BIOPHYSICAL MODELING AND OPTICAL IMAGING TOOLS FOR STUDIES OF CEREBELLAR MOTOR LEARNING

A DISSERTATION

SUBMITTED TO THE DEPARTMENT OF PHYSICS

AND THE COMMITTEE ON GRADUATE STUDIES

OF STANFORD UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Eran Abraham Mukamel

December 2008

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as dissertation for the degree of Doctor of Philosophy.

Mark J. Schnitzer, Principal Advisor

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as dissertation for the degree of Doctor of Philosophy.

Jennifer Raymond

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as dissertation for the degree of Doctor of Philosophy.

Terence D. Sanger

Approved for the University Committee on Graduate Studies

Abstract

Motor learning is a precious capacity that humans, from toddlers to athletes and concert pianists, rely on to move with coordination and grace. The neural circuit in the cerebellum that helps us learn new movements and adjust our performance of familiar reflexes is remarkable for its quasi-crystalline architecture, which has been conserved at the cellular level across vertebrate species. This thesis develops and applies tools based on biophysical and computational modeling of neural circuits, as well as analysis of data sets from optical microscopy experiments, to study mechanisms of motor learning in the cerebellum. In the first part of the thesis, we use theoretical methods, including electrical compartmental modeling of single cells and a statistical theory of learning in networks, to explore hypotheses concerning temporal aspects of learned movements. We propose that the geometric structure of the cerebellar cortex may be useful for improving the neural representation of time. We go on to use a "lock and key" theory to explain a longstanding observation regarding the required causal temporal relationship among sensory stimuli for effective motor training. Our theory relates this timing aspect of motor learning to a biophysical mechanism, rebound conductances, present in the neurons of the deep cerebellar nuclei.

In the second half of the thesis we develop and apply computational and statistical analysis procedures for analyzing experimental data sets from optical imaging experiments to test hypotheses regarding motor learning in the cerebellum. We present tools that can extract intracellular physiological signals, including neuronal spike trains and glial calcium transients, from noisy and high-dimensional imaging data sets recorded in living subjects. Using a novel miniaturized microscope, we measured intracellular signals as well as blood flow in freely moving mice during rest and activity. To examine intracellular signals in finer detail, we examined data recorded by two-photon imaging in head-fixed mice free to run on an exercise wheel. Using this technique, we discovered three forms of dynamic activation in networks of cerebellar Bergmann glia, one of which is associated with the onset of locomotor activity.

High-dimensional optical imaging experiments require new computational approaches to data analysis. We developed automated procedures, including spatio-temporal independent component analysis, to identify the locations and extract the temporal dynamics of individual cells in our movie data sets. Our procedures solve the cell sorting problem that is analogous to the challenge of spike sorting in extracellular electrophysiology. Cell sorting of calcium imaging data revealed patterns of correlated complex spiking in groups of Purkinje cell dendrites. Using *in vivo* optical microscopy we showed that microzones of correlated complex spiking are stable over time and between resting and actively moving behavioral states. Optical imaging in combination with statistical and computational approaches to data analysis may help in coming years to answer longstanding questions concerning the cellular and network mechanisms of motor learning in the cerebellum.

Acknowledgments

The work described in these pages represents the products of a scientific journey I have taken as a graduate student with the mentorship, collaboration and support of many people. I am grateful for the remarkable group of talented, passionate, creative, caring and supportive friends and colleagues with whom I had the privilege of working at Stanford. During the four years of my tenure in Mark Schnitzer's lab, my colleagues constantly impressed and inspired me with their sheer enthusiasm for science, and their energy was an invaluable inspiration and example for me through the years. I thank them not only for our shared scientific experiences whose results are described here, but also for the many intangible lessons and skills I have learned along the way.

Larry Abbott recently described the entry of a wave of physicists into computational neuroscience in the past twenty years in these words: "These new immigrants entered with high confidence-to-knowledge ratios that, hopefully, have been reduced through large growth in the denominators and more modest adjustments of the numerators" [1]. In my case, I certainly owe much of the adjustment in both those parameters to my advisor, Mark Schnitzer. Everyone who works with Mark is awed by his energy and zeal for science, particularly at early hours of the morning when the rest of us can find zeal only for coffee or sleep. He also has great passion and dedication for the people he works with, and all of us in the lab have enjoyed knowing that Mark is on our side.

I was lucky to find in Daniel Wetmore an excellent collaborator and friend. The idea that rebound conductances in the deep cerebellar nuclei might play an important role in the timing of cerebellum dependent behavior (Chapter 2) grew out of Dan's inspiration. In the course of fleshing out that hypothesis, arguing over science and struggling with the arcanae of our code (sometimes at odd hours), I learned a great deal from Dan's broad base of neuroscience knowledge and keenly critical approach to theoretical models. I've learned to rely on Dan's diligence and scientific honesty, as well as his unfailing good nature, and he has set a high standard for my future collaborators.

