LUND UNIVERSITY

Cursed complexity. Computational properties of subcortical neuronal microcircuitry in sensorimotor control

Spanne, Anton

2015

Cursed complexity

Computational properties of subcortical neuronal microcircuitry in sensorimotor control

Anton Spanne
Neural Basis of Sensorimotor Control
Department of Experimental Medical Science

LUND
UNIVERSITY

DOCTORAL DISSERTATION
by due permission of the Faculty of Medicine, Lund University, Sweden.
To be defended at Segerfalksalen on the 7th September 2015, at 09.00 a.m.

Faculty opponent

Prof. Gerald E. Loeb, University of Southern California, Los Angeles, USA

Organization	Document name
LUND UNIVERSITY	DOCTORAL DISSERTATION
	Date of issue: 2015-09-07
Author: Anton Spanne	Sponsoring organization:

Title: Cursed complexity. Computational properties of subcortical neuronal microcircuitry in sensorimotor control

Abstract: One of the big obstacles for understanding the nervous system is its inherent complexity. It poses problems when interpreting both experimental and theoretical studies since we are currently forced to consider only reduced variants of the actual circuitry of the brain. Since there exist problems that do not appear until a system is sufficiently complex, there are no guarantees that the results stemming from such reduced studies can be extrapolated to actually apply to the real brain.

The initial part of the thesis investigates the properties of the spinocerebellar circuitry of the nervous system, and its role in motor control. Especially the cerebellum has been shown to play an important role in the coordination of fast movements, such as reaching and pointing. Paper I uses theoretical reasoning based on previously found experimental studies to show that the cerebellar circuitry should not be studied in isolation if the aim is to explore cerebellar function. The inputs provided by the pre-cerebellar circuits in the spinal cord and brain stem can significantly reduce the complexity of the problem that the cerebellar circuitry needs to solve.

Papers II, IV and V investigate the properties of the mossy fiber pathways. Both the spinal border cell neurons that ascend the ventral spinocerebellar tract with sensorimotor information related to locomotion and the neurons of the cuneate nucleus that process tactile information are studied using behavioral stimulation, either in vivo (Paper V) or through modeling (Paper IV). The results indicate both that the overall activity of the circuitry provides the cerebellum with an easy to interpret encoding, but the individual neurons can at the same time segregate underlying features and details of the stimulus. This result is a parallel to the found statistics of spike generation in Paper III. Even though the neurons have complex electrodynamic properties, their average activity, described by their firing statistics is surprisingly similar between neurons with vastly different morphology.

Paper VI reviews the theoretical grounds for sparse coding, and compares them to recent experimental findings, both in the cerebellum and the neocortex. While there are beneficial properties of certain sparse codes, the experimental results rather indicate that the circuitry both in the cerebellum and the neocortex do not actively maintain a sparse population code.

Key words: Cerebellum, spinal border cells, sparse coding, cuneate nucleus, tactile processing, sensorimotor control, spontaneous activity, spike firing statistics

Classification system and/or index terms (if any)

Supplementary bibliographical information	Language English
ISSN and key title 1652-8220	ISBN 978-91-7619-168-2
Recipient's notes	Number of pages / Price / Security classification

Signature _____Date _____

Cursed complexity

Computational properties of subcortical neuronal microcircuitry in sensorimotor control

Anton Spanne
Neural Basis of Sensorimotor Control
Department of Experimental Medical Science
Faculty of Medicine

LUND
UNIVERSITY

Faculty of Medicine, Department of Experimental Medical Science
ISBN 978-91-7619-168-2
ISSN 1652-8220

Printed in Sweden by Media-Tryck, Lund University
Lund 2015

"Everything should be made as simple as possible, but not simpler."

- Albert Einstein

Preface

One of the big obstacles for understanding the nervous system is its inherent complexity. It poses problems when interpreting both experimental and theoretical studies since we are currently forced to consider only reduced variants of the actual circuitry of the brain. Since there exist problems that do not appear until a system is sufficiently complex, there are no guarantees that the results stemming from such reduced studies can be extrapolated to actually apply to the real brain.

The initial part of the thesis investigates the properties of the spinocerebellar circuitry of the nervous system, and its role in motor control. Especially the cerebellum has been shown to play an important role in the coordination of fast movements, such as reaching and pointing. Paper I uses theoretical reasoning based on previously found experimental studies to show that the cerebellar circuitry should not be studied in isolation if the aim is to explore cerebellar function. The inputs provided by the pre-cerebellar circuits in the spinal cord and brain stem can significantly reduce the complexity of the problem that the cerebellar circuitry needs to solve.

Papers II, IV and V investigate the properties of the mossy fiber pathways. Both the spinal border cell neurons that ascend the ventral spinocerebellar tract with sensorimotor information related to locomotion and the neurons of the cuneate nucleus that process tactile information are studied using behavioral stimulation, either in vivo (Paper V) or through modeling (Paper IV). The results indicate both that the overall activity of the circuitry provides the cerebellum with an easy to interpret encoding, but the individual neurons can at the same time segregate underlying features and details of the stimulus. This result can be seen as a parallel to the found statistics of spike generation in Paper III. Even though the neurons have complex electrodynamic properties, their average activity, described by their firing statistics is surprisingly similar between neurons with vastly different morphology.

Paper VI reviews the theoretical grounds for sparse coding, and compares them to recent experimental findings, both in the cerebellum and the neocortex. While there are beneficial properties of certain sparse codes, the experimental results rather indicate that the circuitry both in the cerebellum and the neocortex do not actively maintain a sparse population code.

List of original papers

I) **Spanne A.**, Jörntell H. (2013). Processing of Multi-dimensional Sensorimotor Information in the Spinal and Cerebellar Neuronal Circuitry: A New Hypothesis. *PLOS Computational Biology* 9(3): e1002979.
doi:10.1371/journal.pcbi.1002979

II) Geborek P., **Spanne A.**, Bengtsson F., Jörntell F. (2013). Cerebellar cortical neuron responses evoked from the spinal border cell tract. *Frontiers in Neural Circuits,* 7, 157.
doi:10.3389/fncir.2013.00157

III) **Spanne A.**, Geborek P., Bengtsson F., Jörntell H. (2014). Spike generation estimated from stationary spike trains in a variety of neurons *in vivo. Frontiers in Cellular Neuroscience*, 8, 199.
doi:10.3389/fncel.2014.00199

IV) **Spanne A.**, Geborek P., Bengtsson F., Jörntell H. (2014). Simulating Spinal Border Cells and Cerebellar Granule Cells under Locomotion – A Case Study of Spinocerebellar Information Processing. *PLoS ONE* 9(9): e107793.
doi:10.1371/journal.pone.0107793

V) Jörntell H., Bengtsson F., Geborek P., **Spanne A.**, Terekhov A. V., Hayward V. (2014). Segregation of Tactile Input Features in Neurons of the Cuneate Nucleus. *Neuron* 83.6: 1444-1452.
doi:10.1016/j.neuron.2014.07.038

VI) **Spanne A.**, Jörntell H. (2015). Questioning the role of sparse coding in the brain. *Trends in Neurosciences* 38 (7), 417-427
doi:10.1016/j.tins.2015.05.005

Contents

Introduction

It is hard to not be impressed by the capabilities of the human mind, hidden within the complex information processing machine that is the brain. In fact, even what could be considered lower level functions, such as reaching to grasp something, or even just maintaining balance, are complex problems of movement control that the brain can manage with surprising ease. Much of the current research effort in robotics is focused on describing so called anthropomorphic control methods, i.e. mimicking principles from human biomechanics and the nervous system. From the perspective of these efforts, a crude definition of what the brain does is simply regulation of motor action. That is, from the outside of a creature with a brain, its nervous system could almost completely be described as a black box controller whose only purpose is to control the creature's movements. All the complexity of the brain is conveniently hidden away inside the black box and in its absence we can allow ourselves to consider to what extent we can currently describe how the nervous system actually works and what problems we need to consider in order to formulate such a description. In order to do this, there are principles that apply to all systems that process complex information that need to be considered. Two of these are presented next, and they form a backbone or thread that connects the various parts of this thesis.

The curse of dimensionality

The biomechanics of the human body contain over 200 bones and 600 skeletal muscles that are under voluntary control. The human hand alone includes 27 bones that are actuated by 36 muscles through a complex network of tendons. Consequently, the hand is by itself a complicated kinematic system, with at least 23 degrees of freedom (Santello et al., 2013). Each degree of freedom provides the nervous system with at least a single intrinsic dimension of sensory input. Due to non-linearities (e.g. hysteresis), noise, and the possible need to sample not only the state, but also its velocity and acceleration, an abundance of sensory input dimensions per degree of freedom is most likely reaching the nervous system. Considering the motor control of for example the hand, the nervous system potentially has to separately control each of the 36 muscles in order to make use of the hand.

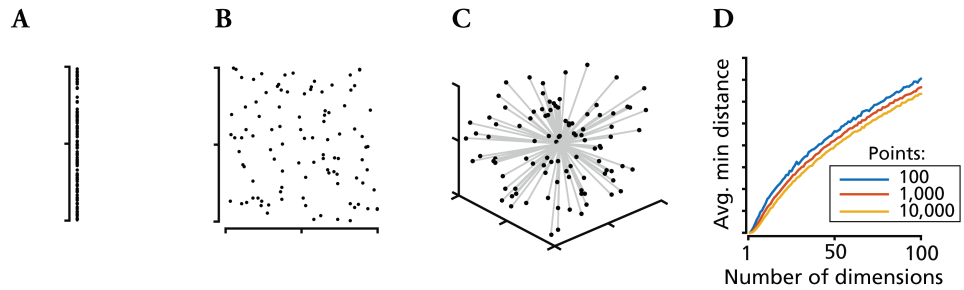

Figure 1. Illustrating the curse of dimensionality.
The curse of dimensionality is related to the number of possible permutations of input values that reach a system. The number of combinations is in turn related to the size of the space that is defined by all the input dimensions. It is not obvious how to compare the size of spaces with varying number of dimensions, but one way is to consider 100 points that have a uniform random distribution across the space in e.g. **(A)** 1 dimension, **(B)** 2 dimensions and **(C)** 3 dimensions. It is clear from the figure that the distance between the points increases from the cramped 1 dimensional line in **(A)** into the spacious 2-dimensional square in **(B)**. Due to the 3 dimensional space being projected onto a 2-dimensional paper in **(C)** the distances looks the same, even though they are actually larger than in the 2-dimensional case. **(D)** If the number of dimensions is increased further, the distance between each of the points will continue to grow. The growth is so fast that it becomes almost impossible to counter the increase in distance, even if the number of distributed points is increased 10 or 100 fold (to 1,000 and 10,000 respectively). In fact, the number of points that are necessary in this example will increase approximately 3 fold for each new dimension introduced. Comparing the 3-dimensional case in **(C)** to the 100-dimensional case in **(D)**, the 100-dimensional case would require approximately 10^{50} points to reach the same small distance. If the randomly distributed points are for example considered to be points within the input space where a neural circuitry has learned how to respond correctly, it is quite clear that it is unreasonable to get the same density of knowledge in the 100-dimensional space as in the 3-dimensional by simply increasing the number of examples (i.e. points in this case) that the circuitry can learn from. Avg. min. distance, the average minimum distance denotes the distance to the closest neighbor averaged across 100 of the points.

The '*Curse of dimensionality*' (COD) refers to the problems that arise when any process or system tries to make sense of high-dimensional information. Whereas it is not primarily a problem of the high-dimensional data, the problems arise when the algorithms or analysis methods that are used do not scale well in relation to the number of dimensions. The COD, as it is used in the thesis, follows from the rate of growth of the space that is spanned by the dimensions of the input to the system. Figure 1 tries to illustrate the growth using points randomly distributed across spaces with increasing number of dimensions. In principle, each time a dimension is added to the total input that the system receives the size of the input space will grow in proportion to the distinct states of the added input, or the size of the range of possible input values. This is a geometric rate of growth, or in other words an exponential relationship between the number of input dimensions and the size of the input space.

As a result, it will also lead to an exponential growth in the number of dynamic states or execution time of the used algorithm in the worst case scenario.

The term COD is commonly attributed to have been coined in relation to dynamic programming by Bellman (1957). Dynamic programming — a name that has an intriguing story of its own (Dreyfus, 2002) — is a method to solve complex optimal decision problems by dividing them to a series of less complex subproblems. Bellman used the COD to describe the problems arising with large state spaces of these subproblems. In the context of the nervous system this is comparable to a high dimensionality of the motor output, and how to activate the motor output to reach a certain target. In the thesis however, it will more commonly be used to describe the problems that arise with a large input space, while the motor problem is only alluded to.

In the context of the input space of the nervous system, each sensor in the periphery could in principle be considered a single dimension, providing unique afferent information. In this view, with millions of input dimensions, the circuits of the nervous system must have found effective ways of sidestepping the COD. Otherwise, the number of neurons in the brain would be far from enough to store enough information in order to cover only a fraction of the total sensory input space, and the experience we would gain during a lifetime would similarly only be enough to give meaning to an extremely small subset of all the possible combinations of inputs that potentially exist.

Minsky and Papert investigated the properties of artificial neural networks in their seminal book '*Perceptrons*' (Minsky and Papert, 1969). The focused on the computational properties of the perceptron model that had been discovered by Rosenblatt (1958). He had in turn used it to describe the properties of sensory processing by reproducing the network topology of the retina. Note that what Rosenblatt investigated and called a perceptron would today be considered a multilayered network of several units, each called a single perceptron, while Minsky and Papert considered only the properties of a single such unit in their book. They realized that it was not possible to solve to classification problems with classes that are not linearly separable using such a single unit or perceptron alone. The results presented in '*Perceptrons*' were part in a sudden widespread realization that the field of artificial intelligence had promised more than it could deliver. It led to a decline in research of artificial neural networks that lasted for a decade if not more (Russell and Norvig, 2009). Eventually it was realized that the constraints using a single perceptron could be easily overcome by adding layers to the networks, essentially returning to the model of Rosenblatt (Rumelhart et al., 1985). In addition, the rediscovery of the backpropagation algorithm gave the artificial intelligence community a tool with which it was possible to train the multilayered networks.

A second point of '*Perceptrons'* does however still remain valid. Minsky and Papert found an exponential complexity of the synaptic weights, number of neurons, and number of training examples needed in relation to the dimensionality of the input space of various classification problems. They pointed out that in order to solve a specific high-dimensional problem using a neural network and avoid the COD, the network should be custom made to handle the specific problem (Minsky and Papert, 1987). While they recognized the capabilities of the multilayered networks, they meant that the examples presented by Rumelhart et al. (1985) where custom designs that solved specific problems and did not take the exponential complexities into account. Minsky would later write: "What magical trick makes us intelligent? The trick is that there is no trick. The power of intelligence stems from our vast diversity, not from any single, perfect principle" (Minsky, 1988).

Following from their description of the undesirable complexities of shallow neural networks, Minsky and Papert cautioned against the use of what they called *toy problems*. The term describes a reduced variant of a complex problem that can be used to investigate a certain limited set of features of the original problem. If the toy problem is scaled up to the size of the original problem however, there are little guarantees that the found properties will be useful to describe the high-dimensional problem. In the original problem formulation, the limiting constraints could be something that had no influence over the toy problem, but which due to exponential complexity has enormous consequences in the high dimensional setting.

The bias-variance dilemma

The obvious first approach to avoid the COD of a high-dimensional problem is to reduce the problem that needs to be solved into a toy problem, and by that simply give up on handling all the details of the original problem. By stripping as many dimensions from the input to the system as possible, the consequences of the exponential complexities can be dramatically reduced. The number of dimensions could be reduced either by simply ignoring them using some heuristic (Gigerenzer and Brighton, 2009), or through the means of a component analysis method that can find the most important or informative intrinsic dimension within the input (Oja et al., 1995). It will however most likely also reduce the possible performance of the system, leading to a trade-off between the negative consequences of a high dimensional system versus the negative consequences of a system that by design can only learn or compute a subset of all possible input-output relations that might be useful.

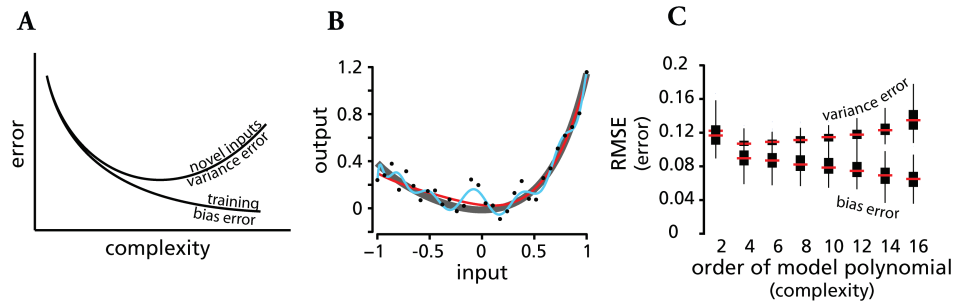

Figure 2. Illustrating the bias-variance dilemma.
(A) The bias-variance dilemma is the trade-off between a low model complexity leading to a high bias error, and a high model complexity leading to a high variance error. In many cases there is a minimum combined error where the model complexity is optimal. **(B)** Points sampled with noise from the gray line. The colored lines indicate the best polynomial fit to the sampled points of a 4th order (red) and a 15th order (blue) polynomial. The 15th order polynomial is better at interpolating the actual points, but end up a longer distance from the underlying relationship. **(C)** The relationship between the bias error and variance error of polynomials of order 2-16. The bias error decreases monotonically while the variance error increase when polynomials of order 6 or more are used. Each bar represents 100 different polynomials each fitted to a random noisy sample, similar to the points in **(B)**. The bias and variance errors are presented as means, 25th percentiles, and 75th percentiles of the 100 polynomials.

To formalize this trade-off, the error of a system that is supposed to produce a specific response in a certain situation can be divided into two parts. The first part is the error that arises due to imperfections of the model that the system uses to internally represent the relationship between a situation and the correct response. Such an error is termed the bias error, since the internal model of the system will have a systematic error due to its limitations representing the actual relationship. With a more complex internal model, the bias error can be reduced as the model will be able to represent more and more details. With a sufficiently high level of complexity, the bias error could in principle be completely removed.

The second part of the error, called the variance error, is related to the system's capability to generalize its current knowledge to handle novel situations. In contrast to the bias error, which is reduced with a high model complexity, the error that relate to the system's ability to generalize will commonly increase after a certain degree of complexity (see Figure 2A). Somewhat counterintuitively, an increase in the complexity of the model will thus degrade the performance of the system. The underlying cause is the COD and the fact that we have to train the system by examples. To understand why, consider the points in Figure 1 as the locations within the input space of the examples that have been given to the system so far. As the complexity of the system increase, e.g. by including more dimensions of the system's environment, the provided examples will be increasingly sparsely distributed within

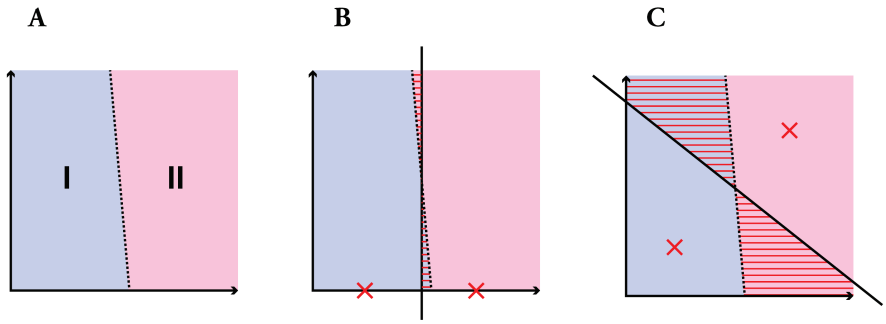

Figure 3. Even meaningful information may degrade performance.
Consider that a system should classify situations encoded in two input dimension into two distinct classes – I and II. **(A)** The 2-dimensional area defined by the two input dimensions and the regions of the area that should be classified as either class I or II. **(B)** The system is only given access to the dimension that has the largest influence upon the correct classification. If the system is given two examples of inputs and the correct class of that input (red crosses) in the middle of each class area, the system can manage to learn a close to ideal classification rule indicated by the black vertical line, even without access to the second dimension. The areas that will give erroneous classification results are indicated by red stripes. **(C)** If the system is given access to the second dimension as well without additional examples, it has no way of knowing whether it should care more about the first or second dimension. As a consequence, the best the system can do is to divide the area by a diagonal line, leading to a much larger portion of the total area that will give erroneous classification results, even though the second dimension also contains relevant information. With additional examples, the system could of course learn the exact shape of the areas, but with only 2 examples the best approach is to simply to ignore the second dimension of the input.

the space of all possible configurations of the input to the system. The consequence for the system will be that the likelihood of it finding the actual relationship between input and desired response will be significantly reduced. Instead, underlying noise or random coincidences will be interpreted as the actual relationships. The system will still respond correctly to the situations that it has encountered and learned from so far, but as soon as it is presented to a novel situation, it will use the erroneous relationships it has learned and produce an erroneous response to the novel situation.

