)

Laurent, G. Olfactory network dynamics and the coding of multidimensional signals. *Nat Rev Neurosci*, 3(11):884–895, November 2002. (Cited on page 123.)

Lee, S., Yoon, B.-E., Berglund, K., Oh, S.-J., Park, H., Shin, H.-S., Augustine, G. J., and Lee, C. J. Channel-mediated tonic gaba release from glia. *Science*, 330(6005):790–796, 2010. (Cited on page 25.)

Lewis, W. H., editor. *Gray's Anatomy of the Human Body*. Lea & Febiger, 20th edition, 1918. (Cited on page 20.)

Li, Z. and Hopfield, J. Modeling the olfactory bulb and its neural oscillatory processings. *Biological Cybernetics*, 61(5):379–392, 1989. (Cited on page 19.)

Lin, A. C., Bygrave, A. M., de Calignon, A., Lee, T., and Miesenbock, G. Sparse, decorrelated odor coding in the mushroom body enhances learned odor discrimination. *Nat Neurosci*, 17(4):559–568, April 2014. (Cited on page 123.)

Llinas, R. *Neurobiology of Cerebellar Evolution and Development*. American Medical Association, 1969. (Cited on page 123.)

Lloyd, S. Least squares quantization in PCM. *Information Theory, IEEE Transactions on*, 28(2):129–137, Mar 1982. (Cited on page 63.)

Lukoševičius, M. and Jaeger, H. Reservoir computing approaches to recurrent neural network training. *Computer Science Review*, 3:127–149, 2009. (Cited on pages 16 and 18.)

Lőrincz, A., Szoboszlay, M., Lanore, F., Vervaeke, K., Silver, R. A., and Nusser, Z. Unpublished observations. Unpublished, 2015. (Cited on pages 90, 116, and 117.)

Maass, W., Natschläger, T., and Markram, H. Real-time computing without stable states: A new framework for neural computation

based on perturbations. *Neural Computation*, 14:2531–2560, 2002. (Cited on page 18.)

Maex, R. and De Schutter, E. Synchronization of golgi and granule cell firing in a detailed network model of the cerebellar granule cell layer. *J Neurophysiol*, 80(5):2521–2537, Nov 1998. (Cited on pages 28 and 30.)

Mainen, Z. F. and Sejnowski, T. J. Influence of dendritic structure on firing pattern in model neocortical neurons. *Nature*, 382(6589): 363–366, July 1996. (Cited on page 19.)

Mann-Metzer, P. and Yarom, Y. Electrotonic coupling interacts with intrinsic properties to generate synchronized activity in cerebellar networks of inhibitory interneurons. *The Journal of Neuroscience*, 19 (9):3298–3306, 1999. (Cited on page 117.)

Mapelli, J., Gandolfi, D., and D'Angelo, E. Combinatorial responses controlled by synaptic inhibition in the cerebellum granular layer. *Journal of Neurophysiology*, 103(1):250–261, 2010. (Cited on page 114.)

Marr, D. A theory of cerebellar cortex. *The Journal of Physiology*, 202 (2):437–470, 1969. (Cited on pages 17, 27, 28, 29, 101, 120, and 122.)

Marr, D. A theory for cerebral neocortex. *Proceedings of the Royal Society of London B*, 176:161–234, 1970. (Cited on page 17.)

Marr, D. Simple memory: A theory for archicortex. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 262: 23–81, 1971. (Cited on page 17.)

McCulloch, W. S. and Pitts, W. A logical calculus of ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5(4):115–133, 1943. (Cited on page 16.)

Medina, J. F. and Mauk, M. D. Computer simulation of cerebellar information processing. *Nat Neurosci*, 3 Suppl:1205–1211, Nov 2000. (Cited on pages 28, 101, and 122.)

Meester, R. *Continuum Percolation*. Cambridge University Press, 1996. (Cited on page 88.)

Mugnaini, E., Osen, K., Dahl, A.-L., Friedrich, J. VictorL., and Korte, G. Fine structure of granule cells and related interneurons (termed golgi cells) in the cochlear nuclear complex of cat, rat and mouse. *Journal of Neurocytology*, 9(4):537–570, 1980. (Cited on page 123.)