I learned early in my career in neuroscience that the study of the brain is still immature, and progress is very much driven by new empirical discoveries. I am grateful to Axel Nimmerjahn for the opportunity that our collaboration has given me to contribute to experimental neuroscience by analyzing and interpreting rich data sets from optical imaging in the living brain. When Axel first showed me the movies he was starting to record, I was amazed by the possibility of seeing the brain's internal processes in action. I have come to appreciate that these remarkable data are the result of Axel's talent and passion for biological experimentation. Axel provided a steady stream of data from ambitious, but carefully crafted, experiments, and he gave me the creative space to do my own experiments (with pencil, paper, and processor) in developing an approach to their analysis. Axel's broad knowledge, his curiosity and intellectual rigor made our collaboration a pleasure, and I could always look forward to stimulating and enlightening discussions with him about science

Several other members of Mark Schnitzer's laboratory were exceptional collaborators and colleagues on a number of projects, as well as fine comrades and companions during our time in the basement of Stanford's Clark Center. I am grateful to Benjie Flusberg, who brought his curiosity and his of clarity to bear on many stimulating conversations, ranging from optics to quantum field theory and even meteorology. Early in my neuroscience career, Amit Mehta and Ed Boyden helped me to delve into the cerebellar literature, and their excitement was infectious. Juergen Jung, Robert Barretto, Eric Cocker, Lynn Sun, and Eunice Cheung were wonderful colleagues from the lab's first years at Stanford. In recent years, I have enjoyed working with Wibool Piyawattanametha, Alessio Attardo, Laurie Burns, Fang Ping Chen, Saeed Hassanpour, Tony Ko, Eric Tattwei Ho, Mike Molineux, Jacob Rinaldi and Brian Wilt. Annette Lewis kept the lab running and helped us on several occasions to sharpen our thinking as we prepared papers. Andrea Lui played a critical role in the projects involving two-photon imaging by assisting with habituation of mice to the head-fixed recording procedure. She also contributed valuable code for performing image registration and reducing movement artifacts in our live brain imaging studies.

In my time at Stanford I was fortunate to interact with several fellow theory students, who satisfied my need to talk about math problems and gave me hope that perhaps something new can truly come out of a chat in front of a whiteboard. James Fitzgerald and Jeremy England are brilliant physical thinkers with whom I worked on theoretical problems in the Schnitzer lab. Before starting my work on neuroscience, I learned a great deal from a rotation experience in the group of Shoucheng Zhang. Chi-Hong Chern, Andrei Bernevig, Congjun Wu and Handong Chen were wonderful colleagues.

My graduate work was funded in part by a National Science Foundation Graduate Research Fellowship. I am also grateful for a scholarship to attend the Methods in Computational Neuroscience course in the summer of 2006 at Woods Hole, Massachussetts, organized by Bard Ermentrout and John White, which was a stimulating and formative experience for me. The community of neuroscience graduate students at Stanford was extremely impressive, and I particularly appreciate the scientific, meta-scientific and even non-scientific interactions I have enjoyed with Daniel Kimmel, Ilana Witten, Christing Cong, Alex Katsov, Rodrigo Alvarez-Icaza, Morteza Shahram, Raag Airan, Leslie Meltzer and Viviana Gradinaru.

A number of scientists have generously helped me and my collaborators refine our work by providing comments and suggestions for our manuscripts. David Linden, Indira Raman, Krishna Shenoy and Byron Yu each helped us in this way. I have also learned a great deal from interactions with Haim Sompolinsky, Terry Sejnowski, Joshua Vogelstein, Jakob Macke and Mark Goldman. Some experimentalists provided sample data and took the time to talk with us about our approaches to data analysis. For this help I am grateful to Shaul Hestrin, Clay Reid, Carlos Portera-Cailliau and Cris Niell.

I owe a debt of thanks to each of the members of my dissertation committee, who have contributed in important ways to my thesis and my education. Jennifer Raymond and the members of her lab were ready interlocutors whenever I wanted to talk about the cerebellum. We always had readers like Jen Raymond and her students in mind when we thought about who would give us the sharpest criticism and the most credible praise. Terry Sanger was a source of advice on machine learning and statistical questions. Seb Doniach was a sounding board for the big questions of neuroscience, and his curiosity about the implications and larger meaning of developments in neuroscience constantly pushed me to think more deeply. I have enjoyed discussing cerebella, both biological and artificial, with Kwabena Boahen and the members of his neuromorphics laboratory.

The staff of the Stanford Physics Department supported me throughout my time at Stanford. I am particularly grateful to our Student Services Manager, Maria Frank, for tireless and dedicated work on behalf of the graduate students. Maria's good cheer made me look forward to visits to her office, and she always helped to solve whatever bureaucratic boondoggle might arise.

I am grateful to a group of friends who provided support and sustenance in all sorts of ways in the past few years. My roommates, Jan Lipfert, Mark Allen and Adam Cohen, were like brothers. I am grateful for the times I've shared with Mina Arvanitaki (possessor of the most infectious laugh on campus), Kathryn Todd and Alex Dunn, Ruben Kogel, Jon Russell, Lindsay Moore, Elene Terry and Matt Pyle. I've had several wonderful musical partners, including Jim Zehnder, Jon Lee, and Naomi Zamir, as well as my incredible teachers, Debra Fong, Chris Costanza, Geoff Nuttall and Barry Shiffman.

My parents, Shaul and Dana, have always inspired me to chase my dreams, and I know that this thesis could not have been completed without their support at many times along the way. Now that I've made it through grad school I hope they won't worry so much about how I'm going to feed myself. I'm lucky that my brother, Ronen, is also a best friend. As my younger brother he expects a lot of me, and as a mathematician he knows all the flaws in my logic. I've recently joined another family, the Glasers. John and Carol Glaser have been like a second set of parents, and being able to hang out with them has been a huge perk of living in the bay area. I always have fun with Bronwyn Glaser, who knows how to find the best things in life and go after them. Unfortunately she's found them in Switzerland, but whenever I do see her I get a huge kick.

Finally, and most of all, I thank Amelia. She is the best thing I've gotten out of Stanford, where I met her when we were neighbors crowded together in the hovels (her word: paradise) designated as graduate housing. She shows me the poetry in life, and she has taught me to see the brain through a poetic lens as well:

THE BRAIN is wider than the sky,
  For, put them side by side,
The one the other will include
  With ease, and you beside.