The trade-off between a bias error due to low model complexity and a variance error due to exaggerated model complexity is commonly called the '*Bias-Variance dilemma*' (Geman et al., 1992). It is illustrated with an example in Figure 2 using polynomials of increasing degree that represent models of varying complexity. The polynomials of high degree will in general be better at interpolating the sampled points, but above a degree of 6, they see a decreased ability to generalize to new samples. It might seem that the cause of the degraded performance using higher-order polynomials in the example is only due to the noise that is introduced during the sampling of the points. Since noise is present in the nervous system, one might ask whether the variance error could be removed by efficient noise reduction.

The simple example in Figure 3 shows that this is not the case. Even the addition of noise-free and relevant information can degrade the performance of the system since it increases the complexity of the model. The bias-variance dilemma arises due to the exponential growth of the number of situations the model can represent in its naïve state, compared to the amount of training situations available. In principle, the consequence is that there is always a fundamental motivation to keep the model complexity as low as possible, e.g. by using different heuristics to reduce the complexity of the model (Gigerenzer and Brighton, 2009).

Complexity of the nervous system

One of the main impediments to our understanding of the brain is the difficulties that exist in describing the properties of complex (high-dimensional) information, and how it can be processed and used by any system. In a description of how the brain works, such an understanding needs to be incorporated. However, the current lack of a useful theoretical framework does not exclude research into the brain, but it should caution us when we design experiments, interpret data and build models of brain circuitry. In all three cases, it is currently not possible to avoid having to reduce the complexity of the studied system. By carelessly doing such a reduction, the consequence might be that the system that is studied is no longer relevant in the context of the original system of high complexity. Simply put, it is not by guarantee possible to study how a system copes with high complexity using a similar low complexity system, such as in a reduced experimental setup, or in far from natural behavioral settings. The low-complexity variant of the same system might completely lack fundamental features that are necessary for the high complexity system to function.

One method to reduce the complexity of the nervous system in order to study it in a functionally relevant way might be to mimic the process of development of a brain from birth to adult state. Unlike many artificial systems, the brain of an animal needs to be functional to some extent already before it reaches its adult state, in certain cases already minutes after birth (Roberts and Rubenstein, 2014). A consequence of this requirement is that the nervous system and the biomechanics of the young animal cannot have the full complexity of the adult animal. The peripheral nervous system and the biomechanics of the body should be constructed in a way so that it is relatively simple to build a primitive controller, which subsequently can become increasingly complex during development. This could be the process that give rise to postural synergies of the hand during grasping (Santello et al., 1998), for example. Due to the complex web of tendons that connect the joints of the hand, the individual degrees of freedom of the hand and joints cannot be controlled in

isolation. Instead, the postures of the human hand when it performs grasping tasks can be relatively well described by just a few linear combinations of the available degrees of freedom. The three most informative of these linear combinations describe basic patterns of hand use, such as opening/closing of the hand involving all fingers (Santello et al., 1998). These are at least superficially similar to the grasps elicited by the grasp reflex of an early infant (Forssberg et al., 1991). The motor control circuitry of the spinal cord and the motor pathways (Santello et al., 2013), support some type of synergistic control.

The following parts of the thesis will explore the original papers in the perspective of complexity reduction. To what extent does the nervous system, primarily in the periphery, the spinal cord and the cerebellum cope with the COD, and manage the balance between bias and variance? What happens when we apply complex stimuli that resemble that which is encountered during natural behavior to systems that have previously only been studied using reduced, artificial methods of stimulation? The two questions illustrate that we need to consider the complexity both when we describe how the brain makes sense of its complex environment, but also when we build our models to make sense of the complex brain.

Outline of the thesis

The core building block of the nervous system is the neuron. In Paper III we investigate the statistical firing properties of a range of neurons in the spinocerebellar circuitry, and find that it is possible to reliably emulate at least the first and second order statistical properties using a two-parameter model. This is despite the relatively complex electrodynamics of a neuron. While the finding does not exclude more complex behavior in the individual neuron, it does indicate that the neurons have a low complexity foundation in their behavior.

The real complexity of the nervous system does however arise due to the network of billions of neurons that are intricately interconnected. The cerebellum should offer interesting insights into how the brain manages this complexity, since it contains a majority of all neurons in the brain, and it also has a feed-forward anisotropic structure that should make it easier to study than the recurrent networks that make up most of the spinal cord and the thalamocortical system. In Paper I, we investigate the properties of the spinocerebellar system, and how it seems to be organized to avoid the potential explosion of complexity that could arise. We find that the cerebellum should not be studied in isolation, but rather together with the systems that provide the cerebellar input through mossy fibers. The organization of these inputs most likely plays a critical role when the spinocerebellar circuitry avoids the COD.

In order to study the response and properties of a system that provide input to the cerebellum in isolation, the properties of spinal border cell (SBC) input to the cerebellar cortex were investigated in Paper II. In line with previous findings regarding the specificity and excitability of cerebellar granule cells, the SBC tract alone could elicit strong responses not only in granule cells, but also Golgi cells, molecular layer interneurons and Purkinje cells. The responses of different granule cells were however quite diverse. In order to investigate how the difference in the response of the granule cells affected their activity during natural behavior, we investigated the response properties of modeled granule cells under locomotion in Paper IV. We found that much of the complexity of the responses could be explained as variable conduction delays within different fibers in the SBC tract. The difference in response seen with the SBC tract stimulation did not seem to influence the overall response of the simulated granule cells to the locomotion related input. Instead, the granule cells had an activity that was close to linearly modulated by the activity of the SBCs. It illustrates that while the system has dynamics that most likely allow complex responses, the overall activity will closely resemble the overall input to the network, allowing the cerebellum to reliably get an initial estimate of the input without considering all the complexity that might be hidden in the details.

The low-complexity interpretation of a circuitry does sometimes directly limit the possibility to transmit complex information along the same channel. The classical view of the somatosensory processing of the nervous system has led to such limitations in our description of the subcortical processing of tactile afferent information. We show in Paper V that the neurons of the cuneate nucleus manage to extract fundamental features from the mechanical contact of tactile stimulation from the primary afferent. The findings in Paper V challenge the functional interpretation of the somatopic map of the primary somatosensory cortex, and also the classical description of the function of secondary somatosensory cortex (Kandel and Schwarz, 2013). The contradicting results stem from the use of a novel spatiotemporal mechanical stimulation (Hayward et al., 2014), rather than pointwise mechanical or electrical stimulation that is commonly used to map the receptive fields of cortical regions. The findings in Paper V illustrate the danger of extrapolating a functional interpretation from experimental data, which has been acquired in an experimental setting with reduced complexity.

Similar issues can be found regarding the concept of sparse coding, which is widely believed to be present in many separate brain regions. We survey the current understanding of sparse coding in Paper VI, in order to investigate the theoretical foundation of sparse codes as well as the experimental evidence. While there are many theoretical studies that show beneficial properties of a sparse code (at moderate sparseness), the experimental results are often built upon reduced experimental settings. In studies where the nervous system is exposed to stimulus that resemble

what would be experienced during natural behavior, the network activity do quite often drastically change into something that rather resemble a dense activity.

Single neurons

Our modern understanding of the nervous system essentially begins with the *neuron doctrine* of Cajal, earning him the Nobel Prize in 1906. He used his exquisitely detailed stainings to state that the neuron is the fundamental physiological unit of the nervous system. From this doctrine, our knowledge of the electrophysiological nature of the neuron developed with the seminal Hodgkin-Huxley model of the giant squid axon, presented in a series of publications during 1952 that later lead to another Nobel Prize in 1963 awarded to Hodgkin and Huxley together with Eccles. The model, with various additions and variants, still serve as the golden standard description of how the neuron membrane potential is regulated through the intricate behavior of ion-channels that perforate the membrane of the neuron. It is worth citing Hodgkin and Huxley (1952) describing one of their figures: "Only one, […], is complete; in the other two the calculation was not carried beyond the middle of the falling phase because of the labor involved …" to illustrate their impressive feat in a time when the abundance of computing power we have today did not exist.

Modeling the neuron

Neurons communicate with action potentials or electric pulses that propagate along the neuron's axon to synaptic terminals and synapses made onto other neurons. By modeling neurons, we try to describe this process at a particular level of detail, which suits our current needs. Due to the fact that the signals between neurons are transmitted as discrete events, a model does not always need to consider the intricacies of the neuron membrane or the complex morphological structure of the neuron. In fact, complex models of neurons potentially suffer from the same bias-variance dilemma as was described in the introduction. Hence, if the purpose of the modeling is anything other than to actually describe all the intricacies in detail, it should not be a purpose of its own to always use as complex models as possible. A reduced complexity does of course also offer substantial advantages in the form of computational cost when the model is actually simulated. The reduced model might also benefit from the possibility to find analytical solutions to the questions that are investigated, essentially avoiding the simulation step completely.

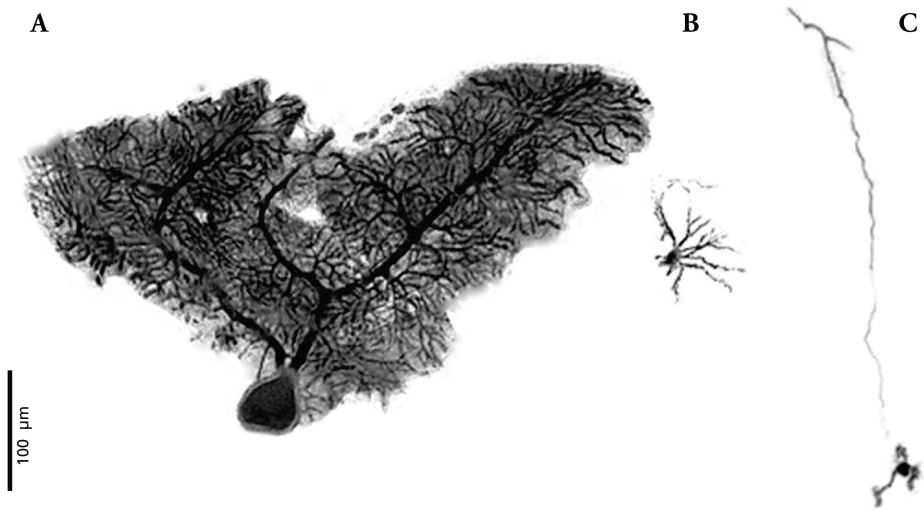

Figure 4. Morphology of cerebellar neurons
Three distinct morphological types of cerebellar neurons. **(A)** A Purkinje cell with its large soma and even larger dendritic tree, being innervated by hundreds of thousands of parallel fibers. **(B)** A stellate molecular layer interneuron, with its star shaped dendritic tree. **(C)** A granule cell with its characteristic axon that transforms into a parallel fiber as it reaches the molecular layer of the cerebellum (at the top). Closer to the soma of the granule cell (the bottom right), at least three dendrites can be seen reaching out to separate mossy fiber rosettes.

It is however not entirely straight forward to reduce the complexity of a detailed model. Consider the neurons in Figure 4. Their morphology is widely different, and it has been shown the same holds true when it comes to at least certain aspects of their electrophysiological properties (McKay et al., 2005, Molineux et al., 2005, Cesana et al., 2006, Zhong et al., 2010). The obvious drawback of using reduced models is that it is necessary to make sure that the reduction does not remove features from the model, which are essential to the function of the neuron, or at least the properties that are being investigated.

A simple and perhaps crude method of measuring model complexity is by the number of parameters of the model. Consider for example the elegant differential equation from Hodgkin and Huxley (1952) that describes the membrane potential of the neuron:

$$C_m \frac{dV_m}{dt} = g_L(E_L - V_m) + g_{Na} m^3 h(E_{Na} - V_m) + g_K n^4 (E_K - V_m).$$

The complete model also contains the following differential equations that describe how the time varying states, m, n, and h change:

$$\begin{aligned} dn/dt &= \alpha_n(1-n) - \beta_n n, \\ dm/dt &= \alpha_m(1-m) - \beta_m m, \\ dh/dt &= \alpha_h(1-h) - \beta_h h. \end{aligned}$$

As well as the equations that express the rates at which the three states change:

$$\begin{aligned} \alpha_n &= 0.01(V_m + 10) \Big/ \left(\exp \frac{V_m + 10}{10} - 1 \right), \\ \beta_n &= 0.125 \exp\left(\frac{V_m}{80}\right), \\ \alpha_m &= 0.1(V + 25) \Big/ \left(\exp \frac{V_m + 25}{10} - 1 \right), \\ \beta_m &= 4 \exp\left(\frac{V_m}{18}\right), \\ \alpha_h &= 0.07 \exp\left(\frac{V_m}{20}\right), \\ \beta_h &= 1 \Big/ \left(\exp \frac{V_m + 30}{10} + 1 \right). \end{aligned}$$

In total, the original Hodgkin-Huxley model has 27 parameters and 4 time varying states. For someone not used to reading equations, it is perhaps southing to learn that a mathematician would react with the same feeling of fright to the equations above as any other person would. The above equations form a complex system of non-linear differential equations. They are non-linear since the membrane potential will determine the rate at which the time varying parameters m, n, and h vary, as can be seen in the α and β equations. The non-linearity introduces difficulties in determining whether or not the membrane potential will remain stable or not, and in determining the properties of the possibly many equilibriums of all the time varying states. It is for example not obvious how to choose values for the parameters of the model in order to avoid the sodium channels being permanently open due to that the threshold of the sodium inactivation is set too high, or never open at all due to an exaggerated potassium conductance.

The fact that neurons are as stable and regular in their electrophysiological properties as they are, even though the underlying dynamics of the membrane potential are described by the complex set of equations above, is fascinating. Since the most violent of the dynamics will only be used to generate the action potential when the sodium channels open in a cascade as a response to depolarization, it is possible to reduce the complexity and construct a simplified model of the neuron without the non-

linearities and a reduced number of parameters. The action potential can be seen as a stereotyped event (see Figure 5E), which allows the model to focus on the subthreshold membrane behavior. In fact, such a simplified model was described by Lapicque already in 1907, and is now known as the leaky integrate-and-fire (LIF) model (Burkitt, 2006). Instead of describing the complex dynamics of the action potential explicitly within the model, the neuron is considered to produce an action potential whenever the membrane potential reaches a certain threshold, after which it is reset to a specified reset membrane potential. The LIF model can be expressed by the following single linear differential equation, the fixed spiking threshold, V_{th}, and the reset potential, V_{reset} :

$$C_m \frac{dV_m}{dt} = g_L(E_L - V_m),$$
$$\text{fire when} \quad V_m > V_{th},$$
$$\text{then reset to} \quad V_m = V_{reset}.$$

The LIF model offers an alternative with only 5 intrinsic parameters, compared to the 27 parameter non-linear model of Hodgkin-Huxley that was described above. The reduction does however come at the expense of the time-dependent and voltage-sensitive conductances and the dynamic refractoriness that real neurons exhibit (Abbott and Kepler, 1990). The choice of model introduces a balance between the biophysical accuracy and the practical and conceptual simplicity of using the model. Furthermore, both the Hodgkin-Huxley and the integrate-and-fire model have given their names to two whole families of models, with various extensions that include additional biological features of neurons to the models (Burkitt, 2006).

Neural noise

The spike trains of most neurons in the nervous system are characterized by irregular intervals between subsequent spikes (see Figures 5A–D). Since there is little definite knowledge of how information is actually encoded within the spike trains across the nervous system, the long standing question to what extent the irregularities are due to intrinsic noise of neurons, or external synaptic noise, remain open (Mainen and Sejnowski, 1995, Chow and White, 1996, van Vreeswijk and Sompolinsky, 1996, Naundorf et al., 2006). For example, irregular but chaotic rather than stochastic spike trains can arise from balanced states emerging in large networks that are sparsely connected, but have strong synaptic weights (van Vreeswijk and Sompolinsky, 1996).

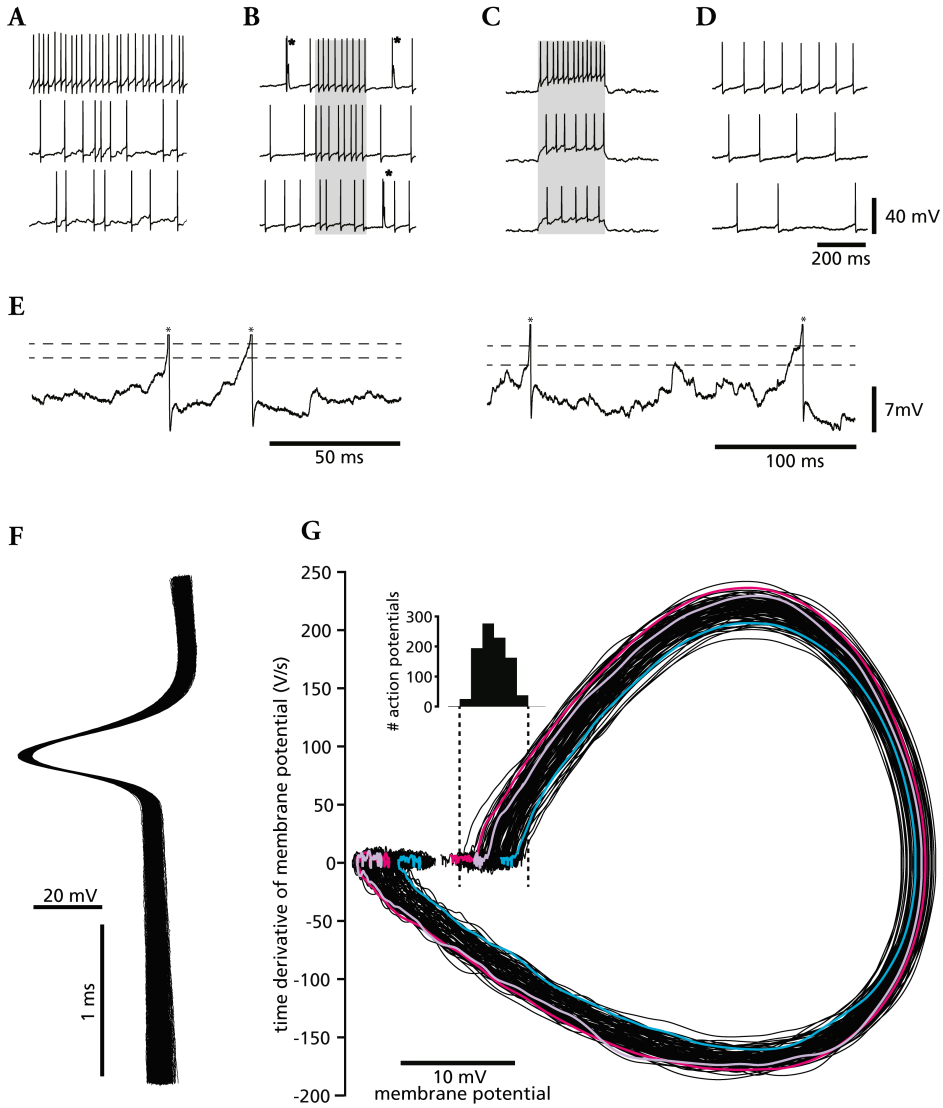

Figure 5. Intracellular spontaneous activity *in vivo*

(A-D) Intracellular spontaneous activity of neurons that are modulated with a bias current across the membrane. The gray areas indicate when the modulating bias current is applied, while each row correspond to a separate amplitude of the applied bias current. **(A)** Spinal interneuron. **(B)** Cerebellar Purkinje cell. Occasional complex spikes are indicated with asterisks. **(C)** Cerebellar molecular layer interneuron. **(D)** Cerebellar Golgi cell. Note the irregular firing of all neurons. **(E)** Two magnified traces from the spinal interneuron in **(A)** where the firing threshold seems to be stochastic, since it significantly change between even subsequent action potentials within 50-200 ms. **(F)** 900 overlaid action-potential traces from the same spinal interneuron as in **(E)**. **(G)** Phase plot of 100 out of the 900 action potentials in **(F)**, with three highlighted action potential traces in color. Note the variation of the firing threshold shown in the inset histogram. The firing thresholds were found by tracing the phase plot of an initiated action potential (dV/dt > 20) back to where it did rise above dV/dt > 2.5. The histogram inset contains all 900 original action potentials.

Figure 5E shows the membrane potential of a spontaneously active spinal interneuron recorded from a decerebrate cat preparation *in vivo*. The neuron in the figure seems to initiate action potential at different thresholds levels, which vary even between consecutive spikes. This behavior is in contrast to the properties of rat cortical neurons in Mainen and Sejnowski (1995), where they find neurons with a nearly flat threshold in a rat cortical slice preparation. It has been proposed that the stochastic firing threshold is due to cooperating sodium channels (Naundorf et al., 2006). If the properties of these channels vary considerably between the *in vivo* and at least some *in vitro* settings, they could perhaps explain the difference in stochastic spike initiation. The onsets or thresholds have been found to vary with up to 10 mV in neurons recorded *in vivo* in the cat visual cortex (Naundorf et al., 2006). The same range of onset potentials can be found in the phase plot of the action potential traces recorded from the spinal interneuron in Figure 5G.

While measuring neurons in slice preparations *in vitro* has the advantage that the synaptic input to the neuron can be under complete control, it also impose a dramatic change to the neuron's environment that might influence the behavior of the neuron. The presence of active synapses does for example to some extent introduce high-frequency fluctuations to the membrane potential that might influence the properties of the spike generation mechanism and the responsiveness of the neuron (Hô and Destexhe, 2000, Destexhe et al., 2001). In the *in vivo* setting, it is not clear to what extent anesthetics influence single neurons, and awake animals will have variations of the global brain activity that change over time outside of the experimenter's control. The decerebrate preparation that were used to collect the data throughout this thesis avoids these issues since the animal is non-anesthetized, the *in vivo*-like synaptic noise is maintained at a steady level, and changes due to global thalamocortical activity over time are reduced due to the decerebration.