Mugnaini, E., Sekerková, G., and Martina, M. The unipolar brush cell: A remarkable neuron finally receiving deserved attention. *Brain Research Reviews*, 66(1–2):220 – 245, 2011. Camillo Golgi and Modern Neuroscience. (Cited on pages 26 and 112.)

Muller, E., Davison, A. P., Brizzi, T., Bruederle, D., Eppler, M. J., Kremkow, J., Pecevski, D., Perrinet, L., Schmuker, M., and Yger, P. NeuralEnsemble.Org: Unifying neural simulators in Python to ease the model complexity bottleneck. In *Frontiers in Neuroscience Conference Abstract: Neuroinformatics 2009*, 2009. (Cited on page 74.)

Mézard, M., Parisi, G., and Zee, A. Spectra of euclidean random matrices. *Nuclear Physics B*, 559(3):689 – 701, 1999. (Cited on page 119.)

Møller, J. and Waagepetersen, R. *Statistical Inference and Simulation for Spatial Point Processes*. Chapman and Hall/CRC, Boca Raton, 2003. (Cited on page 115.)

Møller, J. and Waagepetersen, R. P. Modern statistics for spatial point processes. *Scandinavian Journal of Statistics*, 34(4):643–684, 2007. (Cited on page 32.)

Nemenman, I., Shafee, F., and Bialek, W. Entropy and inference, revisited. In *Advances in Neural Information Processing Systems 14*. MIT Press, 2002. (Cited on pages 63 and 113.)

Newman, M. E. J. *Networks: An Introduction*. Oxford University Press, New York, 2010. (Cited on pages 18, 82, 87, 94, and 96.)

Nielsen, T. A., DiGregorio, D. A., and Silver, R. Modulation of glutamate mobility reveals the mechanism underlying slow-rising ampar epscs and the diffusion coefficient in the synaptic cleft. *Neuron*, 42 (5):757–771, 2004. (Cited on page 25.)

O'Connor, S. M., Berg, R. W., and Kleinfeld, D. Coherent electrical activity between vibrissa sensory areas of cerebellum and neocortex is enhanced during free whisking. *Journal of Neurophysiology*, 87(4): 2137–2148, 2002. (Cited on pages 30 and 117.)

Ohser, J. and Mücklich, F. *Statistical Analysis of Microstructures in Materials Science*. Wiley, 2000. (Cited on page 35.)

Olshausen, B. A. and Field, D. J. Sparse coding of sensory inputs. *Current Opinion in Neurobiology*, 14(4):481–487, August 2004. (Cited on pages 64 and 101.)

Olver, F. W. J., Lozier, D. W., Boisvert, R. F., and Clark, C. W., editors. *NIST Handbook of Mathematical Functions*. Cambridge University Press, New York, NY, 2010. Print companion to DLMF. (Cited on page 131.)

Ostojic, S., Brunel, N., and Hakim, V. Synchronization properties of networks of electrically coupled neurons in the presence of noise and heterogeneities. *Journal of Computational Neuroscience*, 26(3):369–392, 2009. (Cited on pages 31, 87, 91, 100, 116, 118, and 119.)

Paiva, A., Park, I., and Príncipe, J. A comparison of binless spike train measures. *Neural Computing & Applications*, 19:405–419, 2010. 10.1007/s00521-009-0307-6. (Cited on page 68.)

Palay, S. L. and Chan-Palay, V. *Cerebellar cortex: cytology and organization*. Springer, 1974. (Cited on pages 23, 24, 115, and 117.)

Palkovits, M., Magyar, P., and Szentágothai, J. Quantitative histological analysis of the cerebellar cortex in the cat. iv. mossy fiber-purkinje cell numerical transfer. *Brain Research*, 45(1):15 – 29, 1972. (Cited on pages 23, 58, 61, and 123.)

Paninski, L. Estimation of entropy and mutual information. *Neural Computation*, 15(6):1191–1253, June 2003. (Cited on page 113.)

Panzeri, S. and Treves, A. Analytical estimates of limited sampling biases in different information measures. *Network - Comp. Neural.*, 7:87–107, 1996. (Cited on page 63.)