The brain is deeper than the sea,
For, hold them, blue to blue,
The one the other will absorb,
As sponges, buckets do.

The brain is just the weight of God,
For, heft them, pound for pound,
And they will differ, if they do,
As syllable from sound. [2]

Table of Contents

List of tables

Number *Page*

Chapter 2

List of figures

Number *Page*

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Introduction

This thesis presents the stories of a series of investigations in which I and my mentors and collaborators studied the neural mechanisms of motor learning. Each of the chapters presents a different technical approach, ranging from biophysical modeling of individual nerve cells through analysis of large-scale optical imaging data sets, to address a basic scientific puzzle: how do organisms use the brain's capacity for learning to achieve fine control of their movements? This broad question fascinated me in particular because behavioral, anatomical and physiological research on a specific brain system – the cerebellum – have come together in recent years to provide what may be a critical mass of empirical data for understanding an important piece of the puzzle. By studying the mechanisms by which neurons and glial cells in the cerebellum, or "little brain," respond to inputs and generate output on their own, as well as acting together in circuits, I hope it will soon be possible to give a detailed account of the familiar experience of learning a new motor skill. Such an understanding would be practically useful for treating disorders of movement and for informing the design of artificial systems for controlling movement. It would also be an important step on the road to comprehending the function of brain circuits in learning and behavior in general.

The structure of the cerebellum is a remarkable example of intricate but repeating order in the brain. The quasi-crystalline geometry of the handful of neuron and glial cell types in the cerebellum was described in detail over 100 years ago by Ramon y Cajal [3] and continues to inspire beautiful anatomical studies [4]. Although the circuit diagram of the cerebellum can be described fairly completely in a succinct page of text (see circuit diagram, Chapter 1 Figure 1), the network is nevertheless vast: it contains more than half of all the neurons in the brain. The repetition of the same arrangement of cell types, with their axons and dendrites ordered in planes aligned with a local Cartesian coordinate system, suggests that the geometry of the cerebellum in some way serves its function and helps it to coordinate movement [5-7]. This same structure, and circuits with similar architecture in other brain regions, is remarkably conserved among vertebrate species, from fish to primates, and thus seems to be adaptable for a wide range of motor control tasks [8].

This thesis is organized in two halves treating, respectively, theoretical models and experimental tools. In the first two chapters, I use analytic theory and computational simulations to explain aspects of the timing of cerebellum dependent behaviors. Movement is a dynamical process, and proper coordination of movement requires representing and processing temporal information effectively. Experiments have shown the cerebellum to be necessary and sufficient for generating well-timed motor output in a specific experimentally controllable behavior known as delay eyeblink conditioning [9]. In this paradigm of motor learning, repeated pairing of a conditioned stimulus, such as a tone, with an unconditioned stimulus, such as a puff of air to the eye, results in a learned reflexive blink in response to the conditioned stimulus alone. Early theories of the function of the cerebellum by Marr, Albus, and Ito sought to explain the learned association of sensory inputs with appropriate outputs [10-12]. However, the widely used theoretical framework developed by these pioneers does not explain key features of the timing of behavior in eyeblink conditioning.

Behavioral studies of eyeblink conditioning show that time of peak eyelid closure following tone onset during a conditioned blink coincides with the delay between conditioned and unconditioned stimuli used during training [13]. Chapter 1 asks how the geometric organization of cerebellar cortex might create a distributed representation of time that could be used for generating temporally specific learned movements. Although distributed representations of temporal information have been discussed widely in recent theoretical [14, 15], modeling [16] and experimental studies [17-19], we take a novel approach by evaluating the neural representation of time in terms of its usefulness as a resource for statistical estimation. We apply this perspective to a simplified model of the recurrent neural network that receives most of the sensory input to the cerebellum and distributes it sparsely across the vast population of granule cells. Our results suggest a function for the anisotropic organization of the cerebellar cortex in creating cellular dynamics that are more appropriate as a representation of time than those created by alternative anatomical structures.

A second important timing aspect of cerebellum dependent eyeblink conditioning concerns the requirement for a causal relationship between conditioning stimuli. Training is apparently effective only in case the conditioned stimulus precedes the unconditioned stimulus by a sufficient (> 100 ms) interval [20, 21]. Training with conditioned and unconditioned stimuli that coincide in time, or with the reverse (acausal) temporal order, results in no expression of conditioned blinks. Chapter 2 presents a theory, based on the biophysical properties of so-called rebound currents present in neurons of the deep cerebellar nuclei, that accounts for this timing requirement within a “lock and key” framework. The computational modeling and theoretical analysis in this chapter were the result of a close collaboration with Daniel Wetmore and Mark Schnitzer, and our paper was published the Journal of Neurophysiology [22].

The second half of the thesis moves from the realm of theoretical modeling of cerebellar function to developing tools that could be used to test model predictions. Neuroscience has a long history of applying quantitative and statistical analysis tools to complex data sets. Interestingly, some of the tools that have been used for the analysis of multi-electrode recordings and behavioral data sets, and more recently in the growing field of systems-level functional imaging (fMRI and PET), overlap with machine learning and statistical concepts that have been proposed as models of information processing in neural circuits. There is thus a degree of continuity between the tasks of modeling brain activity, and characterizing it by experimental means. In the three experimental projects presented here, we addressed the new challenges that arise in analyzing data sets recorded by optical microscopy as compared with more established techniques of neurophysiology.

In collaboration with my experimentalist colleagues, I focused on the novel challenges of analyzing data sets recorded by optical microscopy in live animals. Chapter 3 presents a novel head-mounted miniature microscope that can be carried by a mouse while moving freely; the work in this chapter appeared in [23]. Our microscope, which was invented by a team in Mark Schnitzer’s laboratory lead by Benjie Flusberg and Erick Cocker, achieved the remarkable feat of recording the activity of neurons as well as the flow of individual red blood cells using a featherweight (1.1 g) microscope. To permit subjects to turn around during the experiment, the microscope incorporates a commutator that allows the fiber bundle carrying optical information to rotate with respect to the camera.