Modeling noisy neurons

Irregular spike trains can be analyzed using the intervals between consecutive spikes called the inter-spike intervals (ISIs). If the spike trains were completely regular as in so called pacemaker cells, all intervals would be equally long, while irregular spike trains lead to distributions of ISIs. The properties of those distributions can be used to investigate and describe the point process that generated the spike trains.

It is possible to model noise, either intrinsic or synaptic, through suitable extensions of deterministic models (Burkitt, 2006). LIF models extended with noise has an advantage over more complex conductance based models, since the LIF models offer analytical solutions that allows it to be fitted to experimental ISI distributions (Rauch et al., 2003, Burkitt, 2006, La Camera et al., 2008). Such biophysical models offer

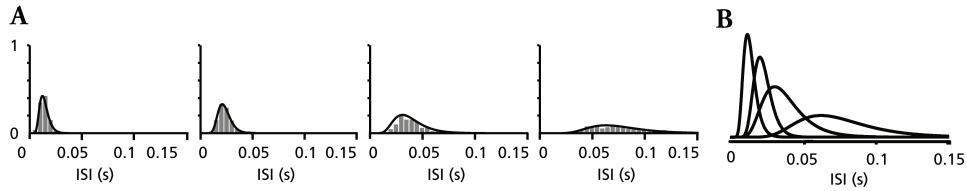

Figure 6. The ISI distributions at different stationary states of firing can be approximated by lognormal distributions.
(A) The empirical ISI distributions of a cerebellar Golgi cell are shown as histograms at four different levels of excitation. The log-normal distributions with the closest fit to the empirical distributions are shown as black lines on top of the histograms. **(B)** All four of the log-normal distributions in **(A)** shown together to illustrate how the distributions change with varying levels of excitation.

direct interpretations to the parameter values as biophysical entities, such as the membrane resistance and capacitance. Those interpretations are however only valid if the effective values of the biophysical parameters of the model correspond or correlate to the actual values of the neurons. In contrast, Rauch et al. (2003) and La Camera et al. (2008) found that the parameter values that corresponded to the best fit ISI distributions of a neuron were not correlated to the directly measured values of the membrane resistance and capacitance of the neuron. It seems that even the relatively simple LIF models suffer either from overfitting, or that the subthreshold voltage gated conductances and the dynamic refractory processes of real neurons influence the firing statistics of the neurons in ways the LIF model cannot emulate (Abbott and Kepler, 1990).

In Paper III, we instead directly investigate the firing statistics of neurons that are spontaneously active. By approaching the firing statistics without considering whether the noise is intrinsic or synaptic by nature, the complexity of the model can be reduced to a minimum. The resulting model will however by design be completely phenomenological with no biophysical interpretations of the model parameters.

By recording the neurons intracellularly, it is possible to inject a current across the membrane, which will either excite or inhibit the neuron relative to its resting membrane potential in order to investigate how the properties of the firing statistics change with various levels of excitation. Figures 5A–D show how four neurons of different types respond to bias currents by changing their average firing rate. The statistics of the different spike trains were then investigated using the distributions of their ISIs. As long as the point process initiating the spikes is renewal (i.e. no correlation between subsequent ISIs) and stationary, the distribution of ISIs will completely describe the statistics of the spike firing of the neuron. Figure 6A shows the empirical ISI distributions of a cerebellar Golgi cell recording at four different

levels of excitation. In order to reduce the parameters describing the distributions, they can be approximated using a parametric distribution, such as the log-normal distribution (see Figure 6B). A log-normal distribution is created by taking the logarithm of a normal distribution. The logarithmic properties of the resulting distribution makes it positive definite and gives it a skewed shape that resembles the properties of empirical ISI distributions, whose skewed shape and positive definite properties arise partly due to the refractory period of the neuron. The log-normal distribution is commonly parametrized by two parameters σ and μ, which are the mean and standard deviation of the underlying normal distribution. Using that parametrization, the probability density function of the log-normal distribution is defined by the following equation:

$$f(t;\sigma,\mu) = \frac{1}{t\sigma\sqrt{2\pi}} \exp\left(\frac{-(\ln\, t-\mu)^2}{2\sigma^2}\right).$$

The mean and standard deviation of the underlying normal distribution can be translated into the mean and standard deviation of the log-normal distribution using the following equations:

$$\begin{aligned} E &= \exp\left(\mu + \frac{1}{2}\sigma^2\right), \\ s.d. &= E\sqrt{\exp\,\sigma^2 - 1}, \end{aligned}$$

where E is the mean and $s.d.$ is the standard deviation of the distribution. By analyzing several regions of spike firing where the neuron exhibit different levels of excitation, it is possible to try to describe the relationship between the firing rate (E^{-1}) and the inverse of the standard deviation ($s.d.^{-1}$) of the ISI distribution as the level of excitation changes. The inverse of the standard deviation is used in order to have the same unit, Hz, on both measures. In Paper III, we present such an approximate relationship. It can be summarized by the following three equations:

$$\begin{aligned} E^{-1} &= c_x\, \ln[1+\exp(x-\Delta_x)], \\ s.d.^{-1} &= \exp\, x, \\ x &= c_I I - \Delta_I, \end{aligned}$$

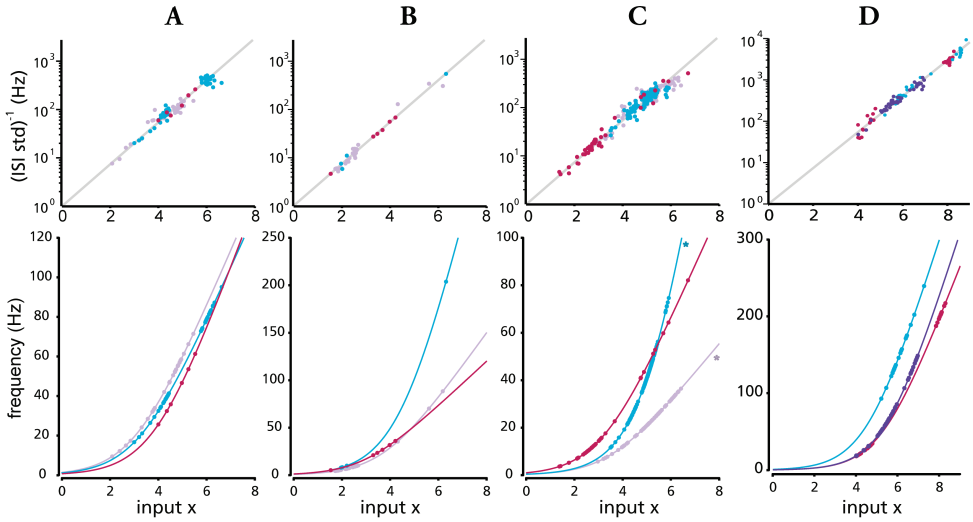

Figure 7. In vivo data from four distinct types of neurons fitted to the model
Model fit of **(A)** 3 Purkinje cells, **(B)** 3 Molecular layer interneurons, **(C)** 3 Golgi cells and **(D)** 3 spinal interneurons. The top row compares the predicted relationship between the internal parameter x, and the inverse standard deviation of the ISI distribution. Each point in the figure corresponds to a single region of spontaneous activity containing 50 action potentials. The neuron exhibits a good model fit if the data align with the line $(ISI\ std)^{-1} = \exp x$. The bottom row shows the relationship between x and the firing frequency of the neuron. The two Golgi cells in **(C)** that failed the statistical test ($p < 0.05$) of the residuals compared to the line in the top row are marked with asterisks in the bottom figure.

where I is the bias current input to the neuron. c_I and Δ_I are two parameters that relate the bias to the dimensionless state x that determine the level of excitation of the model. Finally, c_x and Δ_x are the parameters that regulate the relationship between $s.d.^{-1}$ and E^{-1}.

The lack of any parameters within the exponential relationship between $s.d.^{-1}$ and x allows it to be used as a method to validate the fit of the relationship to experimental data. With a logarithmic scale, the data should always align along the same straight line regardless of the parameter values of the fitted model. In Figure 7 the fit of the model to experimental data is illustrated using 12 different neurons of 4 types. Of all the 18 tested neurons in Paper III, only two cerebellar Golgi cells failed the statistical test of whether the data aligned along the straight line used as validation (Figure 7C).

Considering the differences in morphology of the neurons (see Figure 4 on page 12) and their differing electrophysiological properties (McKay et al., 2005, Molineux et al., 2005, Cesana et al., 2006, Zhong et al., 2010), it is surprising that the firing statistics (i.e. the relationship between $s.d.^{-1}$ and E^{-1}) of all the neurons could be fairly well described by the same rudimentary two-parameter equation. It has

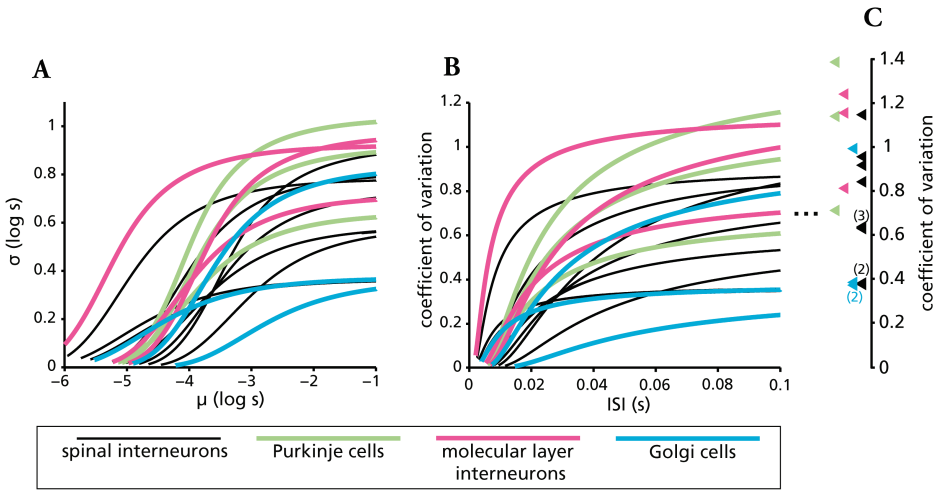

Figure 8. Comparing the firing statistics of different neurons
Comparing the predicted firing statistics of models fitted to the empirical ISI distributions of spinal interneurons (n=9), Purkinje cells (n=3), molecular layer interneurons (n=3) and Golgi cells (n=3). There is no obvious difference between the neuron types when the firing statistics is investigated across the range of activity. **(A)** The relationship between the mean and standard deviation of the underlying normal distributions that define the log-normal distributions. **(B)** The relationship between ISI and the coefficient of variation. **(C)** The asymptotic coefficient of variation as the ISI grows to infinity. The numbers in parentheses indicate if there is more than one value that overlaps in the figure.

previously been shown that it is possible to classify the type of neuron according to their spontaneous firing rate at rest together with the entropy of the ISI distribution (Van Dijck et al., 2013). Since the entire operative range of the neurons is given by the phenomenological model, it allows the same comparison to be made across the operative range, instead of just at their natural resting state. The first two moments of the firing statistics of all the neuron models across their firing range are shown in Figure 8. The variation between the neurons of the same type is comparable to the variation between the different types. The results indicate that it would not be possible to distinguish the neurons using the first two moments of their ISI distributions alone if they were not at their natural resting firing rate.

The use of stationary firing levels does in principle limit the models applicability to situations with homogeneous firing or a static input to the model. It is however possible to create an escape rate model using the hazard functions of the current ISI distribution using the following equation:

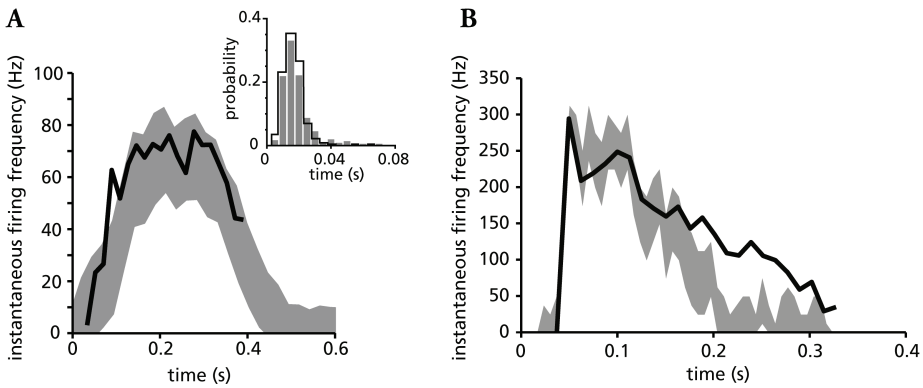

Figure 9. Predicted response by the inhomogeneous model during fictive locomotion.
The response of the inhomogeneous model compared to the actual instantaneous firing rate of a spinal neuron. The black lines show the recorded instantaneous firing rate (IFF)(Fedirchuk et al., 2013) and the gray areas indicate the 95% confidence bounds of the model. The models were driven by the recorded intracellular membrane potential of the same neuron that was also used to measure the extracellular spike trains used for the IFF (Fedirchuk et al., 2013). Just like the experimental IFF, the model output from 12 trials was binned into 30 bins. The simulation was repeated 250 times to compute the confidence intervals. **(A)** DCST neuron. The inset compares the ISIs of the experimental data (gray bars) to the ISIs of the model (black line). **(B)** VSCT neuron. Note that the model manages to reproduce the fast initial transient from 0-300 Hz, and the following slow decay of the IFF. It should be noted that in both **(A)** and **(B)** the input to the model and the measured firing frequencies where naturally not from the same step cycles, which means that perfect fit of model and the experimental data cannot be expected.

$$\lambda(t;\sigma,\mu) = \frac{f(t;\sigma,\mu)}{1-\Phi\left(\frac{\ln t-\mu}{\sigma}\right)},$$

where λ is the instantaneous firing rate, f is the probability density function and Φ is the cumulative density function. As long as the input is slowly modulated, the hazard rate could be used to handle also inhomogeneous situations. In contrast to slow modulation, transient synaptic inputs could cause significant non-renewal behavior, such as spike rate adaptation, that the model cannot handle. In Figure 9, the response of the model during fictive locomotion is compared to both a DSCT and a VSCT spinal neuron that is adopted from spinal neurons recorded during fictive locomotion *in vivo* (Fedirchuk et al., 2013). The model managed to reproduce the behavior both of DSCT neuron with a slowly modulated firing rate (up to ~80 Hz), and that of the VSCT neuron, which has much higher firing rate (up to ~300 Hz) and a far more transient behavior.

Simple in general, but complex in principle

It is fascinating and somewhat surprising that the empirical ISI distributions of such a large set of different neurons can be approximated by the log-normal distribution. Note that the log-normal distribution is not special in this sense. Similar results can be achieved using gamma distributions and inverse Gaussian distribution (Paper III includes a more thorough examination), both of which are commonly used to describe ISI distributions (Kostal and Lánský, 2007). The fascination does rather come from the fact that neurons, at least at a glance, behave similarly despite rather different morphology, anatomical location and assumed function within the nervous system. The fact that the two parameter model presented in Paper III can describe the firing statistics of the neurons across their operative range is fueling the same fascination.

It seems as if the nervous system does its best to not overuse the complexity that is inherent in the biophysical nature of the neuron. The basic response to stimuli is rather a close to linear translation from stimulus strength to firing rate in many of the investigated neurons, across most of their operative range. In another view, if we as observers do not know the details of what the neuron is signaling, the average response of the neuron to relatively slowly modulated inputs can be seen as the first order approximation of the information content. This could perhaps give a developmental advantage since the neuron on the receiving end of the axon can directly make some crude sense of the information it receives without learning the intricacies of the encoding.

Cerebellum

Compared to other regions of the nervous system, the cerebellum has an almost crystalline and uniform structure. The same salient pattern of neurons is repeated over and over across the cerebellar cortex. Furthermore, compared to most of the thalamocortical system and circuitry of the spinal cord, the cerebellum is distinctly anisotropic. The only input to the cerebellum is relayed through two distinct pathways – the mossy fibers and climbing fibers, and the only output projections from the cerebellum originate from the deep cerebellar nuclei (DCN).

Despite being smaller in volume than the neocortex, the cerebellum houses the majority of all neurons in the nervous system in the form of cerebellar granule cells. It also seems to play a ubiquitous role in the nervous system, since it both receives input and relays its output from/to almost all other brain regions. It is more or less involved in as diverse tasks as the timing of eye blink conditioning, motor control of both relatively low-dimensional task such as the vestibulo-ocular reflex (VOR) as well as complex motor coordination (Ito, 2006), and higher cortical functions (Ito, 2008). The regular structure of the cerebellar circuitry, and its involvement in many diverse brain functions, has commonly been attributed to a general functionality, where the cerebellum is thought to learn to correlate input to and correct errors that the body and brain encounters. As a heritage of the early ideas of Marr (1969) and Albus (1971), the input via the climbing fibers is assumed to serve as an error signal to the Purkinje cells. Based on this error signal the cerebellum will learn to avoid repeating the same error by correlating the error signal to input it received from the parallel fibers being active when the error occurred. It would allow the cerebellar circuit to preemptively regulate the activity of downstream neurons the next time the circuitry encounters the same combination of inputs.

The combination of the feed-forward structure of the cerebellum, and the separate input pathway of the climbing fibers serving as an error signal, resemble that of a single hidden layer artificial neural network with an external error or teaching signal. While the similarity is striking, it also poses a risk that all experimental findings are interpreted in this context without considering alternatives. There are in fact several experimental studies that indicate that the Marr-Albus view of the cerebellar circuitry might not be enough to explain all the details of how the cerebellum is used by the nervous system.

The cerebellar circuitry that is involved in both eye-blink conditioning and the VOR is particularly well studied, including the pre-cerebellar circuitry that is involved and the pathways that lead out of the cerebellum from the vestibular nucleus and the anterior interpositus nucleus (Ito, 2006). The plasticity processes of both these systems, which allow the cerebellum to learn how to respond, do not harmonize with the plasticity rule of the Marr-Albus type of models, where the difference between potentiation and depression of the parallel fiber to Purkinje cell synapse completely depend on the signaling of the climbing fibers. In the case of eye-blink conditioning, recent evidence suggest that rather than synaptic plasticity at the parallel fiber to Purkinje cell synapse, the Purkinje cell itself has an intrinsic mechanism that allows it to learn the interval between a parallel fiber activation and a subsequent climbing fiber activation (Johansson et al., 2014). In contrast, the change in synaptic efficacy of the parallel fiber to Purkinje cell synapse during VOR training is strongly correlated to the climbing fiber activity, but only during VOR-increase training (Kimpo et al., 2014). During VOR-decrease training there is no such correlation, even though over time the synaptic efficacy still changes.

The placement of the cerebellum within the neural circuitry that control voluntary movements allows it to take two different roles in order to facilitate the motor control, by learning either the dynamics or the inverse dynamics of the controlled limb (Kawato et al., 1987, Ito, 2006). The dynamics in the forward model would allow the cerebellum to predict the outcome of a motor command, presenting the motor cortex with an artificial feedback signal without the delay of the actual sensory feedback. In contrast, the inverse dynamics in the inverse model could be used to transform the motor command of the motor cortex into suitable torques or muscle activations. The cerebellum could thus be used both in a feed-forward controller as an inverse model, or in a feedback controller by providing the motor cortex with the predicted consequences of a motor command.

It is interesting to note that the Marr-Albus model of cerebellar function potentially fails to explain various different features within VOR and eye-blink conditioning. It is also questionable whether the slow firing rate of the climbing fibers can contain enough information to be a reliable source of an error signal when it comes to more complex voluntary motor control, involving for example coordination. The local spillover plasticity between parallel fiber synapses that has been described by Wang et al. (2000) could for example mean that the plasticity of individual synapses is governed also by the population activity of the parallel fibers, and not only the climbing fiber activity. The differences between the circuits involved in the VOR, eye-blink conditioning, and voluntary movement control, indicate that the three tasks seem to make use of the cerebellar circuitry in distinctly different ways. Also from the complexity perspective, it is perhaps necessary to shift the view of the cerebellum from that of a general-purpose machinery with unlimited capability to learn how to react to any combination of input patterns. A general-purpose structure would assume that

most if not all of the signal processing, such as feature extraction, takes place within the cerebellar circuitry. In the Marr-Albus model, the feature extraction would take place within the granule layer through a massive expansion recoding of the incoming mossy fibers. In their view, each granule cell is assumed to receive a unique (or random) set of approximately 4 mossy fibers, and the granule cell will only be active whenever a majority or all of the mossy fibers are activate simultaneously. However, if the cerebellum is not considered to be a separate structure that has to manage itself, but rather a component in the larger circuitry, the properties and "knowledge" already in the larger circuit could be used to guide the cerebellar circuit. The purpose would be to reduce the possible permutations of mossy fiber to granule cell connections towards combinations of inputs that are functionally relevant to the cerebellar circuit. This would offer a substantial advantage since random combinations of mossy fiber inputs would almost always lead to combinations that has no functional significance to the system due to the astronomical number of combinations that are possible following from the high dimensionality of the input. Even the vast number of granule cells would be far from enough to naively encode the complete input space in any detail. Paper I use this perspective in order to analyze the role of the cerebellum in the circuitry that enable the nervous system to control the movements and the coordination of multi-segmented limbs.