Panzeri, S., Senatore, R., Montemurro, M. A., and Petersen, R. S. Correcting for the sampling bias problem in spike train information measures. *J Neurophysiol*, 98(3):1064–1072, Sep 2007. (Cited on pages 63 and 113.)

Papadopoulou, M., Cassenaer, S., Nowotny, T., and Laurent, G. Normalization for sparse encoding of odors by a wide-field interneuron. *Science*, 332(6030):721–725, 2011. (Cited on page 123.)

Park, S. Y. and Bera, A. K. Maximum entropy autoregressive conditional heteroskedasticity model. *Journal of Econometrics*, 150(2):219 – 230, 2009. Recent Development in Financial Econometrics. (Cited on page 77.)

Parker, D. B. Learning logic. Technical Report TR-47, Center for Computational Research in Economics and Management Science, Massachusetts Institute of Technology, Cambridge, MA, 1985. (Cited on page 17.)

Paulin, M. G. Evolution of the cerebellum as a neuronal machine for bayesian state estimation. *Journal of Neural Engineering*, 2(3):S219, 2005. (Cited on page 28.)

Pecora, L. M. and Carroll, T. L. Master stability functions for synchronized coupled systems. *Phys. Rev. Lett.*, 80:2109–2112, Mar 1998. (Cited on pages 100 and 119.)

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. Scikit-learn: Machine learning in Python. *Journal*

of Machine Learning Research, 12:2825–2830, 2011. (Cited on pages 63 and 74.)

Pellerin, J.-P. and Lamarre, Y. Local field potential oscillations in primate cerebellar cortex during voluntary movement. *Journal of Neurophysiology*, 78(6):3502–3507, 1997. (Cited on pages 30, 91, and 117.)

Pellionisz, A. and Llinas, R. Space-time representation in the brain. the cerebellum as a predictive space-time metric tensor. *Neuroscience*, 7(12):2949 – 2970, 1982. (Cited on page 28.)

Pereira, T., Eroglu, D., Bagci, G. B., Tirnakli, U., and Jensen, H. J. Connectivity-driven coherence in complex networks. *Physical Review Letters*, 110:234103, Jun 2013. (Cited on pages 18, 100, and 119.)

Pikovsky, A., Rosenblum, M., and Khurts, J. *Synchronization: A Universal Concept in Nonlinear Sciences*. Cambridge University Press, 2003. (Cited on pages 18 and 87.)

Poirazi, P., Brannon, T., and Mel, B. W. Pyramidal neuron as two-layer neural network. *Neuron*, 37(6):989 – 999, 2003. (Cited on page 19.)

Pressé, S., Ghosh, K., Lee, J., and Dill, K. Principles of maximum entropy and maximum caliber in statistical physics. *Rev. Mod. Phys.*, 85:1115–1141, Jul 2013. (Cited on page 76.)

Proville, R. D., Spolidoro, M., Guyon, N., Dugue, G. P., Selimi, F., Isope, P., Popa, D., and Lena, C. Cerebellum involvement in cortical sensorimotor circuits for the control of voluntary movements. *Nat Neurosci*, 17(9):1233–1239, September 2014. (Cited on page 27.)

Quian Quiroga, R. and Panzeri, S. Extracting information from neuronal populations: information theory and decoding approaches. *Nat Rev Neurosci*, 10(3):173–185, Mar 2009. (Cited on page 63.)

Rall, W. Electrophysiology of a dendritic neuron model. *Biophysical Journal*, 2(2):145–167, 1962. (Cited on page 19.)

Ramón y Cajal, S. Sobre las fibras nerviosas de la capa granulosa del cerebelo. *Rev. Trim. Histol. Norm. Patol.*, 1:107–118, 1889a. (Cited on page 27.)

Ramón y Cajal, S. Sobre las fibras nerviosas de la capa molecular del cerebelo,. *Rev. Trim. Histol. Norm. Patol.*, 1:33–42, 1889b. (Cited on page 27.)

Ramón y Cajal, S. *Histologie du système nerveux de l'homme & des vertébrés*. A. Maloine, Paris, 1909. (Cited on page 27.)