The flexibility of the miniaturized microscope created an analysis challenge in that movie data had to be computationally re-aligned to a reference frame fixed in the “mouse coordinate system” before extracting the cellular fluorescence signals. This challenge is analogous to a problem faced as a matter of course by our own visual systems while gazing at a fixed scene. The ongoing saccadic movement of the eyes leads to a constantly shifting gaze direction, and our brains combine the information obtained from each of these perspectives to produce the perception of a static world [24]. In analyzing the data from the miniaturized microscope we solved this challenge by a correlation-based image registration algorithm. The successful realignment of rotating movie data sets allowed us to glimpse increases in the complex spike activity of cerebellar Purkinje cells that were associated with locomotion. We also found evidence of increased correlation among complex spike trains during movement. These phenomena had previously been studied only by electrophysiology [25], which could not reveal the pattern of correlations within complete local networks of Purkinje neurons.

Our experience with the miniaturized microscope showed that Ca^{2+}-imaging could provide a new window on the modulation of complex spiking during locomotion. However, the relatively coarse spatial resolution (~2.8-3.9 μm in the lateral dimensions) of epi-fluorescence imaging through the fiberscope, along with the artifacts due to movement that remained even after our best efforts at data processing, limited the accuracy with which we could study complex spiking by this technique. To overcome these limitations, we developed a different imaging modality and created a novel method of data analysis to study the cellular correlates of locomotion with greater precision. Our experimental strategy, which was devised and implemented by Axel Nimmerjahn, revived a technique first developed for studies of insect behavior [26, 27]. We adapted the concept of an animal atop a movable ball to allow two-photon imaging in head-restrained but freely running mice positioned on an exercise wheel [28, 29]. The wheel allowed the mouse to run at will, and measurement of the wheel's rotation provided a temporally precise readout of locomotor behavior that we could correlate with physiological data. With this experimental approach we could record the activity in both neurons and glial cells during locomotion.

Traditional recording techniques report electrical activity in brain cells, and have thus led neuroscientists to neglect the study of cells lacking electrical excitability. This neglect is particularly acute in the case of glial cells, traditionally considered exclusively as supportive of information processing by their neuronal neighbors. The advent of Ca^{2+}-imaging techniques *in vitro* [30] and *in vivo* [31-33] has revealed active dynamical activity in astrocytes evolving over the seconds time scale. Such activity may be spatially organized and responsive to sensory input in a manner very much analogous to neuronal information processing [28, 34]. In Chapter 4, we set out to apply our head-fixed two-photon microscopy technique to study the Ca^{2+} activity in Bergmann glia, the astrocytes of the cerebellar cortex, in awake and behaving mice.

Our studies revealed a rich variety of dynamical activity over a range of spatial and temporal scales in Bergmann glial processes. We classified the transient increases in Ca^{2+} we observed in three categories, one of which was associated with the onset of running and involved activation in the processes of hundreds of glial cells over macroscopic (at least ~500 μm in extent) regions. This form of glial excitation[1] mimics in its initial timecourse the dynamics of blood flow increases [36], suggesting a possible link between intracellular Ca^{2+} transients in Bergmann glia and hemodynamic correlates of locomotion. The other two forms of Ca^{2+} activation we observed were not obviously related to locomotor behavior and occurred at very low spontaneous rates in both isoflurane anesthetized and awake subjects. The use of Ca^{2+}-imaging in behaving rodents should allow in the coming years more detailed investigations of the mechanisms and functional roles of these forms of cellular activation in the cerebellum.

Besides signatures of glial activity, our data sets recorded by two-photon microscopy in head-fixed animals also contained fluorescence transients arising from up to 100 dendrites of individual Purkinje cells. Such data sets offer a unique opportunity to simultaneously observe the intracellular dynamics within a nearly complete local network. However, extracting physiological signals from raw movie data sets poses a considerable challenge for data analysis due to the need to find the locations of cells within a high-dimensional space of movie pixels and subsequently estimate the times of events, such as action potentials, based on observations of Ca^{2+} dependent fluorescence. Whereas multi-electrode recording techniques [37] motivated the development and characterization of automated and semi-automated computational tools for solving the "spike sorting" problem [38], the analogous challenges for optical microscopy data are still immature. Recent studies introduced heuristic and semi-

---

[1] We dubbed the macroscopic form of excitation "flares"; spontaneous transients in individual processes are called "sparkles," and medium-range, stereotyped events involving tens of cells over a range of ~50 μm are called "bursts." We arrived at these monikers after a process of deliberation which generated several other worthy candidates. A particularly fitting alternate name for the "flares" is "glister," a noun meaning "a glistering; a bright light, brilliance, luster" [35]. This word derives from an obsolete English word, "glise," meaning "to glitter, shine." Its similarity to "glia" is entirely fortuitous.

automated approaches for extracting and analyzing neuronal activity signals in Ca^{2+}-imaging data sets [39, 40]. However, there remains a need for automated procedures based on the particular statistical characteristics of physiological signals recorded by optical microscopy for solving the "cell sorting" problem in large scale imaging data sets.