Key experimental findings in the literature

There are at least three classes of experimental evidence that support the view of the pre-cerebellar circuit as a vital part of the cerebellar circuit:

1. There is convergence of different inputs to pre-cerebellar neurons
2. The inhibition of granule cells from Golgi cells is mostly tonic
3. Many granule cells seem to receive functionally similar inputs from more than one, and at least in some cases all four mossy fibers that innervate the granule cell.

The experimental data corresponding to the first point show that the pre-cerebellar circuitry at least has the means to find functionally relevant combinations of different inputs that the cerebellum can use. A large proportion of the mossy fibers that reach the cerebellar regions that contribute to motor control originate from the spinal cord (Oscarsson, 1973). Many of these spinal systems receive both sensory feedback and descending motor commands, either directly or mediated via other spinal interneurons (Jankowska et al., 2010, Hammar et al., 2011, Jankowska et al., 2011a, Jankowska et al., 2011b, Krutki et al., 2011, Shrestha et al., 2012a, Shrestha et al., 2012b). These observations offer an alternative view of where the multi-modal

combination of inputs occurs. Rather than taking place within the granule layer due to random convergence of mossy fibers, it can take place already in the spinal circuitry. The selection of functionally relevant combinations could potentially even be a main function of the neurons that project to the cerebellum. It is not clear how these combinations would be selected, but there are experimental results that indicate the presence of more complex features being available already at the spinal level (Bosco et al., 1996, Poppele et al., 2002). It is also known that the spinal circuitry is established during development and that the underlying plasticity processes take the sensorimotor apparatus into account (Holmberg et al., 1997, Petersson et al., 2003).

The second point considers the properties of the feedback inhibition from Golgi cells to granule cells. The granule cell layer, including the feedback inhibition of the Golgi cells is conspicuously similar to the network structure proposed by Földiak (1990). However, in order for the Golgi cells to actively influence the encoding of the granule cell population, the feedback inhibition should act on the same timescale as the mossy fiber excitation of the granule cells. In contrast, in the adult cerebellum the inhibition is to a large extent carried by a slowly modulated tonic inhibition (Wall and Usowicz, 1997, Jörntell and Ekerot, 2006). Even in juvenile animals up to 98% of the charge is carried by the tonic component (Duguid et al., 2012). Fast inhibitory post-synaptic potentials on the other hand are weak or absent (Jörntell and Ekerot, 2006). Phasic inhibition is still present in young animals, but it is gradually lost at the same time as the tonic component develops (Brickley et al., 1996, Wall and Usowicz, 1997). In line with these findings, *in vivo* studies of the inhibitory response show a presence of tonic inhibition (Jörntell and Ekerot, 2006, Bengtsson et al., 2013), but fast inhibitory post-synaptic potentials are difficult to detect (Chadderton et al., 2004). The lack of fast phasic inhibition in the juvenile and adult animal questions the encoding capabilities of the granule layer during fast movements, which is when feed-forward coordination of the movement by the cerebellar circuit is most likely critical.

The third point considers the focal termination of mossy fibers upon the granule layer of the cerebellum (Alisky and Tolbert, 1997, Garwicz et al., 1998, Tolbert and Knight, 2003, Gebre et al., 2012). The termination pattern of the incoming mossy fiber pathways forms discontinuous sharp clusters, rather than diffuse continuous patterns. These studies are at odds the proposal that granule cells as a rule sample functionally dissimilar mossy fibers. Hence, the topology of the mossy fiber innervation patterns questions the expansion recoding that was proposed by Marr and Albus, where the role of the mossy fiber to granule cell divergence was to decorrelate the granule cell population activity through sampling of randomly combined mossy fibers. Furthermore, individual granule cells receiving cutaneous input have also been reported receiving functionally equivalent input from more than one and sometimes all of the incoming mossy fibers (Jörntell and Ekerot, 2006, Bengtsson and Jörntell, 2009). The third point rather strongly indicates that if there is any substantial combination of different inputs, it often has to take place prior to the granule layer.

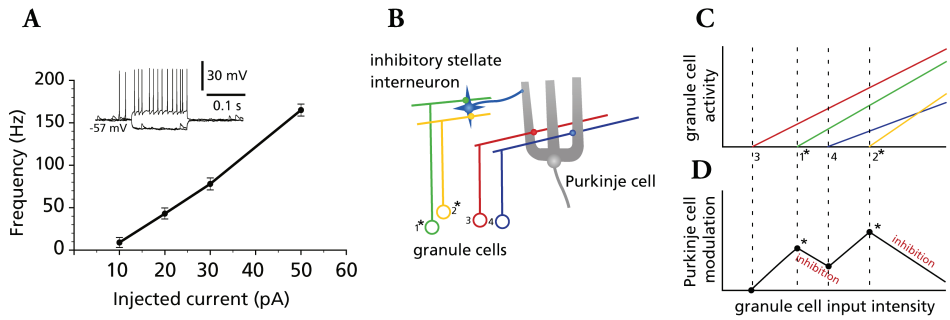

Figure 10. Potential use of the linear rectified response of granule cells
(A) Granule cell recorded *in vivo* which exhibits a linear relationship between firing frequency and injected bias current. The inset shows the response of the granule cell to injected currents of +/- 20 pA. **(B)** Simplified view of the cerebellar circuitry between granule cells and Purkinje cells. The granule cells form parallel fibers that innervate the Purkinje cell either directly with an excitatory synapse (3,4), or through an inhibitory molecular layer interneuron (1,2). **(C)** The thresholds of granule cells vary both due to intrinsic mechanisms and tonic inhibition from Golgi cells. In the example all four granule cells are innervated by the input to the cerebellum, but have their firing thresholds at different levels of input intensity. **(D)** The combined modulation of the Purkinje cell can be shaped to approximate an arbitrary non-linear function with a piece-wise linear function. The thresholds of the granule cells correspond to the knots of the piece-wise linear approximation.

The focal termination is not a strict rule however, since there are separate mossy fiber systems that converge on single granule cells, at least in fringe zones of the separate mossy fiber areas (Huang et al., 2013). It is however not clear if and to what extent they mediate functionally dissimilar inputs.

Redundant granule cells in piecewise linear approximations

If the massive population of granule cells is not used for the commonly proposed expansion recoding, what use does the brain have of them? Some redundancy might be necessary for noise reduction, allowing the Purkinje cell to average across several granule cells transmitting the same information, cancelling out noise. The granule cells do however differ considerably with regard to their resting membrane potential (Chadderton et al., 2004), and most likely also the amount of tonic inhibition from the Golgi cells. As a consequence, each granule cell will reach its firing threshold at different levels of input excitation. Consequently, due to their non-linear firing threshold, each granule cell will still transmit separate information to the Purkinje cell, even compared to other granule cells receiving the exact same inputs. In other words, the granule cells would still perform an expansion recoding of the inputs, and

it does not require that each granule cell is innervated by a unique combination of mossy fibers. Viewing the cerebellum as an adaptive filter (Dean et al., 2010), the non-linear thresholds of the granule cells would allow the granule cell population to act as a bank of filters that transform the incoming mossy fiber information before it reaches the Purkinje cells.

Compared to the other cerebellar neuron types (see Figure 7 in the previous chapter), granule cells have a fairly sharp firing threshold (see Figure 10A for an example), followed by an almost completely linear response to additional input excitation. This response is similar to that of linear rectified units, which have recently been found to have beneficial properties when used in large artificial neural networks (Glorot et al., 2011). In addition, the bidirectional plasticity and the complementary location of the receptive fields in the PCs and interneurons indicate that the interneurons at least approximately act as inhibitory relays of the granule cell activity to the Purkinje cells (Jörntell and Ekerot, 2002, 2003, Ito, 2006, Dean et al., 2010). The inhibitory relay allows the granule cell to both excite and inhibit a Purkinje cell, as illustrated in Figure 10B. This allows the Purkinje cell to combine the activity of several granule cells with varied thresholds into piecewise linear approximations of non-linear functions (see Figures 10C–D).

In order to understand how the population activity of the granule cells can be useful during coordination, it is necessary to also consider the multidimensional case. Instead of approximating a non-linear one-dimensional curve as in Figure 10D, combining two or more projections allow the Purkinje cell to approximate non-linear surfaces as in Figure 11B. Due to the restriction that several granule cells will be aligned along the input dimension defined by the combination of inputs within the incoming mossy fiber, not all surfaces can be approximated given any such combination of inputs. Furthermore, the direction of the projections is not oblivious to the actual non-linear surface that should be approximated. Figure 11C displays the ideal projections across three non-linear surfaces that were approximated by a population of 60 model granule cells. The surfaces were chosen since they appear in the inverse dynamics of double-joint limbs (Hollerbach and Flash, 1982, Kawato et al., 1987), and must somehow be represented by the cerebellum in order for it to perform efficient feed-forward control (Kawato et al., 1987). Figure 11D compares the accuracy of the approximations, given 1-5 mossy fiber projections. Note that the accuracy of the approximations can be fairly good already with two projections, as long as they are not randomly selected. Also note that the accuracy differs significantly between the three surfaces, even though they are superficially very similar. It illustrates the need to consider the very details of the functionality in order to evaluate a model or describe the detailed function of a neural circuitry.

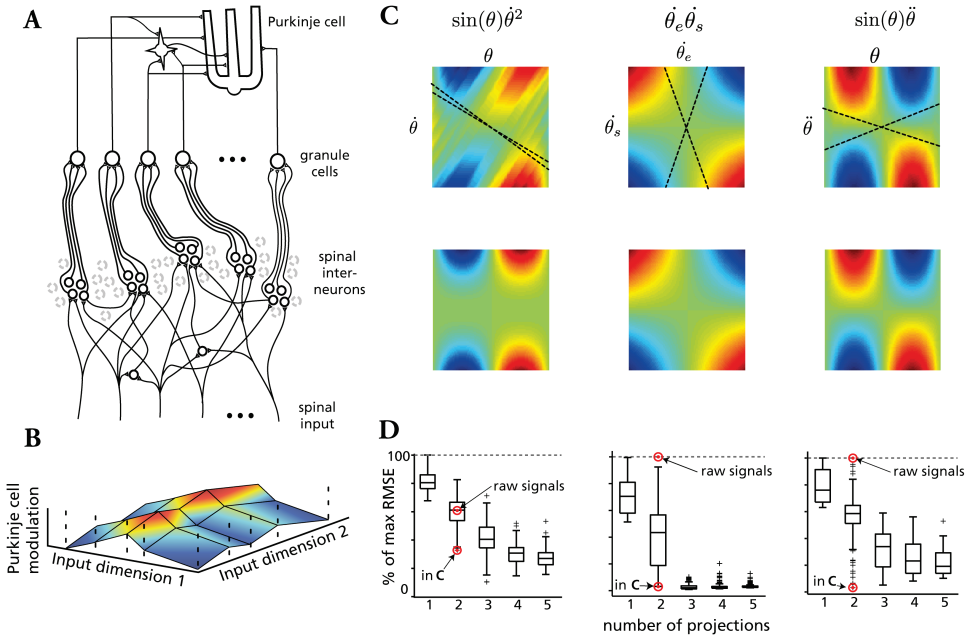

Figure 11. Multidimensional piece-wise linear approximations from projections in input space
(A) The cerebellar input layer together with precerebellar spinal neurons that project to the cerebellum through mossy fibers. Note that signals of different origin can combine already at the spinal level to form projections through the input space. Granule cells receive approximately 4 mossy fibers that often share the same modality, and there are several granule cells per mossy fiber. This means several granule cells might align along each projection from the spinal cord (compare to Figure 10 D). By innervating the same Purkinje cell with granule cells that align along more than one projection, the Purkinje cell can learn multidimensional approximations **(B)**. **(C)** Example of approximations of the 2-dimensional surfaces $\sin(\theta)\dot{\theta}^2$, $\dot{\theta}_e\dot{\theta}_s$, and $sin(\theta)\ddot{\theta}$ (see Paper I for details). The top row illustrates the best approximation using two projections and 60 granule cells along each projection. The dashed lines show the direction of the two optimal projections. **(D)** Percentage of the root mean squared error (RMSE) compared to the maximal error using 1 to 5 projections and a total of 60 granule cells divided among the projections. The distribution of the error is due to the random sampling of projection directions. The RMSE of the optimal directions in **(C)** are indicated by *in C* and the RMSE of using signals without spinal convergence is indicated by *raw signals*. The boxes indicate the mean, and the 25%-75% percentiles, and the whiskers extend to the most extreme data that is not considered to be outliers.

Moving down the hierarchy

In order to make efficient use of the cerebellar circuitry, the pre-cerebellar circuits could maneuver the cerebellum, using it as an extension to for example coordinate the abilities of the pre-cerebellar circuitry. The most basic method would be to send selective projections within the input space through the mossy fibers. The projections should correspond to the functionality that the cerebellum is supposed to learn in the context of that pre-cerebellar circuit. Actively selecting good projections reduces the number of mossy fibers that are necessary, and consequently enables the cerebellar circuitry to put most of the granule cells to good use instead of being ignored after receiving functionally irrelevant combination of mossy fiber inputs.

The mossy fiber systems

Considering the cerebellum as general-purpose computational machine, the spinal part of the spinocerebellar circuitry could just directly relay afferent and efferent information to the cerebellum, assuming the cerebellar part of the circuitry would be able to learn all the intricacies necessary for coordination by itself. The spinal circuitry does however already contain circuitry that performs some level of motor control. From the perspective of the spinal cord it would make more sense to view the cerebellum as an extension of its functionality, to which the spinal circuitry would offer inputs that have functional significance. Such an organizational principle would offer a tremendous benefit to the system from a complexity standpoint, since the cerebellum would not have to re-learn what the spinal circuitry (and potentially also other pre-cerebellar circuits) have already figured out.

The spinocerebellar and spino-reticulo-cerebellar systems are major sources of input to the regions of the cerebellar cortex that have direct connections to the motor systems, the corticospinal, rubrospinal, reticulospinal, tectospinal and vestibulospinal tracts. The input originates from spinal neurons that ascend the spinocerebellar tracts (SCTs). These consist of the ventral spinocerebellar tract (VSCT) neurons including spinal border cells (SBCs), the dorsal spinocerebellar tract (DSCT) neurons, the rostral spinocerebellar tract (RSCT) neurons as well as the spino-reticulo-cerebellar tract (SRCT) neurons. All of these SCT/SRCT neurons receive sensory afferent signals, either directly or via other spinal interneuron circuits. They also receive direct and indirect descending motor commands. The neurons of SCT/SRCT that project to cerebellum are also involved in the local motor control circuit within the spinal cord, as they sometimes project directly to alpha motorneurons and the motor nuclei of the spinal cord (Alstermark et al., 2007). The end result is that the information that the cerebellum receives from these systems is neither distinct sensory feedback, nor efferent motor commands, but a combination that is most likely functionally relevant, since it is in some cases even directly involved in the motor control of the local spinal circuit.

The spinocerebellar pathways can further be subdivided into one direct and one indirect path (Jiang et al., 2015). The indirect path is relayed through the lateral reticular nucleus (LRN). Four separate ascending spinal and brainstem systems converge at the LRN neurons (Alstermark and Ekerot, 2015). The extra pre-cerebellar convergence might be used to allow the cerebellar regions that receive the LRN input

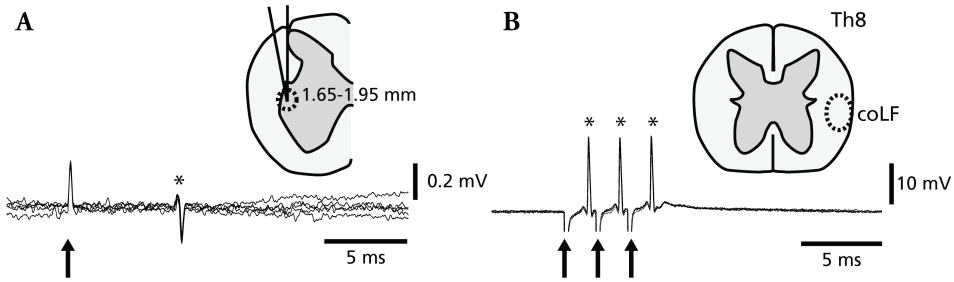

Figure 12. Identifying spinal border cells.
(A) An example of an extracellular recording of a spinal border cell with an antidromic response (indicated by asterisks in the figure) to stimulation in sublobule C1 in the cerebellum. The arrow indicates the stimulation artifact. The inset show a cross-section of the L4 segment indicating two alternative orientations of the electrode tracks and the depths at which the spinal border cells where found. **(B)** An example of a spinal border cell with an antidromic response from the contralateral lateral funiculus (coLF). The coLF was stimulated with a triple pulse, indicated by the three arrows. The target location of the stimulation electrode is indicated with a dashed line in the inset showing a cross-section of the Th8 segment.

to learn to coordinate these functions, including posture (bilateral ventral flexor reflex tract), reaching (C3-C4 propriospinal systems), and grasping (ipsilateral forelimb tract).

Spinal border cells

The SBC compartment of the VSCT offers a unique possibility to study the influence a single spinocerebellar tract in isolation has upon the neurons of the cerebellar cortex. This is possible since the SBC neurons are the only SCT neurons which ascend the contralateral lateral funiculus (coLF) (see Figure 12B), and have terminations within sublobulus C1 of the paramedian lobule in the posterior cerebellum (Matsushita et al., 1979, Matsushita and Ikeda, 1980, Matsushita and Yaginuma, 1989). When the coLF is stimulated while recording the response of cerebellar neurons within sublobulus C1, at least the early recorded responses can be expected to be caused by the SBC mediated inputs alone.

The possibility of activating the ascending SBC fibers through coLF stimulation was verified in Paper II, by first identifying spinal neurons that had an antidromic response to stimulation of the cerebellar cortex (Figure 12A). The region with the lowest threshold for successful antidromic stimulation of the spinal neurons was always found to be within the granule cell layer of the medial part of sublobulus C1,

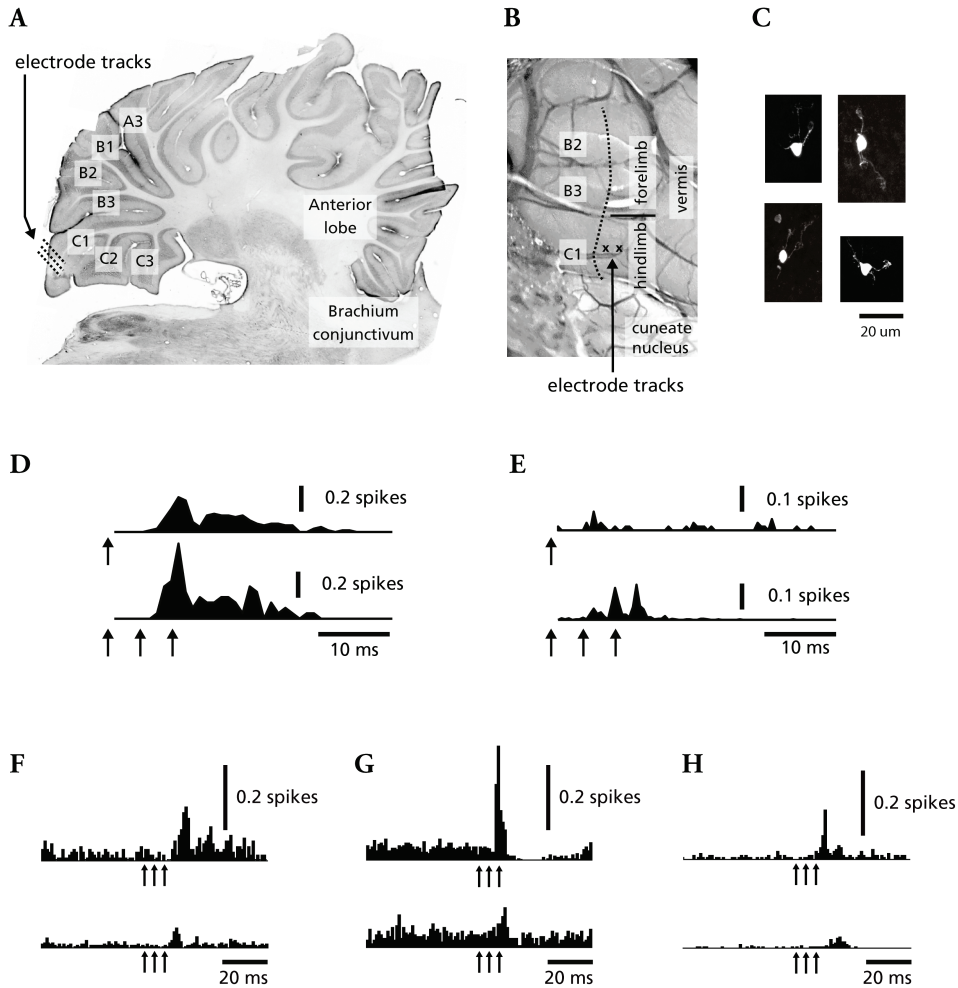

Figure 13. Cerebellar responses to SBC stimulation
(A) Sagittal section of the cerebellum. Microelectrodes and patch clamp recording electrodes were inserted along the dashed lines, at an angle of 30-45° relative to the surface of sublobule C1-C3, B1-B3 indicate the sublobules in the paramedian lobule. **(B)** Surface view of the posterior part of the paramedian lobule. Sublobulus C1, B3 and B2 from **(A)** is indicated to the left. All recordings were made within the C1 zone (different from sublobulus C1) indicated by the arrow. The horizontal line indicates the limit of the forelimb representation of the C1 zone. **(C)** Four morphologically identified granule cells recovered from sublobule C1. **(D)** Peristimulus histograms showing the typical response of a granule cell to single pulse (top) and triple pulse stimuli (bottom) to the coLF. The arrows indicate the time of each stimulation pulse. **(E)** Another typical granule cell response that manages to capture the timing of the three stimulations during the triple-pulse protocol in the peristimulus histogram. **(F-H)** The response evoked by a triple-pulse stimulus to coLF from two Golgi cells in **(F)**, 2 Purkinje cells in **(G)**, and 2 molecular layer interneurons in **(H)**. All histograms have 1ms binwidth.

as expected if the neurons were indeed SBCs. The viability of using coLF stimulation to evoke responses within the cerebellar cortex was further verified by also measuring the antidromic response to coLF stimulation (Figure 12B). The response latency variation was found to be negligible indicating that it was indeed antidromic and not synaptic activation. Using coLF stimulation should therefore resemble synchronous activation of almost the complete SBC population.