Rancz, E. A., Ishikawa, T., Duguid, I., Chadderton, P., Mahon, S., and Häusser, M. High-fidelity transmission of sensory information by single cerebellar mossy fibre boutons. *Nature*, 450(7173):1245–1248, Dec 2007. (Cited on pages 25, 92, 112, and 121.)

Ranjan, R., Khazen, G., Gambazzi, L., Ramaswamy, S., Hill, S. L., Schürmann, F., and Markram, H. Channelpedia: an integrative and interactive database for ion channels. *Frontiers in Neuroinformatics*, 5(36), 2011. (Cited on page 65.)

Ray, S. and Bhalla, U. S. Pymoose: interoperable scripting in python for moose. *Frontiers in Neuroinformatics*, 2(6), 2008. (Cited on page 74.)

Rieubland, S., Roth, A., and Häusser, M. Structured connectivity in cerebellar inhibitory networks. *Neuron*, 81(4):913–929, Feb 2014. (Cited on page 117.)

Ripley, B. D. Tests of 'randomness' for spatial point patterns. *Journal of the Royal Statistical Society. Series B (Methodological)*, 41(3):368–374, January 1979. (Cited on page 34.)

Robins, G. and Alexander, M. Small worlds among interlocking directors: Network structure and distance in bipartite graphs. *Computational & Mathematical Organization Theory*, 10(1):69–94, 2004. (Cited on pages 82 and 85.)

Rosenblatt, F. *Principles of Neurodynamics*. Spartan, New York, 1962. (Cited on page 16.)

Rossi, D. J. and Hamann, M. Spillover-mediated transmission at inhibitory synapses promoted by high affinity a6 subunit {GABAA} receptors and glomerular geometry. *Neuron*, 20(4):783 – 795, 1998. (Cited on pages 23 and 25.)

Rossi, D. J., Hamann, M., and Attwell, D. Multiple modes of gabaergic inhibition of rat cerebellar granule cells. *The Journal of Physiology*, 548(1):97–110, 2003. (Cited on page 122.)

Rothman, J. S. and Silver, R. A. Chapter thirteen - data-driven modeling of synaptic transmission and integration. In Blackwell, K. T., editor, *Computational Neuroscience*, volume 123 of *Progress in Molecular Biology and Translational Science*, pages 305 – 350. Academic Press, 2014. (Cited on pages 50 and 81.)

Rothman, J. S., Cathala, L., Steuber, V., and Silver, R. A. Synaptic depression enables neuronal gain control. *Nature*, 457(7232):1015–1018, Feb 2009. (Cited on pages 45, 48, 53, 54, and 55.)

Roš, H., Sachdev, R. N. S., Yu, Y., Šestan, N., and McCormick, D. A. Neocortical networks entrain neuronal circuits in cerebellar cortex. *The Journal of Neuroscience*, 29(33):10309–10320, 2009. (Cited on pages 30 and 91.)

Ruediger, S., Vittori, C., Bednarek, E., Genoud, C., Strata, P., Sacchetti, B., and Caroni, P. Learning-related feedforward inhibitory connectivity growth required for memory precision. *Nature*, 473(7348): 514–518, May 2011. (Cited on page 27.)

Rumelhart, D. E. and Zipser, D. Feature discovery by competitive learning. *Cognitive Science*, 9:75–112, 1985. (Cited on page 17.)

Rössert, C., Solinas, S., D'Angelo, E., Dean, P., and Porrill, J. Model cerebellar granule cells can faithfully transmit modulated firing rate signals. *Frontiers in Cellular Neuroscience*, 8(304), 2014. (Cited on page 28.)

Salinas, E. and Abbott, L. Transfer of coded information from sensory to motor networks. *The Journal of Neuroscience*, 15(10):6461–6474, 1995. (Cited on page 19.)

Santamaria, F., Tripp, P. G., and Bower, J. M. Feedforward inhibition controls the spread of granule cell-induced purkinje cell activity in the cerebellar cortex. *Journal of Neurophysiology*, 97(1):248–263, 2007. (Cited on page 28.)