Chapter 5 introduces an approach to cell sorting based on independent component analysis (ICA) [41, 42]. We used a spatio-temporal version of ICA to take advantage of all available statistical information in movie data sets [43]. We created simulated fluorescence imaging data with signal and noise characteristics mimicking our cerebellar data sets. Using these simulations, we compared the performance of ICA based cell sorting with a more conventional analysis based on regions of interest (ROI). Our studies characterized the limitations and capabilities of our automated analysis procedures, including the sensitivity to noise and cross-talk among signals from neighboring cells. Automated analysis allowed us to efficiently extract high fidelity estimates of the true cellular signals in artificial data sets with far larger fields of view than could be analyzed by ROI analysis. We further tested cell sorting by using it to analyze optical imaging data recorded in the cerebellum of head-fixed, awake behaving mice.

Our results using automated cell sorting provided a novel view of the organization of "microzones" in the olivo-cerebellar system [25, 44, 45]. Electrophysiological recordings over the past three decades revealed that Purkinje cell complex spikes are weakly synchronized among pairs of cells within rostro-caudal strips, or microzones. One hypothesis proposed that this organization is modifiable during motor behavior, with individual cells shifting in their associations depending on task demands [46]. A different research program, making use of anatomical tracing and receptive field mapping studies, proposed a hard-wired, static organization of the cerebellar microzones that could be used to flexibly control different body regions through the activity of groups of Purkinje cells [47]. Applying automated cell sorting to our Ca^{2+}-imaging data sets revealed no evidence in support of the "dynamic microzones" hypothesis. Instead, we found that groups of correlated Purkinje cells remain structurally fixed during periods of rest and active locomotion. By contrast, both the rate of complex spikes and the degree of correlation were modulated by behavior. Our observation that correlations increase among Purkinje cell complex spike trains during active behavior suggests a distinct information processing strategy in cerebellum as compared with neocortical circuits, where the opposite pattern of modulation of correlations was observed [48].

The biophysical modeling and optical imaging tools and studies described in this thesis represent an attempt to leverage recent advances in the technology of experimental neuroscience to understand the mechanisms of cerebellar motor learning through computational and theoretical techniques. I hope that the hypotheses we put forward will be proven wrong, and the sooner the better, as such proof will surely provide the insight needed to paint a more accurate picture of the workings of the cerebellar cortex. I look forward with excitement to witnessing and, I hope, contributing to these insights in the coming years through the exploitation of remarkable advances in the technology for seeing biological activity at the cellular level.

Chapter 1: Cerebellar network geometry can improve the representation of time for motor learning

ABSTRACT

Coordinated control of movement requires properly timed motor control signals. In many cases, such timing must be ensured by an internal clock within the central nervous system [49-51]. Despite decades of work studying specific mechanisms and general principles for temporal processing in neural circuits [14, 52-56], we still lack a framework that can explain the temporal precision of animal behavior in terms of specific patterns of connectivity. Here we study the mechanism by which well-timed motor control signals are generated in cerebellum-dependent classical eyeblink conditioning. By analyzing this form of motor learning in a statistical estimation framework, we derive a quantitative measure of the number of effective degrees of freedom for temporal processing by a neural network. The number of degrees of freedom is a statistical measure that reflects the richness of the network's temporal dynamics, as well as the noise level and the dynamic range of synaptic plasticity. We simulate the dyn amics of a simple model of cerebellar granule and Golgi cells and show that the striking geometric organization of cerebellar cortex may improve the representation of time by increasing the number of degrees of freedom for temporal processing relative to an isotropic network.

Classical conditioning of the eyelid reflex is a convenient experimental model of a learned motor reflex in which internally generated timing signals play a critical role. In this paradigm, subjects are trained with repeated presentations of a conditioned stimulus (CS), such as a tone or light, together with an unconditioned stimulus (US), such as an airpuff or electrical stimulation of the eyelid. After training with hundreds of such pairings, subjects learn to produce a blink conditioned response (CR) following presentation of the CS alone. When the relative time between the onset of the CS and delivery of the US is fixed during training, subjects learn to produce CRs that peak precisely at the time of the expected onset of the US (Figure 1A,B) [13, 57, 58]. Despite its simplicity and the accumulation of empirical knowledge about this form of motor learning, a basic open question remains how properly timed CRs may be produced by neural mechanisms in the cerebellum.

Experimental studies of classical eyeblink conditioning in rats and ferrets over the past three decades support the conclusion that neural circuits located within the cerebellum are necessary and sufficient for the timing of CRs [59]. Decerebrate ferrets can learn well-timed CRs even when the tone or light CS is replaced with direct electrical stimulation of the afferents to the cerebellum [60]. In these experiments the CS was a train of 15 electrical impulses delivered at 50 Hz and lasting 300 ms. Because each of the 15 pulses was identical, the electrical input to the cerebellum carried no information about the time since the onset of the CS. Furthermore, changes in the activity of sensory afferents responding to tone stimuli apparently also lack any fine temporal structure for guiding the time of the CR [61]. Whereas inputs to the cerebellum lack information about the time since the onset of the CS, cerebellar output signals carried by Purkinje cell axons (Figure 1C) reflect the precise timing of the CR [62]. The experimental evidence thus points to a mechanism within the cerebellar cortex for representing the time since the onset of stimuli in a manner that can be used to produce well-timed motor output signals. Such a mechanism likely involves the recurrent network of granule and Golgi cells which process sensory input to the cerebellum (Figure 1C) [16, 51].

We used statistical learning theory to describe the requirements for a neural mechanism for temporal processing in cerebellar motor learning. In the first part of this chapter, we derive the number of "degrees of freedom," d, carried by a particular neural representation. d depends on both the richness of neural activity patterns, as well as the variability of neural activity and constraints on the learning rule implemented by mechanisms such as synaptic plasticity. We show that d determines how well the representation can be used for motor learning. We then apply this measure to activity traces from a simulation designed to capture basic aspects of the network of cerebellar granule and Golgi cells. We find that the geometric organization of the cerebellum may enhance its ability to represent time and thus improve motor learning.