The synchronous activation of an entire population of neurons within a single tract is of course an extremely artificial stimuli, far from the patterns of activity that could be expected during any type of natural behavior. However, as long as the results are treated with caution, they could provide an alternative to purely anatomical studies, where the functional connectivity (e.g. silent synapses) can also be included. Interpreted together with other findings it is also possible to deduce the influence the single spinal system can have on the upstream neurons within the cerebellar circuitry.

When the granule cells located within sublobulus C1 (see Figures 13A–C) were recorded during the coLF stimulation, only a relatively small subset (~18%, N=264) had a measurable response. The granule cells that did respond could be classified broadly into two classes. The first had a relatively unspecific response with variable spike response times (see Figure 13D), while the second class had one or two spikes at regular response latencies, giving rise to sharp peaks in the peristimulus histograms (see Figure 13E). The relatively strong response of many of the granule cells indicate that more than one of the mossy fibers that innervated the granule cell had a SBC as its origin. Previous studies have additionally found that synchronous activity in three or four of the mossy fibers that innervate a granule cell were required to get a reliable response (Jörntell and Ekerot, 2006, Bengtsson and Jörntell, 2009), suggesting that a majority or even all of the mossy fibers innervating the granule cells that responded strongly to coLF stimulation were SBCs.

Despite the relatively small fraction of the recorded granule cells that had a coLF response, more than 60% of the recorded Purkinje cells (n=10/16) were modulated by the stimulation, even though some of them only had a small modulation. Figure 13G illustrate the difference with peristimulus histograms from two Purkinje cells with detectable modulation – one which was strongly modulated by the coLF stimulation and the other with only a small modulation. Also the fraction of Golgi cells (41%, n=11/27) and molecular layer interneurons (45%, n=5/11) that exhibited responses to the stimulation were considerably larger than that of the granule cells. The fact that the population of granule cells that receive SBC input could elicit a response in a small majority of the Purkinje cells, indicate that the SBC neurons provide relatively important input to the C1 zone of sublobulus C1. The strong response of some Purkinje cells further indicate that a single component of the SCT input to the cerebellum can have a large influence upon the cerebellar output mediated via the Purkinje cells.

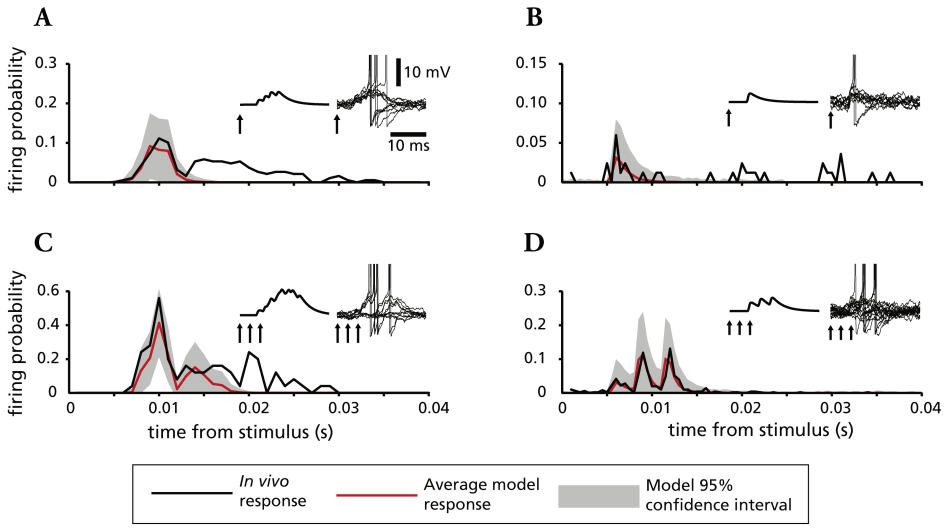

Figure 14. Simulated granule cell responses to coLF stimulation.
Simulated peristimulus response to single (top) and triple (bottom) pulse stimulation to the coLF of two granule cell models - **(A,C)** vs. **(B,D)** - compared to the actual response of two granule cells *in vivo*. The granule cells were modeled as exponential integrate-and-fire models, with four incoming mossy fibers from coLF with varying conduction delays. The only difference between the two models is the neuron resting potential and the conduction delays between coLF and the cerebellum of the four mossy fibers. **(A,C)** The diffuse but strong response of the neuron to the left could be modeled using highly diverse conduction delays. See the simulated synaptic input and granule cell model response in the insets. The left inset shows the synaptic input and the right the simulated membrane potential of the granule cell model. **(B,D)** The sharp response that manages to retain the tripe pulse shape **(D)** could in contrast be modeled using synchronized conduction delays. Note that the late component of the response *in vivo* most prominent in **(A)** is most likely due to recurrent spinal connections that were also activated during the stimulation of the coLF *in vivo*. The model has no such connections, which means it is not expected to reproduce the late component of the response.

Modeling the spinocerebellar system

In order to investigate the behavior of the cerebellar input layer during more natural conditions than the synchronous coLF stimulation, a model of the SBC spinocerebellar circuit was developed in Paper IV, including a LIF granule cell model. Similar to other components of the VSCT, it is likely that the SBCs participate in the control of locomotion (Arshavsky et al., 1972, Fedirchuk et al., 2013). Hence, the activity of the model granule cells was investigated during afferent and efferent input to the model SBCs that have been recorded during fictive locomotion.

The granule cells were simulated as LIF neuron models with an additional exponential part that mimic the spike generation of the neuron (Fourcaud-Trocmé et al., 2003). The model also included a stochastic component adding noise to the simulated membrane potential. The model parameters were selected to reproduce the behavior of the recorded granule cells in Figures 13D−E when exposed to single and triple pulse stimuli. Each pulse of the stimulus was delivered as one post-synaptic current to the LIF model from each of the four mossy fibers that were assumed to innervate the model. Two sets of model parameters allowed the model to reproduce both the sharp response of the neuron in Figure 13D and the indistinct response of the neuron in Figure 13E. Figure 14 illustrate the behavior of the two models during the two stimulation patterns, compared to the experimental *in vivo* recordings.

Despite the distinct difference between the responses of the two granule cells, the two models could capture the difference with using slightly different resting potential (-58 mV compared to -57 mV) and by varying the synchronization between the four incoming mossy fibers. It is plausible that this is in fact the difference between the recorded neurons as well, since the conduction delays of the SBCs were found to vary considerably, comparable to those of other parts of the VSCT (Geborek et al., 2013). The influence of the synchronization can be seen in the synaptic input to the neuron models, shown in the insets of the panels in Figure 14.

The spinal border cells were modeled using the inhomogeneous point process described in Paper III, using the model that reproduced the behavior of a VSCT neuron in Figure 9B on page 21. The reasons for not using the same type of model for both SBCs and granule cells were twofold. Firstly, VCST neurons (and thereby the SBCs) have been found to operate as rate coders during fictive locomotion (Fedirchuk et al., 2013), whereas the distance to the firing threshold (Chadderton et al., 2004, Jörntell and Ekerot, 2006) as well as the sharp non-linear onset of the f-I curve (see Figure 10A on page 27) allow the granule cells to respond to fast transient events that the model in Paper III cannot handle with certainty. Secondly, the experimental data describing the behavior of spinal neurons comes as slowly modulated membrane potentials or spike trains from massive input populations with small unitary synaptic inputs that suits the escape rate model, while the single pulse and triple pulse stimulation suits a model with explicit spike generation, which can handle the large unitary synaptic inputs that are normally delivered from mossy fibers to granule cells (Cathala et al., 2003, Chadderton et al., 2004, Jörntell and Ekerot, 2006).

SBCs receive efferent motor command from a large population of neuron descending the reticulospinal tract (RST). They also receive Ib afferent input mediated via inhibitory interneurons, as well as excitatory input from group II and possibly group Ia afferents (Jankowska et al., 2011b, Shrestha et al., 2012a, Shrestha et al., 2012b). About half the population does however not receive any excitatory group I input,

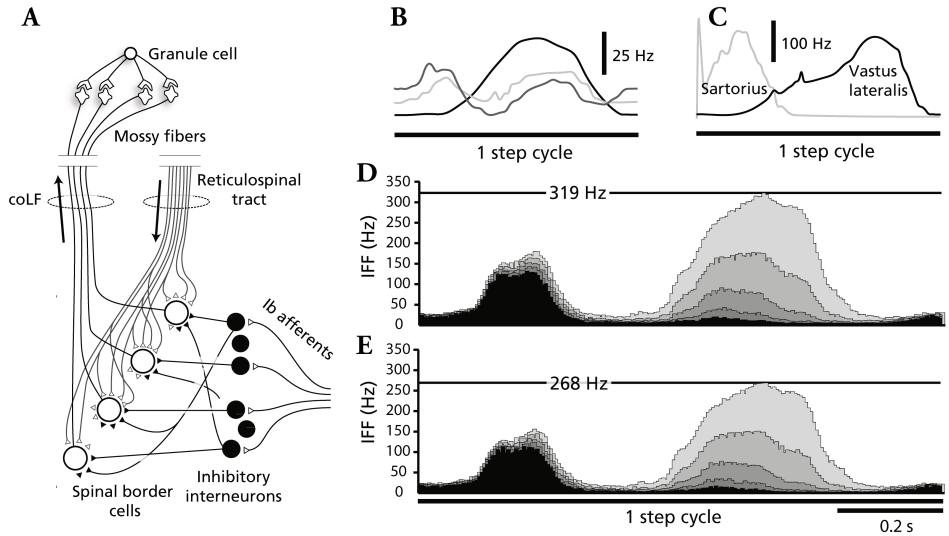

Figure 15. Simulating cerebellar granule cells with spinal border cell input during locomotion. **(A)** The network structure of the model that was used to simulate the granule-cell response to spinal border cell input. The granule cells were modeled as in Figure 12, with four mossy fibers relaying spike trains from four spinal border cells that were modeled using escape rate models. The activity of the spinal border cell models were modulated by simulated input from the reticulospinal tract **(B)** and various degree of Ib inhibition from the Sartorius and Vastus lateralis muscles **(C)**. **(B)** The activity of three different resticulospinal tract neurons during a single step cycle adopted from Matsuyama and Drew (2000a, b). **(C)** The constructed activity of Ib interneurons carrying afferent Ib sensor information from the Sartorius and Vastus lateralis muscles during a single step cycle. Constructed using data from Matsuyama and Drew (2000a) **(D)** The response histogram of the first granule cell model (see Figure 12 A and C). The gray histograms correspond to increasing Ib inhibition from the vastus lateralis muscle, going from none (light gray) to maximal (black). **(E)** The response histogram of the second granule cell model (see Figures 12B,12D). Note that while the granule cell models had distinct responses to the single and triple-pulse coLF stimulation, they only vary with regard to their firing rate amplitude when exposed to slowly modulated mossy fiber inputs. IFF, instantaneous firing frequency determined by the frequency of spikes within each bin of the histograms.

even though the inhibitory Ib input remains from combinations of synergistic muscles (Burke et al., 1971, Lundberg and Weight, 1971, Oscarsson, 1973). This subpopulation of the SBCs was used as a prototype to model the input that was provided to the SBC models. The Ib convergence from synergistic muscles was also limited to that of the vastus lateralis/quadriceps and the sartorius muscles. Convergence of Ib inhibition of these muscles has been shown for the subset of SBC neurons that lack group I excitation (Burke et al., 1971).

The structure of the simulated circuitry, from the RST and Ib input that converge on the SBC to the granule cells in the cerebellar input layer, is illustrated in Figure 15A.

Due to the massive number of synaptic inputs to each SBC, both excitatory via the RST (approximately 5000) and a comparable amount of inhibitory synapses from spinal interneurons (Shrestha et al., 2012a, Shrestha et al., 2012b), each synaptic event most likely have a very small amplitude. Given that the maximum depolarization reaches 5 mV and the maximum hyperpolarization -7 mV (Hammar et al., 2011, Shrestha et al., 2012a, Shrestha et al., 2012b), the synaptic events will have an average amplitude in the order of μV, which is comparable to the synaptic strengths found in other synapses of the cat spinal cord (Jankowska, 1992). The sum of the population activity will however give rise to large modulation of the membrane potential of the SBCs. Such modulations under fictive locomotion have been shown for VSCT neurons in general, where the maximal depolarization during a step cycle was 7.1 mV (Fedirchuk et al., 2013).

The total RST and Ib inhibitory input the SBC model could therefore be modeled as a single synaptic input current that combine the average input from all the excitatory and inhibitory synaptic inputs. The time courses of the RST input was constructed using the activity of three RST neurons adapted from Matsuyama and Drew (2000a, b). Their activity during a step cycle of fictive locomotion can be seen in Figure 15B. The simulated Ib inhibitory input was adapted from EMG recordings from the specific muscles from Matsuyama and Drew (2000a). The time course of the simulated Ib activity is shown in Figure 15C. The EMG data could be used instead of neural activity, since it has been shown that the Ib activity essentially mirrors the EMG activity of a muscle (Prochazka and Gorassini, 1998).

The instantaneous firing frequency of the two granule cell models during one step cycle can be seen in Figures 15D–E. Note that the responses of both the granule cells share the same topology despite their different transient responses, while their maximum firing rate differ slightly. In order to investigate the influence from varying amount of Ib inhibition of the SBCs, the simulation was repeated with amplitudes of the inhibition corresponding to 0%, 20%, 40%, 60%, 80% and 100% of the maximal inhibition. Such a change of the relative contribution from the efferent motor commands from the RST and the inhibitory Ib sensory feedback correspond to changing the direction of the projections that were described in the previous chapter.

The simple fact that the efferent commands and afferent feedback converge already at the spinal level, allow the cerebellar circuitry to learn non-linear interactions between the afferent and efferent signals. Through the indirect pathway via LRN, these non-linear interactions could also be between higher level aspects of motor control, such as reaching, grasping and keeping posture (Alstermark and Wessberg, 1985). The predicted activity of the granule cell model can also be compared to the activity of granule cells of awake mice under locomotion. As is the case in Figure 15, the granule cells seem to have a dense activation profile that is modulated according to the step cycle (Powell et al., 2015), possible due to the feedback from individual muscles.

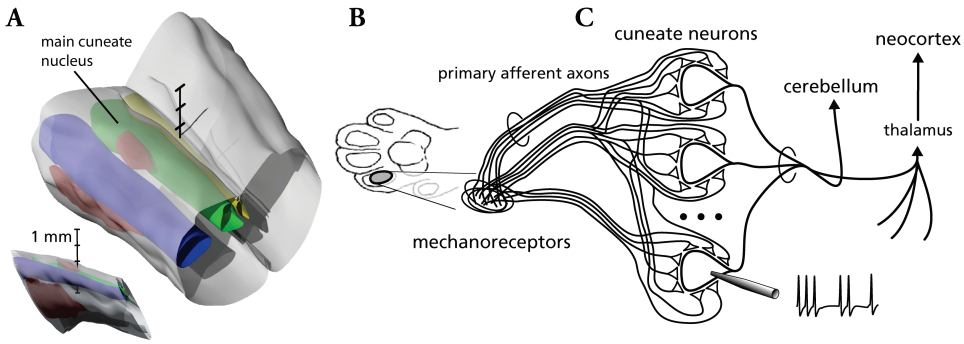

Figure 16. Anatomy and connectivity of the cuneate nucleus.
(A) 3D reconstruction of the caudal brainstem of a cat, indicating the anatomical location of the cuneate nucleus in relation to other structures and nuclei. The large green volume indicate the main cuneate nucleus, while the smaller red volume just lateral of the cuneate nucleus indicate the external cuneate nucleus, which is a separate structure. The gracile nucleus (yellow), the trigeminal nucleus (blue) and the lateral reticular nucleus (large red) are also indicated as colored volumes. Adopted from Geborek et al. (2012). **(B)** The stimulation site upon one of the digits of the paw. The stimulated skin regions will contain the receptive fields of 100s of primary afferent sensors that sample the relevant mechanoreceptor population. **(C)** Each afferent make approximately 1000 synapses within the cuneate nucleus. The cuneate neurons then project to the neocortex via thalamus and to the cerebellum as mossy fibers. It is possible to make both intracellular and extracellular recordings of the cuneate neurons using patch clamp electrodes.

The cuneate nucleus

The cuneate nucleus lies anatomically close to the LRN, which can be seen in the 3D reconstruction of the cat brainstem in Figure 16A. Like the LRN, it also projects to the cerebellum, but unlike LRN, the cuneate nucleus only processes input from tactile primary afferents. The nucleus is in principle structured as a feed-forward network of inhibitory interneurons and projection neurons. The afferents that reach the cuneate nucleus innervate both the projection neurons and the inhibitory interneurons. The inhibitory interneurons do in turn innervate the projection neurons, which send axons both to the cerebellum and the thalamus. In contrast to studying the fundamental processing of spinal circuits which was done with the SBC system in Paper II and Paper IV, where both efferent motor commands and antagonist Ib inhibition were involved, the feed-forward network structure, and the fact that the cuneate nucleus receives only tactile primary afferent inputs allows an experimental setup where complex stimuli that resemble what would be encountered during natural behavior can be used. As will be shown further down, it is possible to reliably

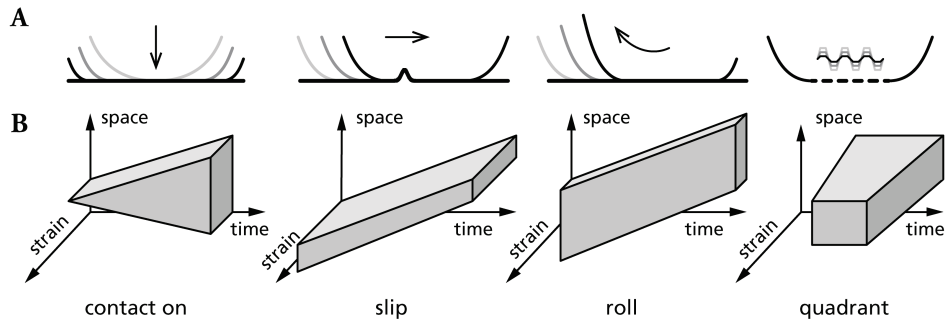

Figure 17. Spatiotemporal mechanical stimulation patterns
Drawings illustrating the stimulus patterns that were delivered using the tactile display. *Contact on* illustrates the initiation of a contact, characterized by an area of strained tissue that increase over time. *Slip*, correspond to a narrow area of strained skin that moves across the fingertip, while *roll* correspond to a broader area with lower strain. The *quadrant* stimulus was used as a control stimulus, and activates a small region of skin (1mm^2) with fast stretch (5 ms), a subsequent delay (5 ms), followed by a fast relaxation of the stretch (5 ms). The quadrant stimulus could be delivered at one location within a 3x3 grid covering the tactile display **(A)** The surface displacement of the fingertip/digit during the actual stimulus. **(B)** The corresponding time evolution of the surface strain along one single somatotopic coordinate.

mechanically stimulate the receptive field of cuneate neurons, and record a highly reproducible response. Furthermore, studying the cuneate nucleus serves a twofold purpose, since it relays the tactile information not only to the cerebellum, but also to the neocortex via thalamus. The details of the information processing of the cuneate nucleus could therefore not only provide information about the functionality of the cerebellum, but also the properties of the somatosensory areas of the neocortex.

Classically, the cuneate nucleus can be described as overlooked, at most credited to perform some type of contrast enhancement of the primary afferent information via lateral inhibition from the local inhibitory interneurons (Kandel and Schwarz, 2013). The assumed role of the cuneate nucleus follows directly from the classical view of primary sensory cortex (S1). There is supposed to be a close to one-to-one mapping from primary afferent sensors and a column of S1, organized as a somatotopic map (Zachariah et al., 2001, Kandel and Schwarz, 2013). The one-to-one mapping does of course not allow that any relay between the sensors and the cortical area performs any complex processing of the primary afferent data. Instead, such processing is not thought to take place until the secondary somatosensory areas (Kandel and Schwarz, 2013). While there are experimental evidence for this view of the early somatosensory processing, most of it has been produced in a reduced experimental setting, which fails to reproduce the complex dynamics of both the skin and the contact deformations that arise during natural use of the skin area. The purpose of most

studies so far, has been to characterize the receptive fields and tactile submodalities using single punctate stimuli (Mountcastle, 1997, Friedman et al., 2004, Johansson and Flanagan, 2009). The results from such studies indicate a somatotopic map, but should perhaps be considered more of description of the average connectivity than the fundamental functional principle of S1.