Sargent, P. B., Saviane, C., Nielsen, T. A., DiGregorio, D. A., and Silver, R. A. Rapid vesicular release, quantal variability, and spillover contribute to the precision and reliability of transmission at a glomerular synapse. *J Neurosci*, 25(36):8173–8187, Sep 2005. (Cited on pages 24, 48, 53, 107, and 112.)

Saviane, C. and Silver, R. A. Fast vesicle reloading and a large pool sustain high bandwidth transmission at a central synapse. *Nature*, 439(7079):983–987, Feb 2006. (Cited on page 121.)

Sawtell, N. B. Multimodal integration in granule cells as a basis for associative plasticity and sensory prediction in a cerebellum-like circuit. *Neuron*, 66(4):573 – 584, 2010. (Cited on page 27.)

Schmidhuber, J. Deep learning in neural networks: An overview. *Neural Networks*, 61(0):85 – 117, 2015. (Cited on pages 16 and 17.)

Schonewille, M., Gao, Z., Boele, H.-J., Vinueza Veloz, M. F., Amerika, W. E., Šimek, A. A., De Jeu, M. T., Steinberg, J. P., Takamiya, K., Hoebeek, F. E., Linden, D. J., Huganir, R. L., and De Zeeuw, C. I. Reevaluating the role of ltd in cerebellar motor learning. *Neuron*, 70(1):43–50, 2011. (Cited on page 30.)

Schrauwen, B. and Van Campenhout, J. Linking non-binned spike train kernels to several existing spike train metrics. *Neurocomputing*, 70(7-9):1247 – 1253, 2007. Advances in Computational Intelligence and Learning. 14th European Symposium on Artificial Neural Networks 2006. (Cited on page 68.)

Schreiber, S., Fellous, J., Whitmer, D., Tiesinga, P., and Sejnowski, T. A new correlation-based measure of spike timing reliability. *Neurocomputing*, 52-54(0):925 – 931, 2003. Computational Neuroscience: Trends in Research 2003. (Cited on page 68.)

Schwartz, E. J., Rothman, J. S., Dugué, G. P., Diana, M., Rousseau, C., Silver, R. A., and Dieudonné, S. Nmda receptors with incomplete mg2+ block enable low-frequency transmission through the cerebellar cortex. *The Journal of Neuroscience*, 32(20):6878–6893, 2012. (Cited on pages 25, 45, 48, 53, 54, 55, 61, 101, 107, 112, and 122.)

Schwarz, G. Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464, 03 1978. (Cited on page 58.)

Schweighofer, N., Doya, K., and Lay, F. Unsupervised learning of granule cell sparse codes enhances cerebellar adaptive control. *Neuroscience*, 103(1):35 – 50, 2001. (Cited on pages 30, 101, and 122.)

Scott, D. W. *Multivariate density estimation: theory, practice, and visualization*, volume 383. John Wiley & Sons, 2009. (Cited on page 110.)

Segev, I. and Schneidman, E. Axons as computing devices: Basic insights gained from models. *Journal of Physiology-Paris*, 93(4):263 – 270, 1999. (Cited on page 19.)

Seja, P., Schonewille, M., Spitzmaul, G., Badura, A., Klein, I., Rudhard, Y., Wisden, W., Hubner, C. A., De Zeeuw, C. I., and Jentsch, T. J. Raising cytosolic Cl- in cerebellar granule cells affects their excitability and vestibulo-ocular learning. *EMBO J*, 31(5):1217–1230, March 2012. (Cited on pages 45, 54, and 122.)

Seung, H. S. How the brain keeps the eyes still. *Proceedings of the National Academy of Sciences*, 93(23):13339–13344, 1996. (Cited on page 20.)

Shambes, G., Gibson, J., and Welker, W. Fractured somatotopy in granule cell tactile areas of rat cerebellar hemispheres revealed by micromapping; pp. 106–115. *Brain Behav Evol*, 15(2):106–115, 1978. (Cited on pages 25 and 27.)

Shamir, M. and Sompolinsky, H. Implications of neuronal diversity on population coding. *Neural Computation*, 18(8):1951–1986, June 2006. (Cited on page 107.)