RESULTS

Motor learning is a statistical estimation problem

To produce well-timed motor output, as observed in cerebellar Purkinje cells [62], neural circuits must accurately represent time delays spanning some range. An analog clock provides a representation of time that can be read by an observer to infer the current time with a specific precision. If the clock has a minute hand, an observer could use the clock to time a boiling egg. If a seconds hand is also available, the clock could be used to pull a shot of espresso (ideally for 20-25 s). The precision with which a clock or a neural circuit represents time determines its range of functional uses. Although it might be possible to use a clock with only a minutes hand in a timing task a few seconds, the accuracy of performance will be limited by the representation of time.

We developed a statistical learning framework to describe the function of the cerebellar circuit in motor learning tasks such as eyeblink conditioning. The goal of motor learning is to produce a specific time-varying motor control signal, $P(t)$. The output of the cerebellum, carried by axons of Purkinje cells, should approximate the desired motor control signal. We thus denote the activity of Purkinje cells by $\hat{P}(t)$ to emphasize their function as statistical estimators of the correct output at each point in time.

The cerebellum performs this estimation by processing information about sensory input arriving via mossy fiber afferents (Figure 1C). After training, the network implements a mapping from mossy fiber (sensory) inputs to Purkinje cell (motor) outputs. This mapping comprises two transformations. In the first stage, the network of granule and Golgi cells recodes sensory afferent signals into a set of activity patterns distributed across the population of granule cells, $\{g_i(t)\}$ (Figure 2A). Because of the enormous number of granule cells, this stage of processing effectively expands the sensory information into a representation in terms of a high-dimensional basis set. Second, the output of thousands of granule cells converges at individual Purkinje cells (Figure 2B). Each granule cell parallel fiber makes a synaptic connection with a weight that can be adjusted during training and which determine the map from inputs to outputs. After training, Purkinje cells map their inputs to specific motor control signals (Figure 2C), which are then relayed out of the cerebellum. The aim of learning is to adjust this mapping so that it approximates the desired motor program: $\hat{P}(t) \approx P(t)$.

Linear regression provides a simple statistical framework for analyzing the transformation of granule cell sensory information to Purkinje cell motor output. In fact, the cerebellum has often been described as performing a linear transformation of its inputs to produce appropriate output [63]. In the case of binary output, this is the computational paradigm of a perceptron [64-66]. The output motor control signal at time t is thus a linear combination of the inputs to the Purkinje cell:

$$\hat{P}(t) = \sum_i J_i g_i(t)$$

Here, a Purkinje cell's activity, $\hat{P}(t)$, depends on the instantaneous activity of all the pre-synaptic granule cells, $\{g_i(t)\}$, as well as the parallel fiber to Purkinje cell synaptic weights, $\{J_i\}$. For simplicity, we will restrict our attention to the output of one Purkinje cell. Our model thus concerns the network of ~10^5 granule cells and ~100 Golgi cells which shape each Purkinje cell's input. The goal of learning is to minimize the mean square error of the output control signal,

$$\varepsilon[J_i] = \int dt[P(t) - \hat{P}(t)]^2.$$

This error function defines a quadratic landscape for learning (Figure3B) [67]. The landscape is a function of the set of synaptic weights, $\{J_i\}$, which number in the thousands for each Purkinje cell. We can describe the geometry of the error landscape in terms of the granule cell patterns of activity by using a discrete representation in which each signal is determined by its value at a set of time points, $\{t_j\}$: $\mathbf{P}_j = P(t_j)$ and $\mathbf{G}_{ji} = g_i(t_j)$. The error function is thus $\varepsilon[\mathbf{J}] = \|\mathbf{P} - \mathbf{GJ}\|^2 = (\mathbf{J} - \mathbf{J}^*)^T \mathbf{H}(\mathbf{J} - \mathbf{J}^*)$, where we define the curvature matrix, $\mathbf{H} = \mathbf{G}^T\mathbf{G}$. Within this high-dimensional space, the landscape has a unique minimum at the optimal configuration of synaptic weights for a given desired motor output pattern: $J_i^* = [\mathbf{H}^{-1}\mathbf{G}^T\mathbf{P}]_i$. The Purkinje cell's task is to find the error minimum by some suitable process of updating the synaptic weights (Figure 3A); for example, backpropagation provides a biologically plausible learning rule for descending the gradient of the error landscape [68]. The weights should be adjusted during training in response to sensory input in a manner consistent with biological processes, such as long-term plasticity [10, 69]. However, such processes may not be capable of achieving the error minimum, $J_i = J_i^*$.

A main reason the network may not be able to exactly reproduce the desired motor output is variability in the activity of granule cells from trial to trial. Although the responses of granule cells are remarkably reliable [70], there will inevitably be some fluctuations in their response to similar patterns of sensory input. This variability introduces a noise term in the error landscape:

$$\varepsilon_{noise}[\mathbf{J}] = \|\mathbf{P} - (\mathbf{G} + \eta)\mathbf{J}\|^2 = (\mathbf{J} - \mathbf{J}^*)^T \mathbf{H}(\mathbf{J} - \mathbf{J}^*) + \sigma_{noise}^2 \|\mathbf{J}\|^2$$

where $\eta_{ij} = \eta_i(t_j)$ is a stochastic term representing trial-to-trial fluctuations, and we assumed independent noise in each cell, $\langle \eta_{ij} \rangle = 0$ and $\langle \eta^T \eta \rangle_{ik} = \sigma^2_{noise} \delta_{ik}$. The new term in the error function that is proportional to the noise variance penalizes large synaptic weights. The reason for this penalty is that using very large values for the J_i will amplify the fluctuations in granule cell activity on each trial and thus degrade the accuracy of Purkinje cell output. This noise-induced penalty balances the first term in the error function, which is minimized when $\mathbf{J} = \mathbf{J}^*$. These two terms are analogous to the bias and variance contributions to the statistical error in multi-dimensional regression [71].