Mechanical spatiotemporal stimulation

In a description of somatosensory processing it is necessary to include the contribution from contact mechanics during the skin-object interaction that define touch. Using this view, it is clear that any mechanical stimulation that lead to deformation of the skin will be non-local with respect to the receptive fields of individual receptors (Hayward, 2011). The processing of such non-local sensory data would likely integrate the information from several receptive fields and submodalities. The integrated information of the skin-object interaction could subsequently be used to extract important features of the touched objects, i.e. texture, shape, friction and similar properties. In Paper V, the response of cuneate projection neurons to fundamental components of such mechanical skin-object interactions are investigated using a tactile display that can deliver patterns of strain to the skin that resemble the strain patterns during natural moving skin-object interactions (Hayward et al., 2014). The stimulation patterns that were delivered using the tactile display are illustrated in Figure 17, and include contact initiation, contact cessation as well as sliding object interactions at various speeds. Together these stimuli excite what can be considered fundamental features of spatiotemporal skin-object interactions (Hayward, 2011, Hayward et al., 2014).

Initially the intracellular membrane potential of five neurons was recorded using whole-cell patch-clamp electrodes. Since the purpose was to initially examine the synaptic input to the neurons, action-potentials were avoided by applying a hyperpolarizing current over the membrane. Samples of the intracellular results are shown in Figures 18A–B. Note the silence of the primary afferents between stimulations, and the reliability of the synaptic input during stimulations. As is shown by example in Figure 18B, both the pattern of synaptic input to a single neuron during the different stimulation patterns, and the response of each neuron to a single stimulus were easily discernable (see Paper V for details).

Figure 18C illustrate the similarity between the intracellular response to a stimuli and the peristimulus histogram created from spike times extracted from extracellular recordings of the same neuron. The close resemblance motivated the use of extracellular recordings to obtain data from an additional 13 neurons (in total n=18), available due to the relative simplicity of maintaining an extracellular recording. This

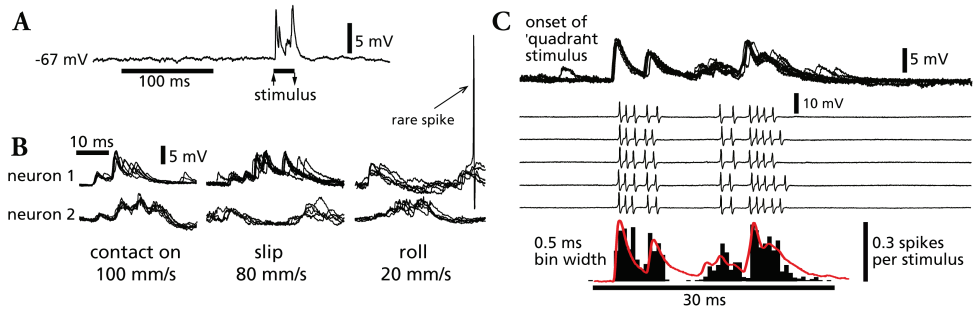

Figure 18. Response of cuneate neurons to spatiotemporal tactile stimulation
(A) Example of an intracellular recording of a cuneate neuron. The membrane potential is significantly modulated by excitatory synaptic input during stimulations to the receptive field of the neuron. **(B)** 5 overlaid traces of intracellular synaptic responses from two neurons responding to three of the applied stimulation patterns. Note the reliability of the synaptic responses and the quality of the recording indicated by the shape of the rare spike of neuron 1. **(C)** Overlaid intracellular response traces to a quadrant stimulus (top), and 5 example extracellular recordings of spike trains from the same neuron and stimulation pattern (middle). The average intracellular response (red line) closely resembles the peristimulus histogram of the extracellular spike trains (bottom).

allowed the number of repeated stimulations per stimulus and neuron to be increased from 7 to 50, since the same neuron could be recorded for a significantly longer period of time.

In order to investigate the information content of the extracellular spike trains, the spikes evoked during a response were counted and binned into non-overlapping spike bins. Binned spikes were used because it cannot be assumed that the exact timing of each spike is preserved and is available to cortical regions that process the cuneate input, since the signal is indirectly relayed to there through the thalamus. The average spike frequency during one of the bins is however more likely to reach the cortex unperturbed. As a consequence of the binning, the granularity of the possible information is reduced, but this should in principle only mean that it is theoretically possible to reach even better classification results using another better encoding.

The resulting features of the spike trains, i.e. the approximate spike counts during each bin, were evaluated using a classifier. It was given the task to correctly classify the responses of the stimuli using the responses from a population of cuneate neurons. The classification results are presented in Figure 19A, where it can be seen that when the full 150ms observation period was used, the correct classification rate approached 100%, even when using as few as 3 out of the 18 neurons. When all 18 neurons where used, the correct classification rate reached above 70% already after 30ms. In addition to the labeled classification, the heterogeneity of the responses to the different stimuli was investigated using unlabeled clustering of the responses (see Figure 19B). The fact

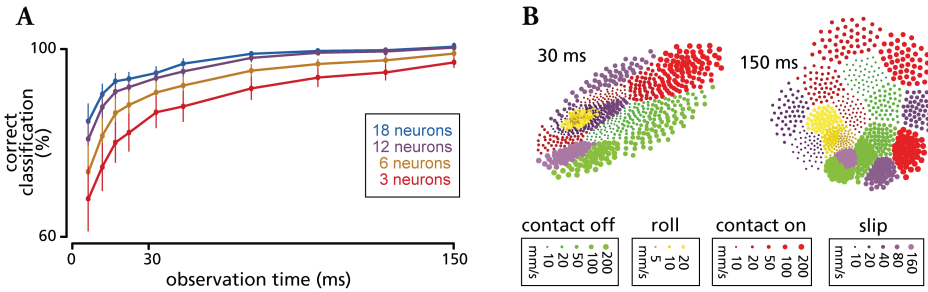

Figure 19. Population response and segregation capabilities
(A) Classification result using linear classification and binned spikes. The figure illustrates the relationship between the length of time used for analysis after stimulation onset and the correct classification rate. **(B)** Unlabeled clustering of the population response to all stimuli using curvilinear component analysis. Almost all the patterns could be discriminated already after 30 ms. At 150 ms the segregation was perfect, even for the different speeds of roll, which proved to be the least informative stimulus family.

that the unlabeled method managed to segregate the responses without knowledge of the underlying stimuli indicate that the population response of the cuneate nucleus is tuned to segregate the fundamental features of the skin-object interactions that were used to construct the different stimulus patterns.

A common principle of the pre-cerebellar circuits

The spinocerebellar system, including the LRN and cuneate nucleus in the brainstem were carelessly grouped under the same label in this chapter. It is obvious by considering the biomechanics they interact with, that the separate systems have different roles and detailed functionality. The VSCT and DSCT do for example seem to be interested in rhythmic movements, while the indirect relay through the LRN sample the state of execution of voluntary movements together with posture (Alstermark and Ekerot, 2013).

These systems have previously been overlooked, perhaps due to a seeming simplicity in their behavior that can be seen in Paper II and Paper IV for the SBC system and Paper III for the individual spinal interneurons. It is possible to only consider their rate code and still deduce the strength of a stimuli or the phase of a step cycle. Overall they also seem to relay a mapping from distal location within the biomechanics to somatotopic and motor maps in the motor cortex and the primary somatosensory cortex.

The complex response of the cuneate neurons to natural stimuli adds another layer to their story. While the rate code or topological map can be useful for the brain, it is likely that the individual neurons within the rate code also signal complex features of the biomechanical state, but without ways of delivering natural stimuli that excite these features, they remain hidden. Paper V illustrate how a close to natural stimulus, delivered as patterns of strains to the skin reveals this encoding in a circuitry that was previously though to perform little complex processing of the incoming sensory information.

Sparse coding

Searching for coding principles offers the tempting possibility to find general-purpose methods that the brain can utilize to encode and decode data within the nervous system. Investigating such coding principles also offers the possibility to study circuits before descriptions of the actual detailed content of the incoming signals exist, or before the detailed functionality of a circuit is known.

Sparse coding is one such encoding principle that is explored in Paper VI, and follows from the spike threshold that all neurons have. Since single neurons can be inhibited from generating action potentials, it is possible to have inactive neurons within a population of neurons that do not signal any information other than that they are inactive. The most extreme population code that utilizes inactive neurons is the *local code* (see Figure 20A, left). In a local code only a single or small subset of all neurons in a population is active at any time. Furthermore, these same neurons are only active during a single pattern of input or context. Since there is no interference between the active neurons during different contexts using a local code, it is possible to learn how to respond to a certain context encoded via a local code after a single presentation of that context. The consequence of this simplicity is that the population of neurons can at most encode a single context per neuron.

With no inactive neurons, the code of the neural population activity is called a *dense code* (see Figure 20A, middle). Using a dense code the population activity is allowed to use any combination of neural activity to encode a context. In general, this means that it is possible to encode an exponential number of contexts in relation to the number of neurons included in the population. At the same time, the interference between the representations of different contexts leads to an enormous increase in complexity, leading to a slow learning rate.

The region of population activity between the local and dense code is in principle a sparse code. The idea is to find the ideal trade-off between the benefits and drawbacks of the local and dense code. The idea of utilizing a sparse activity in a neural encoding to increase beneficial properties has a long history in the theoretical descriptions of both the cerebellar cortex (Marr, 1969, Albus, 1971) and the neocortex (Barlow, 1972). It is straight forward to understand why a sparse code provides benefits when the population activity needs to be decoded. With fewer active neurons in any given situation, it is not necessary to consider all the quiet neurons when the correct

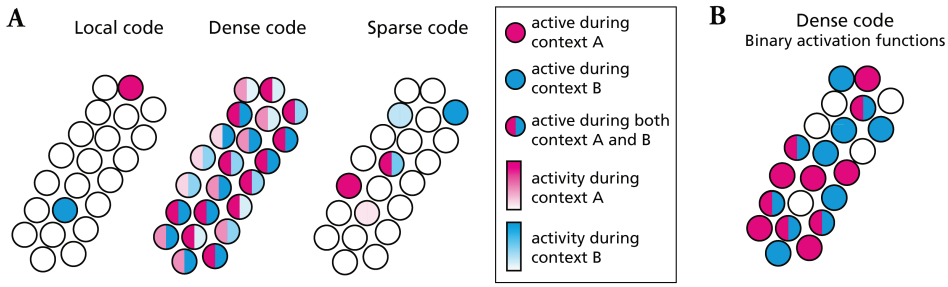

Figure 20. Sparse code
(A) A sparse code (right) can be considered to be any code with an activation ratio between that of a local code (left) and a dense code (middle). The sparse code to the right is an example with an activation ratio of approximately 10%. **(B)** Note the difference between a dense code with binary activation functions (neurons that are either on or off) and continuous thresholded activation functions. In the binary setting, a dense code will have an average activation ratio of 50%, while it will have a 100% activation ratio in the continuous setting.

decoding or response is learned. A sparse code can thus in principle reduce the complexity of the population activity, by reducing the number of combinations of neurons that are allowed to be active simultaneously.

Theoretical considerations

From a complexity perspective, a sparse code can efficiently reduce the possible complexity of the encoding since it essentially reduces the number of possible combinations of activity that are allowed within the population of neurons. At the same time, it is almost always necessary to have additional circuitry to construct the sparse representation of the incoming information. That circuitry could in principle introduce additional complexity to the circuitry. Hence, not only the level of sparseness, but also the method to enforce or create the sparse representation most likely matters.

In an artificial setting, where each neuron is active or inactive independently of each other, the sparseness of the encoding is directly influencing the maximum convergence rate of a decoder that is using a gradient descent method to learn to decode the population activity (Schweighofer et al., 2001). The convergence rate of a gradient descent method determines how far it is possible to move towards the correct configuration of synaptic weights from the current configuration, given one step of learning. In Paper VI, these results are expanded upon and it turns out that while a sparser code will indeed increase the maximum convergence rate, it is not the ratio of

active neurons that matters, but the number of active neurons. It means that in order to get the same beneficial properties with a large neural population as with a smaller one, the sparseness must increase in proportion to the population size.

Results obtained using models with only a few neurons compared to the actual circuitry in the real nervous system, might therefore not be applicable. One could consider to simply increase the sparseness to compensate for the increase in population size, but this might lead to unrealistically low activity ratios bordering the local code scheme. Concrete image classification studies have shown that the model accuracy of the constructed classifiers decrease significantly when the activation ratio gets less than 5-20% (Glorot et al., 2011, Thom and Palm, 2013). Interestingly, the variation in the threshold activity where the decrease occur depend strongly on the method that was used to enforce the sparseness of the code (Thom and Palm, 2013). A sparseness of 5-20% can also be compared to the 1% activation ratio that was assumed in the cerebellar model of Albus (1971).

Sparseness in the cerebellar cortex

Sparse coding has been proposed as one of the dominant encoding principles of the both the cerebellar input layer, and multiple cortical circuitries including primary visual cortex. In the framework of the Marr-Albus theory of cerebellar cortex, Albus hypothesized that the population of granule cells had binary activation functions and a 1% activity ratio (Albus, 1971). The low activity ratio was chosen to allow the granule cells to differentiate between similar mossy-fiber patterns, and Albus also noted that it should also facilitate learning speed due to less interference between different patterns. In contrast to Albus, Marr assumed that only the excitatory parallel fiber to Purkinje cell synapses carried information to the Purkinje cells. This led him to conclude that the granule cell activity had to be sparse in order to encode the necessary amount of patterns (Marr, 1969). Like Albus, Marr also assumed a rather sparse activity of the incoming mossy fibers in order to reach the necessary low activity ratio of the granule cells. Marr did however underestimate the ratio of silent synapses to be approximately 30%, which is much lower than the recent estimate of as much as 80%-98% (Ito, 2006, Dean et al., 2010). Increasing the fraction of silent synapses to 95% will lead to that the calculated necessary sparseness using the equations of Marr would drop to 5% in order to store a single pattern, or 0.25% in order to store as little as 20 separate patterns (see the separate box on the next page for details). While the overall functionality of the cerebellar circuitry most likely correspond to the role it has in the Marr-Albus models, the assumptions regarding whether the granule cell population activity is sparse or not should be used with caution.

Analysis of Marr (1969)

Marr used common combinatorics to investigate the proportion of active granule cells, and the subsequent influence that the level of activity should have upon the memory capacity of a cerebellar-like circuit. While the reasoning is completely sound, Marr made assumptions regarding the activity patterns of the mossy fibers and the fraction of silent parallel fiber to Purkinje cell synapses that has since been disproven. Using Marr's model to motivate a sparse activity of the granule cell layer should therefore be avoided.

The mossy fiber activity is sparse

Marr assumed that the number of active mossy fibers out of the 7000 whose activity reach a single Purkinje cell is sparse. "The essential point is that the numbers are all nearer 0 than 7000 (on an arithmetical scale)". Marr used the following equation to calculate the sparseness of the granule cells using the number of active mossy fibers:

$$\binom{C}{R}\binom{L}{R}\Big/\binom{7000}{R},$$

where R is the firing threshold of the granule cells, C the number of mossy fibers innervating a single granule cell, and L the number of active mossy fibers. Using R=3, C=4 and R =3500, the activity ratio of the granule cell population will reach 50%, far above the activity ratio that is required below.

95% rather than 30% silent synapses

Marr used the following equation to investigate the influence of silent synapses upon the number of patterns the Purkinje cell could learn:

$$\left(1-\frac{n}{200{,}000}\right)^{x} > p,$$

where 200,000 is the number of granule cells, n is the number of active granule cells, x is the number of patterns that are stored, and p is the fraction of silent synapses. The following table shows the number of patterns that can be stored at different levels of sparseness ($n' = n/200{,}000$):

n	n'	x (p=95%)	x(p=75%)	x (p=30%)
500	0.25%	20	64	481
1000	0.5%	10	32	240
2000	1%	5	16	120
5000	2.5%	2	6	48
10000	5%	1	3	23
20000	10%	0,5	1,5	11

Using a realistic fraction of silent synapses, Marr's model would require an unrealistic high sparseness in order to store more than a few patterns.

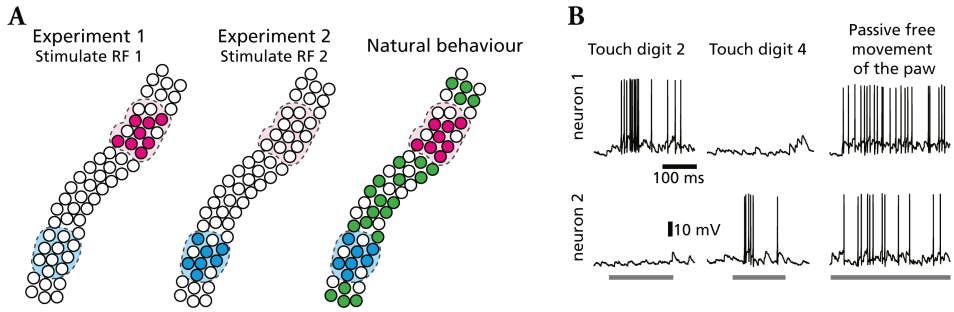

Figure 21. Experimental considerations.
Reduced stimuli can in some cases indicate a sparser code than a natural stimulation. **(A)** Stimulating a single receptive field with focal termination within the population of neurons will yield a sparse code if the receptive fields are stimulated individually. However, during natural behavior the receptive fields might never be stimulated in isolation. The code would thus be much less sparse during natural behavior than during the reduced experimental setting. **(B)** Example of cerebellar granule cells and tactile stimulation of individual receptive fields, compared to a natural passive movement of the entire paw. Note that the passive movement seems to engage receptors within receptive fields across the entire paw, and not only those closely related to the joint. The gray lines under the bottom row indicate the stimulus duration.

cerebellar-like circuitry that enforces optimal memory capacity, assuming there are no inhibitory interneurons that contribute with more than a modulatory blanket inhibition upon the Purkinje cells (Clopath and Brunel, 2013). The same circuitry would also exhibit the distribution of synaptic weights, including the large fraction of silent synapses that have been observed in the cerebellum (Brunel et al., 2004). The model does however also predict that the granule cell activity should be binary, which does not seem to be the case (see Figure 10A on page 27). Detailed granule cell models have also been found to reliably transmit modulated signals (Rössert et al., 2014). The properties of the molecular layer interneurons also indicate that they do transmit detailed parallel fiber information, rather than act through a modulatory blanket inhibition (Jörntell and Ekerot, 2002, 2003, Dean et al., 2010).

An alternative explanation of the large fraction of silent synapses is the co-variance learning rule (Sejnowski, 1977, Porrill and Dean, 2008). In this view, uncorrelated synaptic inputs will decay to zero. It is however possible that the cerebellum utilize the possibility of silent synapses as a method to reduce the complexity. Compared to the co-variance rule, which may still be the underlying rule used to update the informative neurons, the cerebellum would actively force a large fraction of the synaptic weights to be zero. This is a common method to keep the model complexity low of artificial neural networks (Hastie et al., 2009), sacrificing some bias to keep the variance error low. In fact, enforcing silent synapses is comparable in efficiency to the

use of sparse coding (Thom and Palm, 2013), and might be a method the cerebellum utilizes to retain the relatively dense activity of the granule cell population.

The experimental findings that contradicts many of the predictions made from the sparse coding hypotheses of granule layer of the cerebellar cortex has already been mentioned in the previous chapter concerning the cerebellum. In particular, many of the models require fast inhibitory feedback from the Golgi cell in order to maintain the sparse code also under conditions where a large portion of the mossy fibers are active. The mostly tonic inhibition that has been found experimentally contradicts this requirement. It is however possible to produce a sparse code with mostly tonic inhibition (Billings et al., 2014). Like the Marr-Albus model, it would require that each granule cell would be innervated by functionally dissimilar mossy fibers. As was described, such combinations of mossy fiber input are implausible, not only from a complexity view-point, but also considering the termination patterns of the incoming mossy fiber pathways.

In addition, direct measurements of granule cells show that many granule cells do not require unique combinations of input to respond (Chadderton et al., 2004, Jörntell and Ekerot, 2006), there seems to be a redundancy of granule cells that respond to the same input (Garwicz et al., 1998, Jörntell and Ekerot, 2006). As was shown with the SBC tract in Paper II, it is possible to drive granule cells stimulating a single mossy fiber tract. The stimulus of a single tract can also significantly modulate the activity of individual Purkinje cell, indicating that the fraction of the granule cell population that does respond is not insignificant. Individual granule cells can also respond strongly to low intensity stimulation from distinct receptive fields. Such responses could be misinterpreted as sparse codes when point-wise stimuli of individual receptive fields are used, while a natural stimulus would rather give rise to a dense code (see Figure 21). Granule cells in particular seem to have a behavior with generally low spontaneous activity, that seem to be strongly regulated even by intrinsic mechanisms in the absence of other inhibition (Brickley et al., 2001). In line with this, the population activity of granule cells in awake mice under locomotion were found to be quiet at rest, but have a dense activity during behavior, modulated in phase with the step cycle (Powell et al., 2015). The activity can be compared with the dense activity predicted by the model of the SBC system in Paper IV (see Figure 15 on page 37).