Silver, R. A., Cull-Candy, S. G., and Takahashi, T. Non-nmda glutamate receptor occupancy and open probability at a rat cerebellar synapse with single and multiple release sites. *The Journal of Physiology*, 494(1):231–250, 1996. (Cited on pages 24 and 101.)

Silver, R. A., Traynelis, S. F., and Cull-Candy, S. G. Rapid-time-course miniature and evoked excitatory currents at cerebellar synapses in situ. *Nature*, 355(6356):163–166, January 1992. (Cited on pages 24 and 101.)

Simat, M., Parpan, F., and Fritschy, J.-M. Heterogeneity of glycinergic and gabaergic interneurons in the granule cell layer of mouse cerebellum. *The Journal of Comparative Neurology*, 500(1):71–83, 2007. (Cited on pages 24, 115, and 117.)

Solinas, S., Forti, L., Cesana, E., Mapelli, J., De Schutter, E., and D'Angelo, E. Computational reconstruction of pacemaking and intrinsic electroresponsiveness in cerebellar golgi cells. *Frontiers in Cellular Neuroscience*, 1(2), 2007. (Cited on pages 46 and 47.)

Solinas, S., Nieus, T., and D'Angelo, E. A realistic large-scale model of the cerebellum granular layer predicts circuit spatio-temporal filtering properties. *Front Cell Neurosci*, 4:12, 2010. (Cited on page 28.)

Sompolinsky, H. and Shapley, R. New perspectives on the mechanisms for orientation selectivity. *Current Opinion in Neurobiology*, 7 (4):514 – 522, 1997. (Cited on page 19.)

Stallman, R. M. *Free Software, Free Society*. Free Software Foundation, Boston, MA, second edition, 2010. (Cited on page 74.)

Stein, R. B. A theoretical analysis of neuronal variability. *Biophysical Journal*, 5(2):173 – 194, 1965. (Cited on page 77.)

Strogatz, S. H. From Kuramoto to Crawford: exploring the onset of synchronization in populations of coupled oscillators. *Physica D: Nonlinear Phenomena*, 143(1–4):1 – 20, 2000. (Cited on page 18.)

Strogatz, S. H. Exploring complex networks. *Nature*, 410(6825):268–276, March 2001. (Cited on page 18.)

Strong, S. P., Koberle, R., de Ruyter van Steveninck, R. R., and Bialek, W. Entropy and information in neural spike trains. *Phys. Rev. Lett.*, 80:197–200, Jan 1998. (Cited on page 63.)

Stuart, G., Spruston, N., and Häusser, M., editors. *Dendrites*. Oxford University Press, New York, 2008. (Cited on page 19.)

Subramaniyam, S., Solinas, S., Perin, P., Locatelli, F., Masetto, S., and D'Angelo, E. Computational modeling predicts the ionic mechanism of late-onset responses in unipolar brush cells. *Frontiers in Cellular Neuroscience*, 8:237–, July 2014. (Cited on page 112.)

Sultan, F. Distribution of mossy fibre rosettes in the cerebellum of cat and mice: evidence for a parasagittal organization at the single fibre level. *European Journal of Neuroscience*, 13(11):2123–2130, 2001. (Cited on pages 23 and 121.)

Sultan, F. and Glickstein, M. The cerebellum: Comparative and animal studies. *The Cerebellum*, 6(3):168–176, 2007. (Cited on page 23.)

Torben-Nielsen, B. and Stiefel, K. M. An inverse approach for elucidating dendritic function. *Front Comput Neurosci*, 4:128, 2010. (Cited on page 19.)

Traub, R. D., Kopell, N., Bibbig, A., Buhl, E. H., LeBeau, F. E. N., and Whittington, M. A. Gap junctions between interneuron dendrites can enhance synchrony of gamma oscillations in distributed networks. *The Journal of Neuroscience*, 21(23):9478–9486, 2001. (Cited on pages 30 and 91.)

Treves, A. and Panzeri, S. The upward bias in measures of information derived from limited data samples. *Neural Computation*, 2: 399–407, 1995. (Cited on page 113.)

Treves, A. and Rolls, E. T. What determines the capacity of autoassociative memories in the brain? *Network: Computation in Neural Systems*, 2(4):371–397, 1991. (Cited on page 64.)