The biological mechanisms of synaptic plasticity provide a second important motivation for introducing a penalty term in the error function. Without such a term, the synaptic weights may be required to vary over a broad range in order to change achieve the desired motor output pattern. However, plasticity processes such as long-term potentiation and depression can only vary the magnitude of synaptic weights over a limited dynamic range (Figure 4A). A quadratic penalty term in the error function like that introduced due to noise in the above equation is one way of solving the original optimization problem, $\min_{\{J_i\}} \varepsilon[\mathbf{J}]$, subject to a constraint on the synaptic weights: $\frac{1}{N}\sum_{i=1}^{N} J_i^2 \le J_{max}$. This constraint is imposed by adding a term, $\sigma^2_{syn} \| \mathbf{J} \|^2$, to the error function, where σ^2_{syn} is a Lagrange multiplier and is monotonically related to the constraint, J_{max}. This type of optimization is known as ridge regression, and it is one way of regularizing the original learning problem. Because of the formal similarity between the effects of noise and the constraint on the dynamic range of synaptic plasticity, we can lump these two effects together into a quadratic penalty term with total weight $1/\Lambda^2 = \sigma^2_{noise} + \sigma^2_{syn}$. With this penalty, the optimal synaptic weight configuration is given by $\mathbf{J}^{*,ridge} = (\mathbf{H} + \mathbf{Id}/\Lambda^2)^{-1}\mathbf{G}^T\mathbf{P}$ (Figure 4A).

The net effect of constraints on the synaptic weights is to prevent the network from exactly reproducing the desired motor output. However, the severity of this limitation depends on the statistics of granule cell activity, and in particular on correlations among the population of granule cells. For example, if all granule cells produced identical patterns of activity, motor output would necessarily be proportional to this pattern regardless of the configuration of synaptic weights. This would be the case in networks with line-attractor dynamics, in which all network cells quickly relax to a one-dimensional subspace [72, 73]. Because of the enormous number of granule cells, even slight differences in their activity patterns may be exploited to accurately fit a wide range of motor output signals. However, constraints on the dynamic range of synaptic plasticity and granule cell noise compete with this goal. The residual error of the motor output, $\varepsilon^*[\Lambda;\mathbf{G}] = \varepsilon[\mathbf{J}^{*,ridge}]$, is determined both by the constraint, Λ, as well as the representation of temporal information in the granule cells, $\mathbf{G}$.

The geometric picture of the error landscape as a quadratic function of the synaptic weights provides a simple interpretation of the residual motor error (Figure 4A). The matrix, $\mathbf{H}$, determined by the covariance of the granule cell activity traces, sets the curvature of the error landscape along each dimension of synaptic weight space [67]. Eigenvectors of $\mathbf{H}$ with large eigenvalue correspond to dimensions with steeply curved error landscape, while small eigenvalues correspond to dimensions with shallow error curvature. Dimensions with very shallow curvature pose a challenge for learning because the synaptic weights must be varied over a large range to significantly reduce the error. By contrast, steeply curved dimensions may be easily optimized by small adjustments of the synaptic weights. A more detailed mathematical description of this geometric picture of the error is provided in Appendix A.

The final accuracy of motor output after learning thus depends on the number of dimensions, d, of synaptic weight space for which the error landscape's curvature is steeper than a cutoff, $1/\Lambda$. The cutoff is inversely related to the dynamic range of synaptic plasticity and directly related to the noise power. This concept can be quantified for ridge regression in terms of the granule cell activity patterns (Figure 4B):

$$d = \mathrm{Tr}[\mathbf{G}(\mathbf{H} + \mathbf{Id}/\Lambda^2)^{-1}\mathbf{G}^T] = \sum_{i=1}^{N} 1/[1 + (S_i\Lambda)^{-2}].$$

Here, $\{S_i^2\}$ are the eigenvalues of $\mathbf{H}$. d corresponds to the number of effective "degrees of freedom" in statistical regression, and is a measure of the complexity of the range of motor output signals that can be learned given the representation of time in the granule cell network [74].

The steepness of the error landscape was related to the speed of learning in [67]. By contrast, we consider the effect of the error curvature on the limit of motor accuracy that can be achieved in the presence of noise and using limited synaptic plasticity.

Effect of network geometry on temporal processing

A prominent feature of the cerebellar cortex is the regular, quasi-crystalline arrangement of the granule cell parallel fibers [7]. Parallel fibers extend several millimeters along the cerebellar folia, making contact with the perpendicularly oriented Purkinje cell dendritic arbors (Figure 5A). Unlike neocortical brain areas, the cerebellar cortex is strongly anisotropic, with distinct cellular structures aligned along each of the three local dimensions. We used computational simulations to test the effect of this geometric organization on the temporal dynamics of granule cells, and hence on their ability to represent time for motor learning.

Our model includes recurrent synaptic connections between cerebellar granule and Golgi cells (Figure 5B). The length of parallel fibers, and the extent of Golgi cell dendritic arbors, together determined the range over which each granule cell could contact a Golgi cell [75]. Within this range, synaptic connections were chosen randomly. For comparison, we also created a network model with "isotropic" synaptic connections. In this network, granule cells contacted neighboring cells in each direction with equal likelihood. Cellular dynamics were simulated using a simple linear rate model [76] (see Appendix B for detailed methods). Although relatively simple in their architecture, both of the network models we studied incorporate important geometric constraints that have so far been neglected in the bulk of theoretical work on randomly connected, all-to-all neural networks [77-79].