Sparseness in the neocortex

Similar to the random combinations of mossy fiber input to granule cells, Barlow (1972) proposed that neurons in the visual cortex act as spurious coincidence detectors. The organizational principle said that the higher up the hierarchy of the

cortical areas the neuron was located; the coincidence the neuron was interested in should become more and more specific. The ideas of Barlow have given rise to sparse coding models of the visual cortex that reproduce the properties of the receptive fields of the cortical neurons that have been found experimentally (Olshausen, 1996, Vinje and Gallant, 2000, Olshausen and Field, 2004, Zylberberg and DeWeese, 2013). Both the models and the experiments do however lack the influence from motor activities, and assumes that vision is a passive process, at least at the initial cortical areas of the visual cortex. It is for example common to only include experimental data where the eye remains fixed (Vinje and Gallant, 2000). Saccades and eye movements are instead introduced using a moving stimulus delivered to a fixed eye. It is however clear from recent studies that motor activity strongly influences the firing of neurons in V1 (Keller et al., 2012, Ayaz et al., 2013, Saleem et al., 2013, Erisken et al., 2014), and even changes the structure of their receptive fields (Ayaz et al., 2013). From these studies, it seems that the natural stimuli of V1 neurons should include motor activity. It is also unclear whether sparseness is an epiphenomenon in this circuit, or whether the circuitry actively tries to enforce it. Recent studies suggest that the sparseness of V1 decrease during development (Berkes et al., 2009, Berkes et al., 2011), which would require that the circuitry does not strive to reduce sparseness, but rather uphold a homeostatic equilibrium (Zylberberg and DeWeese, 2013). It can be questioned whether the sparse activity of the resulting circuitry should be used as a defining property of the circuitry in that case.

Sparseness to describe population activity

It should be clear that the activation ratio of the neural population influence several key properties of the encoding. It is at the same time not possible to fully describe a coding scheme and its beneficial or detrimental properties using only the activation ratio. In order to compare the results of both experimental and modeling studies in a meaningful way, at least the size of the population, the method to enforce the sparseness and the spatial structure of the input termination should always be considered. The focus upon the sparseness or activation ratio as the fundamental underlying feature of the encoding might lead to that completely disparate coding schemes, models, or experimental setups are compared simply because they have a similar sparseness. The wide definition of a sparse code (i.e. anything between a dense and local code), further complicates such comparisons.

Conclusions

From an experimental perspective, the complexity of the nervous system leads to difficulties to map the functional properties of different neural circuitry components. Simple point-wise stimuli might provide a functional connectivity map, but it is not possible to extrapolate such data to the functional properties of the circuitry without further data. Paper V illustrates the problem using natural mechanical stimuli to investigate the processing capability of the cuneate nucleus. Classically the cuneate nucleus has been overlooked, mostly due to stimuli that failed to reproduce natural spatiotemporal strain pattern of the skin. Similar problems also arise when the response of granule cells of the cerebellum is investigated using point-wise stimuli (see Figure 21 on page 49), and in the primary visual cortex where the activity and the receptive field properties of the cortical neurons change when motor activity is also included in the stimulation protocol.

The theoretical perspective of Paper I, the modeling perspective of Paper IV and the experimental results of Paper II and V, all indicate the importance of the biomechanics, spinal cord and brainstem for a complete description of the nervous system. In a sense, behavior is shaped already at these levels, before it reaches the cortex or cerebellum. Despite of this, it is unclear what these regions actually do in any detail. Recent models and interpretations of experimental data is however working on constructing functional models of different spinal systems and nuclei in the brainstem (Azim et al., 2014, Alstermark and Ekerot, 2015, Jiang et al., 2015), which together with studies of the spinal circuitry using natural complex stimuli as in as in Paper V could provide more insights into the detailed properties of these systems. In parallel to these studies, the capabilities of the spinal circuitry can be explored through detailed circuitry models. Such models have already shown that the spinal circuitry alone seems to be capable facilitating the control of multi-segmented limbs (Raphael et al., 2010). A fundamental limitation with the decerebrate preparation that was used throughout the thesis is the obvious lack of descending motor command to the spinal and cerebellar circuits. It is known that for example active versus passive touch impose different cortical activity patterns (Ackerley et al., 2012), similar to how motor command or sensory feedback strongly influence the activity of visual processing in V1(Keller et al., 2012, Ayaz et al., 2013, Saleem et al., 2013), including an effect that can be traced as early as the dorsolateral geniculate nucleus of the thalamus (Erisken et al., 2014). From the opposite point of view, the

sensory processing and active exploration might prove to be a fundamental component in motor control (Loeb and Fishel, 2014), where the goal of the motor behavior is not simply to move a limb or grasp an object, but rather to efficiently explore features of the limb's surroundings. It bears close resemblance to the organization of the visual system, where motor activity through saccades and other eye movements are integral in our exploration of our visual environment, even though it is something we barely consciously perceive.

It is perhaps also possible to approach the subcortical circuitry from a developmental perspective. It is clear that simple but useful behavior must arise quickly, and subsequently be increasingly refined during development. Can this perhaps be formulated to a principle of information organization in ascending pathways? Consider the somatotopic map for example. Even though the activity of the cortical neurons must also reflect the complex features of the mechanical touch that were found already in the cuneate nucleus in Paper V, the average activity resembles a mapping of where upon the body the mechanical contact was located. The same is true for the neuronal activity described in Paper III. The neurons have the possibility to have complex response to stimuli, yet they have a regular and reliable underlying activity that can be described by the relatively simple model in Paper III. Neurons do however also process data with impressive detail in their spike timing, which the intracellular and extracellular responses of the cuneate neurons in Paper V show (see Figure 18 on page 42). The ability to encode from a coarse to an increasingly detailed message within the same population activity might be one way of avoiding unnecessary complexity. Initially the receiving end would care only about the average population activity and the overall firing frequency modulation of the incoming signals. Through development, this crude decoding could become increasingly granular, caring about the activity of a decreasing number of neurons with an increasing time resolution of the firing frequency modulation. The useful nature of such an organization in relation to motor control and the inherent complexity has previously been described by Loeb (1983).

The findings in all six papers of the thesis encourage the use of both the following perspectives when the circuits of the nervous system and the biomechanics of the body are explored:

- Inside out: How can the central parts of the nervous system shape itself, both evolutionary and during development, in order to make sense of the afferent input it receives?
- Outside in: How is the biomechanics, afferent sensors, peripheral nervous system, spinal and brainstem circuitry shaped, and the information they convey further up the hierarchy of the nervous system organized, in order to allow the subsequent circuitry to make as much sense of the afferent/ascending input it receives as possible?

The central structures of the nervous system should perhaps not be considered general-purpose computational structures that can resolve any type of input they receive, but rather a structure with constraints on the maximal possible complexity it can handle. Crucially, the brain has coevolved together with the rest of the body. The properties of the afferent input encoding then becomes at least as important as the central circuitry, a sort of preprocessing that enables the hierarchy of the brain, from sensors, through spinal cord, brainstem and the cortex, to efficiently decode and correlate the incoming information. Just as the brain can evolve to improve how efficient it can be controlling the biomechanics, the biomechanics and the periphery of the nervous system can evolve to make itself more accessible to the developing brain.

References

Abbott L, Kepler TB (1990) "Model neurons: From hodgkin-huxley to hopfield. " In: Statistical mechanics of neural networks, pp 5-18: Springer.

Ackerley R, Hassan E, Curran A, Wessberg J, Olausson H, McGlone F (2012) "An fMRI study on cortical responses during active self-touch and passive touch from others." *Frontiers in behavioral neuroscience 6*.

Albus JS (1971) "A theory of cerebellar function." *Mathematical Biosciences 10*.

Alisky J, Tolbert D (1997) "Quantitative analysis of converging spinal and cuneate mossy fibre afferent projections to the rat cerebellar anterior lobe." *Neuroscience 80*.

Alstermark B, Ekerot C-F (2015) "The lateral reticular nucleus; integration of descending and ascending systems regulating voluntary forelimb movements." *Frontiers in computational neuroscience 9*.

Alstermark B, Ekerot CF (2013) "The lateral reticular nucleus: a precerebellar centre providing the cerebellum with overview and integration of motor functions at systems level. A new hypothesis." *The Journal of physiology 591*.

Alstermark B, Isa T, Pettersson LG, Sasaki S (2007) "The C3–C4 propriospinal system in the cat and monkey: a spinal pre-motoneuronal centre for voluntary motor control." *Acta physiologica 189*.

Alstermark B, Wessberg J (1985) "Timing of postural adjustment in relation to forelimb target-reaching in cats." *Acta physiologica scandinavica 125*.

Arshavsky YI, Berkinblit M, Fukson O, Gelfand I, Orlovsky G (1972) "Origin of modulation in neurones of the ventral spinocerebellar tract during locomotion." *Brain research 43*.

Ayaz A, Saleem AB, Schölvinck ML, Carandini M (2013) "Locomotion controls spatial integration in mouse visual cortex." *Current Biology 23*.

Azim E, Jiang J, Alstermark B, Jessell TM (2014) "Skilled reaching relies on a V2a propriospinal internal copy circuit." *Nature 508*.

Barlow HB (1972) "Single units and sensation: A neuron doctrine for perceptual psychology?" *Perception 1*.

Bellman R (1957) "Dynamic Programming." Princeton University Press.

Bengtsson F, Geborek P, Jörntell H (2013) "Cross-correlations between pairs of neurons in cerebellar cortex in vivo." *Neural Networks 47*.

Bengtsson F, Jörntell H (2009) "Sensory transmission in cerebellar granule cells relies on similarly coded mossy fiber inputs." *Proceedings of the National Academy of Sciences 106*.

Berkes P, Orbán G, Lengyel M, Fiser J (2011) "Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment." *Science 331*.

Berkes P, White B, Fiser J (2009) "No evidence for active sparsification in the visual cortex." In: Advances in Neural Information Processing Systems, pp 108-116.

Billings G, Piasini E, Lőrincz A, Nusser Z, Silver RA (2014) "Network structure within the cerebellar input layer enables lossless sparse encoding." *Neuron 83*.

Bosco G, Rankin A, Poppele R (1996) "Representation of passive hindlimb postures in cat spinocerebellar activity." *Journal of neurophysiology 76*.

Brickley SG, Cull-Candy SG, Farrant M (1996) "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*.

Brickley SG, Revilla V, Cull-Candy SG, Wisden W, Farrant M (2001) "Adaptive regulation of neuronal excitability by a voltage-independent potassium conductance." *Nature 409*.

Brunel N, Hakim V, Isope P, Nadal J-P, Barbour B (2004) "Optimal information storage and the distribution of synaptic weights: perceptron versus Purkinje cell." *Neuron 43*.

Burke R, Lundberg A, Weight F (1971) "Spinal border cell origin of the ventral spinocerebellar tract." *Experimental brain research 12*.

Burkitt AN (2006) "A review of the integrate-and-fire neuron model: I. Homogeneous synaptic input." *Biological cybernetics 95*.

Cathala L, Brickley S, Cull-Candy S, Farrant M (2003) "Maturation of EPSCs and intrinsic membrane properties enhances precision at a cerebellar synapse." *The Journal of Neuroscience 23*.

Cesana E, Forti L, Mapelli J, D'Angelo E (2006) "Ionic mechanisms of autorhythmic firing in rat cerebellar Golgi cells." *The Journal of physiology 574*.

Chadderton P, Margrie TW, Häusser M (2004) "Integration of quanta in cerebellar granule cells during sensory processing." *Nature 428*.

Chow CC, White JA (1996) "Spontaneous action potentials due to channel fluctuations." *Biophysical journal 71*.

Clopath C, Brunel N (2013) "Optimal Properties of Analog Perceptrons with Excitatory Weights." *PLoS Comput Biol 9*.

Dean P, Porrill J, Ekerot C-F, Jörntell H (2010) "The cerebellar microcircuit as an adaptive filter: experimental and computational evidence." *Nature Reviews Neuroscience 11*.

Destexhe A, Rudolph M, Fellous J-M, Sejnowski TJ (2001) "Fluctuating synaptic conductances recreate in vivo-like activity in neocortical neurons." *Neuroscience 107*.

Dreyfus S (2002) "Richard Bellman on the birth of dynamic programming." *Operations Research 50*.

Duguid I, Branco T, London M, Chadderton P, Häusser M (2012) "Tonic inhibition enhances fidelity of sensory information transmission in the cerebellar cortex." *The Journal of Neuroscience 32*.

Erisken S, Vaiceliunaite A, Jurjut O, Fiorini M, Katzner S, Busse L (2014) "Effects of locomotion extend throughout the mouse early visual system." *Current Biology 24*.

Fedirchuk B, Stecina K, Kristensen KK, Zhang M, Meehan CF, Bennett DJ, Hultborn H (2013) "Rhythmic activity of feline dorsal and ventral spinocerebellar tract neurons during fictive motor actions." *Journal of neurophysiology 109*.

Forssberg H, Eliasson A, Kinoshita H, Johansson R, Westling G (1991) "Development of human precision grip I: basic coordination of force." *Experimental brain research 85*.

Fourcaud-Trocmé N, Hansel D, Van Vreeswijk C, Brunel N (2003) "How spike generation mechanisms determine the neuronal response to fluctuating inputs." *The Journal of Neuroscience 23*.

Friedman RM, Chen LM, Roe AW (2004) "Modality maps within primate somatosensory cortex." *Proceedings of the National Academy of Sciences of the United States of America 101*.

Földiak P (1990) "Forming sparse representations by local anti-Hebbian learning." *Biological cybernetics 64*.

Garwicz M, Jörntell H, Ekerot CF (1998) "Cutaneous receptive fields and topography of mossy fibres and climbing fibres projecting to cat cerebellar C3 zone." *The Journal of physiology 512*.

Geborek P, Jörntell H, Bengtsson F (2012) "Stimulation within the cuneate nucleus suppresses synaptic activation of climbing fibers." *Frontiers in neural circuits 6*.

Geborek P, Nilsson E, Bolzoni F, Jankowska E (2013) "A survey of spinal collateral actions of feline ventral spinocerebellar tract neurons." *European Journal of Neuroscience 37*.

Gebre SA, Reeber SL, Sillitoe RV (2012) "Parasagittal compartmentation of cerebellar mossy fibers as revealed by the patterned expression of vesicular glutamate transporters VGLUT1 and VGLUT2." *Brain Structure and Function 217*.

Geman S, Bienenstock E, Doursat R (1992) "Neural networks and the bias/variance dilemma." *Neural computation 4*.

Gigerenzer G, Brighton H (2009) "Homo heuristicus: Why biased minds make better inferences." *Topics in Cognitive Science 1*.

Glorot X, Bordes A, Bengio Y (2011) "Deep sparse rectifier neural networks." In: International Conference on Artificial Intelligence and Statistics, pp 315-323.

Hammar I, Krutki P, Drzymala-Celichowska H, Nilsson E, Jankowska E (2011) "A trans-spinal loop between neurones in the reticular formation and in the cerebellum." *The Journal of physiology 589*.

Hastie T, Tibshirani R, Friedman J (2009) "The elements of statistical learning." Springer-Verlag.

Hayward V (2011) "Is there a 'plenhaptic'function?" *Philosophical Transactions of the Royal Society B: Biological Sciences 366*.

Hayward V, Terekhov AV, Wong S-C, Geborek P, Bengtsson F, Jörntell H (2014) "Spatio-temporal skin strain distributions evoke low variability spike responses in cuneate neurons." *Journal of The Royal Society Interface 11*.

Hô N, Destexhe A (2000) "Synaptic background activity enhances the responsiveness of neocortical pyramidal neurons." *Journal of neurophysiology 84*.

Hodgkin AL, Huxley AF (1952) "A quantitative description of membrane current and its application to conduction and excitation in nerve." *The Journal of physiology 117*.

Hollerbach JM, Flash T (1982) "Dynamic interactions between limb segments during planar arm movement." *Biological cybernetics 44*.

Holmberg H, Schouenborg J, Yu Y-B, Weng H-R (1997) "Developmental adaptation of rat nociceptive withdrawal reflexes after neonatal tendon transfer." *The Journal of Neuroscience 17*.

Huang C-C, Sugino K, Shima Y, Guo C, Bai S, Mensh BD, Nelson SB, Hantman AW (2013) "Convergence of pontine and proprioceptive streams onto multimodal cerebellar granule cells." *Elife 2*.

Ito M (2006) "Cerebellar circuitry as a neuronal machine." *Progress in neurobiology 78*.

Ito M (2008) "Control of mental activities by internal models in the cerebellum." *Nature Reviews Neuroscience 9*.

Jankowska E (1992) "Interneuronal relay in spinal pathways from proprioceptors." *Progress in neurobiology 38*.

Jankowska E, Krutki P, Hammar I (2010) "Collateral actions of premotor interneurons on ventral spinocerebellar tract neurons in the cat." *Journal of neurophysiology 104*.

Jankowska E, Nilsson E, Hammar I (2011a) "Do spinocerebellar neurones forward information on spinal actions of neurones in the feline red nucleus?" *The Journal of physiology 589*.

Jankowska E, Nilsson E, Hammar I (2011b) "Processing information related to centrally initiated locomotor and voluntary movements by feline spinocerebellar neurones." *The Journal of physiology 589*.

Jiang J, Azim E, Ekerot C-F, Alstermark B (2015) "Direct and indirect spino-cerebellar pathways: shared ideas but different functions in motor control." *Frontiers in computational neuroscience 9*.

Johansson F, Jirenhed D-A, Rasmussen A, Zucca R, Hesslow G (2014) "Memory trace and timing mechanism localized to cerebellar Purkinje cells." *Proceedings of the National Academy of Sciences 111*.

Johansson RS, Flanagan JR (2009) "Coding and use of tactile signals from the fingertips in object manipulation tasks." *Nature Reviews Neuroscience 10*.

Jörntell H, Ekerot C-F (2002) "Reciprocal bidirectional plasticity of parallel fiber receptive fields in cerebellar Purkinje cells and their afferent interneurons." *Neuron 34*.

Jörntell H, Ekerot C-F (2003) "Receptive field plasticity profoundly alters the cutaneous parallel fiber synaptic input to cerebellar interneurons in vivo." *The Journal of Neuroscience 23*.

Jörntell H, Ekerot C-F (2006) "Properties of somatosensory synaptic integration in cerebellar granule cells in vivo." *The Journal of Neuroscience 26*.

Kandel ER, Schwarz JH (2013) "Principles of Neural Science." McGraw-Hill Medical.

Kawato M, Furukawa K, Suzuki R (1987) "A hierarchical neural-network model for control and learning of voluntary movement." *Biological cybernetics 57*.

Keller GB, Bonhoeffer T, Hübener M (2012) "Sensorimotor mismatch signals in primary visual cortex of the behaving mouse." *Neuron 74*.

Kimpo RR, Rinaldi JM, Kim CK, Payne HL, Raymond JL (2014) "Gating of neural error signals during motor learning." *Elife 3*.

Kostal L, Lánský P (2007) "Variability and randomness in stationary neuronal activity." *Biosystems 89*.

Krutki P, Jelen S, Jankowska E (2011) "Do premotor interneurons act in parallel on spinal motoneurons and on dorsal horn spinocerebellar and spinocervical tract neurons in the cat?" *Journal of neurophysiology 105*.

La Camera G, Giugliano M, Senn W, Fusi S (2008) "The response of cortical neurons to in vivo-like input current: theory and experiment." *Biological cybernetics 99*.

Loeb GE (1983) "Finding common groud between robotics and physiology." *Trends in Neurosciences 6*.

Loeb GE, Fishel JA (2014) "Bayesian Action&Perception: Representing the World in the Brain." *Frontiers in neuroscience 8*.

Lundberg A, Weight F (1971) "Functional organization of connexions to the ventral spinocerebellar tract." *Experimental brain research 12*.

Mainen ZF, Sejnowski TJ (1995) "Reliability of spike timing in neocortical neurons." *Science 268*.

Marr D (1969) “A theory of cerebellar cortex.” *The Journal of physiology 202*.

Matsushita M, Hosoya Y, Ikeda M (1979) “Anatomical organization of the spinocerebellar system in the cat, as studied by retrograde transport of horseradish peroxidase.” *Journal of Comparative Neurology 184*.

Matsushita M, Ikeda M (1980) “Spinocerebellar projections to the vermis of the posterior lobe and the paramedian lobule in the cat, as studied by retrograde transport of horseradish peroxidase.” *Journal of Comparative Neurology 192*.

Matsushita M, Yaginuma H (1989) “Spinocerebellar projections from spinal border cells in the cat as studied by anterograde transport of wheat germ agglutinin-horseradish peroxidase.” *Journal of Comparative Neurology 288*.

Matsuyama K, Drew T (2000a) “Vestibulospinal and reticulospinal neuronal activity during locomotion in the intact cat. I. Walking on a level surface.” *Journal of neurophysiology 84*.

Matsuyama K, Drew T (2000b) “Vestibulospinal and reticulospinal neuronal activity during locomotion in the intact cat. II. Walking on an inclined plane.” *Journal of neurophysiology 84*.