Tripathy, S. J., Savitskaya, J., Burton, S. D., Urban, N. N., and Gerkin, R. C. Neuroelectro: A window to the world's neuron electrophysiology data. *Frontiers in Neuroinformatics*, 8(40), 2014. (Cited on page 67.)

Tsodyks, M., Pawelzik, K., and Markram, H. Neural networks with dynamic synapses. *Neural Computation*, 10(4):821–835, May 1998. (Cited on pages 50 and 66.)

Tsodyks, M. V. and 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. (Cited on pages 50 and 52.)

Turing, A. M. On computable numbers, with an application to the entscheidungsproblem. *J. of Math*, 58(345-363):5, 1936. (Cited on page 16.)

Tyrrell, T. and Willshaw, D. Cerebellar cortex: Its simulation and the relevance of marr's theory. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 336(1277):239–257, 1992. (Cited on pages 29, 30, 101, 111, and 120.)

Valera, A., Dupont, J., Heloise, C., Binda, F., Pawlowski, S., Léna, C., Rothstein, J., Poulain, B., and Isope, P. Functional organization of synaptic inputs on mouse cerebellar golgi cells related to zebrin band patterns. Program no. 580.10/nn16. In *2012 Neuroscience Meeting Planner*, New Orleans, LA, 2012. Society for Neuroscience. Online. (Cited on pages 35, 85, 86, 115, and 117.)

Van Kan, P. L., Gibson, A. R., and Houk, J. C. Movement-related inputs to intermediate cerebellum of the monkey. *Journal of Neurophysiology*, 69(1):74–94, 1993. (Cited on pages 24, 25, 61, and 121.)

Van Rossum, M. C. A novel spike distance. *Neural Comput*, 13(4): 751–763, Apr 2001. (Cited on page 68.)

Vargas-Irwin, C. E., Brandman, D. M., Zimmermann, J. B., Donoghue, J. P., and Black, M. J. Spike train similarity space (ssims): A framework for single neuron and ensemble data analysis. *Neural Computation*, 27(1):1–31, November 2014. (Cited on page 68.)

Vella, M., Cannon, R. C., Crook, S., Davison, A. P., Ganapathy, G., Robinson, H. P. C., Silver, R. A., and Gleeson, P. libneuroml and pylems: using python to combine imperative and declarative modelling approaches in computational neuroscience. *Frontiers in Neuroinformatics*, 8(38), 2014. (Cited on pages 65 and 74.)

Vervaeke, K., Lőrincz, A., Gleeson, P., Farinella, M., Nusser, Z., and Silver, R. A. Rapid desynchronization of an electrically coupled interneuron network with sparse excitatory synaptic input. *Neuron*, 67(3):435 – 451, 2010. (Cited on pages 13, 14, 28, 30, 31, 36, 37, 40, 45, 47, 52, 53, 85, 87, 88, 89, 90, 91, 92, 93, 96, 97, 98, 99, 100, 115, 116, 117, and 119.)

Vervaeke, K., Lorincz, A., Nusser, Z., and Silver, R. A. Gap junctions compensate for sublinear dendritic integration in an inhibitory network. *Science*, 335(6076):1624–1628, Mar 2012. (Cited on pages 28, 46, 47, 48, 87, 90, 115, 116, 117, 118, and 119.)

Vetter, P., Roth, A., and Häusser, M. Propagation of action potentials in dendrites depends on dendritic morphology. *Journal of Neurophysiology*, 85(2):926–937, 2001. (Cited on page 19.)

Victor, J. D. and Purpura, K. P. Nature and precision of temporal coding in visual cortex: a metric-space analysis. *Journal of Neurophysiology*, 76(2):1310–1326, 1996. (Cited on page 68.)

Vinje, W. E. and Gallant, J. L. Sparse coding and decorrelation in primary visual cortex during natural vision. *Science*, 287(5456): 1273–1276, 2000. (Cited on pages 64 and 103.)

Voogd, J. and Glickstein, M. The anatomy of the cerebellum. *Trends in Cognitive Sciences*, 2(9):307 – 313, 1998. (Cited on page 27.)