McKay BE, Molineux ML, Mehaffey WH, Turner RW (2005) “Kv1 K+ channels control Purkinje cell output to facilitate postsynaptic rebound discharge in deep cerebellar neurons.” *The Journal of Neuroscience 25*.

Minsky M (1988) “Society of mind.” Simon and Schuster.

Minsky ML, Papert SA (1969) “Perceptrons.” MIT press Boston, MA.

Minsky ML, Papert SA (1987) “Perceptrons - Expanded Edition: An Introduction to Computational Geometry.” MIT press Boston, MA.

Molineux ML, Fernandez FR, Mehaffey WH, Turner RW (2005) “A-type and T-type currents interact to produce a novel spike latency-voltage relationship in cerebellar stellate cells.” *The Journal of Neuroscience 25*.

Mountcastle VB (1997) “The columnar organization of the neocortex.” *Brain 120*.

Naundorf B, Wolf F, Volgushev M (2006) “Unique features of action potential initiation in cortical neurons.” *Nature 440*.

Oja E, Karhunen J, Wang L, Vigario R (1995) “Principal and independent components in neural networks-recent developments.” In: Seventh Italian Workshop on Neural Networks, Vietri, pp 16-35: Citeseer.

Olshausen BA (1996) “Emergence of simple-cell receptive field properties by learning a sparse code for natural images.” *Nature 381*.

Olshausen BA, Field DJ (2004) “Sparse coding of sensory inputs.” *Current opinion in neurobiology 14*.

Oscarsson O (1973) “Functional organization of spinocerebellar paths.” In: Somatosensory system, pp 339-380: Springer.

Petersson P, Waldenström A, Fåhraeus C, Schouenborg J (2003) “Spontaneous muscle twitches during sleep guide spinal self-organization.” *Nature 424*.

Poppele R, Bosco G, Rankin A (2002) “Independent representations of limb axis length and orientation in spinocerebellar response components.” *Journal of neurophysiology 87*.

Porrill J, Dean P (2008) “Silent synapses, LTP, and the indirect parallel-fibre pathway: computational consequences of optimal cerebellar noise-processing.” *PLoS Comput Biol 4*.

Powell K, Mathy A, Duguid I, Häusser M (2015) “Synaptic representation of locomotion in single cerebellar granule cells.” *Elife 4*.

Prochazka A, Gorassini M (1998) “Ensemble firing of muscle afferents recorded during normal locomotion in cats.” *The Journal of physiology 507*.

Raphael G, Tsianos GA, Loeb GE (2010) “Spinal-like regulator facilitates control of a two-degree-of-freedom wrist.” *The Journal of Neuroscience 30*.

Rauch A, La Camera G, Lüscher H-R, Senn W, Fusi S (2003) “Neocortical pyramidal cells respond as integrate-and-fire neurons to in vivo–like input currents.” *Journal of neurophysiology 90*.

Roberts BA, Rubenstein DI (2014) “Maternal tactics for mitigating neonate predation risk during the postpartum period in Thomson’s gazelle.” *Behaviour 151*.

Rosenblatt F (1958) “The perceptron: A probabilistic model for information storage and organization in the brain.” *Psychological Review 65*.

Rumelhart DE, Hinton GE, Williams RJ (1985) “Learning internal representations by error propagation.” DTIC Document.

Russell S, Norvig P (2009) “Artificial intelligence: a modern approach, 3rd Edition.” Pearson.

Rössert C, Solinas S, D'Angelo E, Dean P, Porrill J (2014) “Model cerebellar granule cells can faithfully transmit modulated firing rate signals.” *Frontiers in cellular neuroscience 8*.

Saleem AB, Ayaz A, Jeffery KJ, Harris KD, Carandini M (2013) “Integration of visual motion and locomotion in mouse visual cortex.” *Nature neuroscience 16*.

Santello M, Baud-Bovy G, Jorntell H (2013) “Neural bases of hand synergies.” *Frontiers in computational neuroscience 7*.

Santello M, Flanders M, Soechting JF (1998) “Postural hand synergies for tool use.” *The Journal of Neuroscience 18*.

Schweighofer N, Doya K, Lay F (2001) “Unsupervised learning of granule cell sparse codes enhances cerebellar adaptive control.” *Neuroscience 103*.

Sejnowski TJ (1977) “Storing covariance with nonlinearly interacting neurons.” *Journal of mathematical biology 4*.

Shrestha SS, Bannatyne B, Jankowska E, Hammar I, Nilsson E, Maxwell D (2012a) "Inhibitory inputs to four types of spinocerebellar tract neurons in the cat spinal cord." *Neuroscience 226*.

Shrestha SS, Bannatyne BA, Jankowska E, Hammar I, Nilsson E, Maxwell DJ (2012b) "Excitatory inputs to four types of spinocerebellar tract neurons in the cat and the rat thoraco-lumbar spinal cord." *The Journal of physiology 590*.

Thom M, Palm G (2013) "Sparse activity and sparse connectivity in supervised learning." *The Journal of Machine Learning Research 14*.

Tolbert DL, Knight TL (2003) "Persistence of spinocerebellar afferent topography following hereditary Purkinje cell degeneration." *The Cerebellum 2*.

Wall MJ, Usowicz MM (1997) "Development of Action Potential-dependent and Independent Spontaneous GABAA Receptor-mediated Currents in Granule Cells of Postnatal Rat Cerebellum." *European Journal of Neuroscience 9*.

Van Dijck G, Van Hulle MM, Heiney SA, Blazquez PM, Meng H, Angelaki DE, Arenz A, Margrie TW, Mostofi A, Edgley S (2013) "Probabilistic identification of cerebellar cortical neurones across species." *PloS one 8*.

van Vreeswijk C, Sompolinsky H (1996) "Chaos in neuronal networks with balanced excitatory and inhibitory activity." *Science 274*.

Wang SS-H, Khiroug L, Augustine GJ (2000) "Quantification of spread of cerebellar long-term depression with chemical two-photon uncaging of glutamate." *Proceedings of the National Academy of Sciences 97*.

Vinje WE, Gallant JL (2000) "Sparse coding and decorrelation in primary visual cortex during natural vision." *Science 287*.

Zachariah M, Coleman G, Mahns D, Zhang H, Rowe M (2001) "Transmission Security for Single, Hair Follicle–Related Tactile Afferent Fibers and Their Target Cuneate Neurons in Cat." *Journal of neurophysiology 86*.

Zhong G, Droho S, Crone SA, Dietz S, Kwan AC, Webb WW, Sharma K, Harris-Warrick RM (2010) "Electrophysiological characterization of V2a interneurons and their locomotor-related activity in the neonatal mouse spinal cord." *The Journal of Neuroscience 30*.

Zylberberg J, DeWeese MR (2013) "Sparse coding models can exhibit decreasing sparseness while learning sparse codes for natural images." *PLoS computational biology 9*.

Populärvetenskaplig sammanfattning

Ett utav de stora hinder som fortfarande ligger i vägen för våra försök att beskriva hjärnan, är hur komplex den är. En människohjärna består av miljarder nervceller och sinsemellan är de sammankopplade i nätverk med trillioner kopplingar. Komplexitet behöver i allmänhet inte vara ett problem. Vi kan till exempel beskriva hur gaser och vätskor fungerar, trots att det ofta rör sig om ofattbart många atomer som interagerar och samverkar för att gasen eller vätskan ska få sina egenskaper. Problemet med att reducera hjärnans komplexitet på samma sätt är att när vi reducerar komplexiteten hos något, förutsätter det att det inte är själva komplexiteten vi är intresserade av. Tyvärr gäller inte det hjärnan.

Det som är så imponerande med en hjärna är nämligen att den klarar av att ge mening och interagera med vår omvärld. För om det är något som är mer komplext än vår hjärna, så är det vår omvärld. I den finns ju trots allt tusentals andra hjärnor som man kan interagera med. Komplexiteten i omvärlden ligger emellertid också närmre hjärnan än så. I vår egen kropp finns hundratusentals sensorer utspridda som mäter temperatur, ljus (och bilder!), ljud, beröring – till och med hur spänd varje enskild muskelfiber är i varenda muskel i kroppen. Inte nog med att hjärnan belastas med information från alla dessa sensorer, den styr samtidigt en komplex kropp med över 200 ben och 600 skelettmuskler.

Sanningen är att vi inte riktigt vet hur man ska göra för att hantera något så komplext. Robotforskare bygger till exempel numera robotar vars mekanik försöker härma människokroppen för att försöka lista ut vilka knep som döljer sig bakom alla senor och muskler. Sanningen är också att vi måste försöka förstå ändå. Antalet sjukdomar som har kopplingar till nervsystemet där vi hade haft nytta av att veta hur det faktiskt fungerar är många.

Den här avhandlingen försöker hitta metoder med en reducerad komplexitet, men med vilka man kan undersöka hjärnan utan att man förlorar möjligheten att förutsäga hur den egentligen fungerar. Ett viktigt perspektiv är att en hjärna själv måste reducera komplexiteten i det den gör. Antalet sensorer och muskler är så många att hjärnan inte räcker till annars. Genom att försöka hitta de sätt som hjärnan själv använder för att hantera omvärldens komplexitet, kan det möjligtvis gå att beskriva hur det går till när hjärnan växer från något enkelt till något komplext. Den andra observationen man kan göra är nämligen att det tar tid för en hjärna att bli färdig,

men trots det måste den fungera på en grundläggande nivå redan vid födseln. Vi förutsätter inte att en nyfödd ska kunna spela piano, men den kan redan från början kontrollera en relativ komplicerad kropp.

Genom att undersöka egenskaperna hos nervsystemet nedanför själva hjärnan, i hjärnstammen, ryggmärgen och lillhjärnan, kan man försöka hitta de principer som hjärnan använder för att organisera informationen som når hjärnan, för att redan ett nyfött djur kan börja dra nytta av den. I de sex delarbetena i den här avhandlingen visar vi exempel på hur hjärnan tjänar på att utnyttja de ofta förbisedda delarna av nervsystemet som finns under hjärnan. Vi ger också exempel på hur det i hjärnstammen finns nervceller som utför uppgifter som man tidigare trott skett först uppe i hjärnan.

I) Det första delarbetet handlar om lillhjärnan, som volymmässigt är stor som en knytnäve och därmed ganska liten jämfört med neocortex eller 'storhjärnan'. Trots sin lilla volym finns här en majoritet av alla nervceller. Den gängse bilden av vad lillhjärnan gör är bland annat att förbättra koordination mellan funktioner i andra delar av nervsystemet. Att det finns så många nervceller just här skulle i sådana fall kunna beror på att många utav hjärnans andra delar skickar information hit, som måste kombineras på olika sätt för att lillhjärnan ska kunna använda dem effektivt. Det visar sig emellertid att om man inte kombinerar alla signaler som kommer från vitt skilda håll i nervsystemet på ett smart sätt, kommer det bli för komplext att hantera till och med för lillhjärnan, trots alla miljarder nervceller som finns till hands. En del av lösningen som vi föreslår är att informationen kombineras redan innan den når lillhjärnan, i de nätverk av nervceller som redan finns i ryggmärg och hjärnstam.

II) Det andra delarbetet handlar om precis ett sådant nätverk av nervceller som skickar signaler från nedre delen av ryggmärgen till lillhjärnan. Det visar sig att det räcker med att stimulera och aktivera bara ett enskilt knippe av nervceller för att en oväntat stor andel av nervcellerna i lillhjärnan ska reagera. Precis som förväntat om den mesta av informationen kombineras redan innan lillhjärnan. När man en gång spelat in hur nervcellerna i lillhjärnan svarar på stimuleringen av nervcellsknippet i ryggmärgen, kan man försöka bygga matematiska modeller som förklarar varför cellerna svarar på ett visst sätt, eller som helt enkelt klarar av att reproducera liknande svar. Sedan kan man använda modellerna som beskriver hur en enskild nervcell beter sig som byggstenar i en nätverksmodell som i sin tur beskriver hur sammankopplade nervceller beter sig.

III) Många nervceller i nervsystemet är spontant aktiva. Det betyder att de skickar signaler till varandra även om ingen stimulerat dem. På sidan 15 i Figur 5 finns exempel på fyra sådana spontant aktiva nervceller som spelats in. Trots att de inte stimuleras aktivt skickar de iväg en puls av aktivitet till andra nervceller då och då. Närbilden i Figur 5E visar dessutom hur spontanaktiviteten ser ut när den tas emot av en nervcell. I stort sett allt det brus som man kan se mellan pulserna beror på att

tusentals andra nervceller hela tiden bombarderar den inspelade nervcellen med sin spontanaktivitet. Genom att stimulera en spontanaktiv nervcell kan man modulera den spontana aktiviteten så att den skickar signaler mer eller mindre sällan. Det tredje delarbetet handlar om hur nervcellers spontanaktivitet ser ut. Trots väldigt skilt utseende (se till exempel Figur 4 på sidan 12) och därmed många andra skilda egenskaper, tyder resultaten på att de nervceller från ryggmärgen och lillhjärnan som undersöktes har ett grundläggande beteende som inte skiljer sig särskilt mycket mellan de olika nervcellstyperna.

IV) I det fjärde delarbetet användes beskrivningen av nervcellers spontanaktivitet från delarbete tre, tillsammans med hur nervcellerna i lillhjärnan beter sig från delarbete två, för att bygga en modell som simulerade hur nervcellerna i lillhjärnan, som tar emot signalerna från ryggmärgen, beter sig när vi promenerar. Resultaten visar att den genomgripande aktiviteten på ett enkelt sätt återspeglar grundläggande egenskaper i stegcykeln. Nervcellskretsen verkar alltså på det övergripande planet försöka undvika att göra informationen som kommer in mer komplicerad än vad som behövs. Anledningen till att vi inte hittade något mer komplext beteende host nervcellerna var kanske helt enkelt att vi undersökte dem i ett sammanhang som inte kräver särskilt mycket komplexitet.

V) Eftersom hjärnan trots allt måste klara av komplicerade beräkningar måste nervcellerna också klara av att skicka detaljerade signaler. Anledningen till att det inte dök upp några sådana tecken i delarbete fyra kan helt enkelt vara att kretsen i det läget inte hade någon nytta av komplicerade signaler och därför undvek dem. I det femte delarbetet användes istället en mekanisk display som kan simulera att huden som rör vid displayen istället tar på ett föremål. Med denna typ av stimulering visade det sig ske informationsbearbetning redan i hjärnstammen, där nervcellerna kunde urskilja egenskaperna hos olika typer av mekanisk kontakt med huden. Den klassiska bilden är att sådana saker inte sker förrän tre steg senare i en del av neocortex som heter sekundära somatosensoriska cortex. Om de beräkningarna kan utföras redan i hjärnstammen innebär det att delarna av neocortex där man tidigare trodde beräkningarna gjordes kan användas till annat.

VI) I det sjätte delarbetet undersöks huruvida grupper av nervceller har som mål att vara så lite aktiva som möjligt. Tanken är att om man som nervcell bara är aktiv i vissa speciella situationer, kan man undvika att störa andra nervceller när de är aktiva. Det är en princip som kallas sparse coding (gles kodning på svenska). Problemet är att det finns nackdelar med en för gles aktivitet, framför allt när det gäller hjärnans förmåga att generalisera redan inlärd kunskap till nya situationer. För att uppnå de positiva egenskaperna av en gles aktivitet måste dessutom aktiviteten bli mer och mer gles ju fler nervceller som är inblandade. Det innebär att man i nätverk som är lika stora som de som finns i hjärnan måste ha en extremt gles kod. Sammantaget tyder det på att hjärnan har hittat andra sätt än en gles aktivitet för att undvika problemen med

komplexitet. Slutligen verkar det som om de experimentella observationer som är gjorda när nervcellskretsarna utsätts för situationer som liknar naturligt beteende inte har en gles aktivitet under tiden de stimuleras.

De sex delstudierna tillsammans inbjuder till en syn där man ser på hjärnan från ett annat perspektiv än som är vanligt. Istället för att hjärnan finns för att lära sig för att förstå allt som sker i omvärlden och hur den egna kroppen fungerar, kan man snarare vända på det och ge kroppen och de tidiga delarna av nervsystemet uppdraget att forma hjärnan med hjälp av de sensoriska signaler som genereras ute i muskler och hud, som sedan moduleras i ryggmärg och hjärnstam. I grunden handlar det om att kropp, ryggmärg, hjärnstam och hjärna utvecklats evolutionärt ihop. Lika gärna som hjärnan kan vara byggd för att förstå sig på kroppen, kan kroppen vara byggd för att göra sig mer förståelig för resten av nervsystemet.

Acknowledgements

For several years, through my master thesis and my time as a Ph.D. student, I have had the privilege to work in a group of colleagues that all share the epitome of the classical scientist. Your passion to the field of neuroscience has been contagious. Without that contagion, what started as a master thesis five years ago would never had led to this thesis.

Henrik Jörntell

If you read through the acknowledgements chapters of almost all Ph.D. theses in neurophysiology from Lund produced over the last decade (perhaps more), Henrik's name will appear again and again, with praises of his phenomenal knowledge. I cannot do worse of course. Without the library of knowledge that resides within Henrik's head, not a single part of this thesis would exist. Having had Henrik as my supervisor however, I would also like to thank him for sharing his passion for neuroscience, and his drive to understand how it all works and fits together. Throughout my time in his lab, he has also never failed to display an amazing amount of trust in me and his other Ph.D. students. I have never felt like I have worked on Henrik's projects – through his leadership he has always made it feel as if they were ours.

Rolf Johansson

Without the knowledge of mathematical modeling that you have shared with me, through courses, projects, books and discussions, I would have had none of the insights that have been absolutely vital during these years as a Ph.D. student. You have tried to teach me how to use math in practice to some success, but most of all you have inspired me to continue to learn and explore the intricacies of the world. The world would be such a boring and unexplainable place without the mathematical toolset you have given me.

Fredrik Bengtsson, Pontus Geborek, Hannes Mogensen, Jonas Enander

Without all of your hard work in the lab, the very foundation of this thesis would not exist. **Pontus Geborek**, your amazing amount of work and dedication to the lab has led to not only brilliant experiments and results, but you managed to redesign most of the experimental setup of the lab during your time here. I admire your practical

problem solving skills, and I truly enjoyed being part of your team, developing more or less strange gadgets and devices. **Fredrik Bengtsson**, through your sharp intuition and somewhat averse view upon modeling, you have kept me on my toes and forced me to value any experimental finding over even the fanciest network models and computational principles. Furthermore, thank you for allowing me to torment your medical students with basic electrodynamics and electrophysiology. Many of my insights have been gained through the perspectives that they have shared with me as naïve students. **Hannes Mogensen**, it took your creative mind to deprecate the last remnant of the 90s in the lab. With a small amount of Swedish 'vemod' leaving blisteringly fast assembly programs behind, we are now heading into the 21th century ready for new amazing experiments. **Jonas Enander**, I am at awe of your organizational skills, perhaps only surpassed by Fredrik Johansson's. I wish that I could say that I have learned at least some planning skills from you, but I am a lost cause. With the hard work you put into everything you do, I do not need to say good luck, as we know it will come to you sooner or later.

Fredrik Johansson, Dan-Anders Jirenhed, Anders Rasmussen, Germund Hesslow

Our dear neighbors at F10. All four of you serve as brilliant scientific role models. Your tireless fight against the wind of the scientific community with your recent discoveries only reinforces the view I have of you as scientists. You stand at the brink of an untouched field of experimental neuroscience and I truly look forward to all of your coming discoveries.

Kersti Larsson

Without your selfless dedication to take care of both our groups, using your long experience and your up-to-date expertise, chaos would soon erupt at F10.

Max Martinsson, Johannes Ulén, Stefan Evertsson

My reliable lunch partners through all these years. Thank you for listening to my close to crazy ramblings each time I try to persuade you to believe in my new ideas. Thank you for not always believing in them. **Max**, your impressive ability to realize what I actually mean with my ramblings and that you manage to ask the questions that completely shoot the new idea to the ground, has been an invaluable tool in my research. If an idea lasts through the lunch break, it might have some bearing. **Johannes**, in addition to the typographical advice and expertize, your unique and stringent mathematical perspectives has been immensely useful for me. Thank you for always trying to answer my questions. **Stefan**, without your hard work, dedication, and belief into what Max and I (and now you as well) have created, I would never have had time to finish this thesis in time.

Min kära familj

Mamma och Jon, tack för att ni alltid funnits där och för allt viktigt ni lärt mig om hur man ska vara och vad som spelar roll i livet. Farmor, var stolt över dina underbara barnbarn och barnbarnsbarn. Utan din livsglädje och den kärlek som du, farfar och mormor gett så villkorslöst hade världen varit en sämre plats. Du är en förebild som man bara kan hoppas att man lever upp till.

Linnea, Bosse, Sanna, Göran, Ulla, Bob och Bruna

Tack till er alla sju, för utan er hade min värld varit platt och utan mening. Från er hämtar jag all inspiration och energi. Ett speciellt tack till min kära fru som stått ut med mig längst, och till skillnad från de stackars djuren självmant valt att bo med mig.

Knerten

Du är mitt och Linneas efterlängtade nya äventyr som kommer förändra vårt liv i grunden, samtidigt som det alltid varit självklart att du varit på väg. Tack också för att du tvingade mig att skriva klart den här avhandlingen i tid.