Vos, B. P., Volny-Luraghi, A., and De Schutter, E. Cerebellar golgi cells in the rat: receptive fields and timing of responses to facial stimulation. *European Journal of Neuroscience*, 11(8):2621–2634, 1999. (Cited on pages 24 and 30.)

Wager, C. G., Coull, B. A., and Lange, N. Modelling spatial intensity for replicated inhomogeneous point patterns in brain imaging. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 66(2):429–446, 2004. (Cited on page 35.)

Ward, D. R. *Golgi cell mediated inhibition in the cerebellar granule cell layer*. PhD thesis, University College London, 2012. (Cited on pages 25, 51, 52, 53, 55, 107, 112, and 122.)

Watts, D. J. and Strogatz, S. H. Collective dynamics of 'small-world' networks. *Nature*, 393:440, 1998. (Cited on pages 18, 94, and 96.)

Webster, S., Diggle, P., Clough, H., Green, R., and French, N. Strain-typing transmissible spongiform encephalopathies using replicated spatial data. In Baddeley, A., Gregori, P., Mateu, J., Stoica, R., and Stoyan, D., editors, *Case Studies in Spatial Point Process Modeling*, volume 185 of *Lecture Notes in Statistics*, pages 197–214. Springer New York, 2006. ISBN 978-0-387-28311-1. (Cited on page 35.)

Werbos, P. *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences*. PhD thesis, Harvard University, 1974. (Cited on page 17.)

Widrow, B. and Hoff, M. E. Adaptive switching circuits. In *1960 IRE WESCON Convention Records*, 1960. (Cited on page 17.)

Wiegand, T. and A. Moloney, K. Rings, circles, and null-models for point pattern analysis in ecology. *Oikos*, 104(2):209–229, 2004. (Cited on page 34.)

Williams, R. W. and Herrup, K. The control of neuron number. *Annual Review of Neuroscience*, 11(1):423–453, 1988. PMID: 3284447. (Cited on page 123.)

Wilson, G. Where's the real bottleneck in scientific computing? scientists would do well to pick up some tools widely used in the software industry. *Am Sci. v94 i1*, 2006. (Cited on page 74.)

Winfree, A. T. Biological rhytms and the behavior of populations of coupled oscillators. *Journal of Theoretical Biology*, 16:15–42, 1967. (Cited on page 18.)

Wittenberg, G. M. and Wang, S. S.-H. *Evolution and Scaling of Dendrites*, chapter 2, pages 43–67. Oxford University Press, 2008. (Cited on pages 23, 114, and 123.)

Wolpert, D. M., Miall, R., and Kawato, M. Internal models in the cerebellum. *Trends in Cognitive Sciences*, 2(9):338 – 347, 1998. (Cited on page 27.)

Wu, C. W. Perturbation of coupling matrices and its effect on the synchronizability in arrays of coupled chaotic systems. *Physics Letters A*, 319(5–6):495 – 503, 2003. (Cited on pages 18 and 119.)

Wu, H.-S., Sugihara, I., and Shinoda, Y. Projection patterns of single mossy fibers originating from the lateral reticular nucleus in the rat cerebellar cortex and nuclei. *The Journal of Comparative Neurology*, 411(1):97–118, 1999. (Cited on page 23.)

Zhang, J., Han, V. Z., Meek, J., and Bell, C. C. Granular cells of the mormyrid electrosensory lobe and postsynaptic control over presynaptic spike occurrence and amplitude through an electrical synapse. *Journal of Neurophysiology*, 97(3):2191–2203, 2007. (Cited on page 123.)

Final digital version created in London in March 2015.

This document was typeset in LaTeX with a style derived from André Miede's `classicthesis`, using GNU Emacs and AUCTeX.

Hermann Zapf's *Palatino* and *Euler* type faces (Type 1 PostScript fonts *URW Palladio L* and *FPL*) are used. The sans-serif and "typewriter" text are typeset, respectively, in *Bera Sans* and *Bera Mono*, originally developed by Bitstream, Inc. as "Bitstream Vera" (Type 1 PostScript fonts were made available by Malte Rosenau and Ulrich Dirr).