A key feature of the synaptic weight matrix of the granule and Golgi cell network is that it is bipartite. That is, granule cells synapse on Golgi cells but not on other granule cells, and likewise Golgi cells do not synapse onto other Golgi cells. This structure greatly reduces the dimensionality of the dynamical activity induced in the network by the recurrent connections. In a bipartite network with N_{Go} Golgi cells and N_{gr} granule cells, where $N_{gr} \gg N_{Go}$, the total number of dynamical modes corresponding to non-zero eigenvalues of the synaptic weight matrix is no more than N_{Go} (see Appendix B). Intuitively, each granule cell can only interact with other granule cells by first influencing the activity of a Golgi cell. Since each Golgi cell diverges in its synaptic connections onto hundreds of granule cells, each of these granule cells effectively experiences the same dynamic input.

We examined the transient activity in granule cells following the onset of a stimulus, which induced an immediate increase in external input to the network. We designed the desired motor output signal to reflect the expected activity of a Purkinje cell involved in a conditioned eyeblink response. Following training, such cells should exhibit a timed pause in spiking at the moment of the expected US [22]. Details about the simulation methods are given in Appendix B.

If no constraint is imposed on the dynamic range of synaptic plasticity, both the geometric and the isotropic model networks are capable of learning to accurately generate well-timed motor output signals (Figure 6A). However, as noted above, this accuracy is bought at the expense of using large changes in synaptic weights to compensate for correlations in the granule cell activity traces. Such correlations would make it hard to learn accurate motor responses with a limited range of synaptic plasticity. In fact, we found that only the geometrically organized network could accurately learn the motor output using constrained synaptic weights (Figure 6B). To assess the general significance of this result, we varied the parameter, Λ, controlling the dynamic range of plasticity over several orders of magnitude. The results showed that the geometric network achieved low residual motor error over a broad parameter range in which the isotropic network failed to learn accurate responses (Figure 6C). These results show that network geometry can have a profound effect on the representation of time for learning.

To assess the impact of network connectivity in greater detail, we systematically varied the range of granule cell projections along (L_x) and across (L_y) the cerebellar folium. These two parameters defined a phase space for network geometry, with isotropic networks lying along the line $L_x = L_y$. For each value of (L_x, L_y) we created a model network and simulated its dynamics in response to a step change in input. The temporal pattern of activity in the network cells was then used to calculate the effective number of degrees of freedom, d (see above and Appendix A). Figure 7A shows the map of these simulation results as a function of the geometric parameters for a network with relatively weak synaptic connections. We found that as we increased the strength of all synapses by one (Figure 7B) or two (Figure 7C) orders of magnitude, while keeping the geometric pattern of synaptic projections the same, the number of degrees of freedom increased correspondingly. However, when the synaptic weights were increased too much the network became unstable, with some of the normal modes growing exponentially over time (Figure 7C).

We hypothesized that d depends on the diversity of oscillations in granule cell activity. Such oscillations are created by recurrent feedback, and thus depend on the geometry of the network. We thus examined the set of frequencies of the network normal modes for different geometric configurations. Figure 7D shows the cumulative distribution of mode frequencies for two isotropic networks, as well as two anisotropic networks. The networks with isotropic connections had most of their oscillation frequencies distributed within a narrow range. By contrast, anisotropic networks had a broader range of oscillation frequencies and a more even distribution of those frequencies within that range. However, in the limit of a very anisotropic network $(L_y = L_x / 8)$, the normal mode frequencies broke into discrete groups (red curve in Figure 7D). This is because such a network is effectively one-dimensional, with each beam of granule cells effectively disconnected from neighboring beams. The degeneracy of mode frequencies in such a network reduced the effective degrees of freedom compared with a slightly less anisotropic network (optimal network shown in green curve).

CONCLUSION

In this chapter we developed a theoretical framework for analyzing the representation of time by a distributed neural circuit. We focus on cerebellum dependent classical eyeblink conditioning as a prototype for timed motor learning tasks. We relate learning well timed motor output signals to a statistical estimation problem. The theory of ridge regression provides mathematical tools for analyzing the ability of a network to accurately generate timed output signals by adjusting synaptic weights. The limited dynamic range of synaptic plasticity provides a key constraint on learning. We show that the pattern of activity in the granule cells can be evaluated in terms of the number of “degrees of freedom,” which measures the complexity of motor output patterns that can be learned.

We applied our theoretical framework to analyze simulations of a cerebellum-like network of granule and Golgi cells. By varying the relative spread of granule cell synaptic connections along each of the two dimensions of the cerebellar cortex, we evaluated the effect of geometry on motor learning.

APPENDIX A: GEOMETRY OF THE ERROR LANDSCAPE

In this appendix we calculate the residual motor error after training in a network with a constrained range of synaptic plasticity:

$$\begin{aligned}\varepsilon_{res}[\Lambda;\mathbf{G}] &= \left\|\mathbf{P}-\hat{\mathbf{P}}\right\|^2 = \left\|\mathbf{P}-\mathbf{G}\mathbf{J}^{*,ridge}\right\|^2 \\ &= \left\|\mathbf{P}-\mathbf{G}(\mathbf{H}+\mathbf{Id}/\Lambda^2)^{-1}\mathbf{G}^T\mathbf{P}\right\|^2\end{aligned}$$

We can simplify this expression by performing singular value decomposition on the granule cell activity matrix: $\mathbf{G}=\mathbf{U}\mathbf{S}\mathbf{V}^T$ where $\mathbf{U},\mathbf{V}$ are orthogonal matrices and $\mathbf{S}$ consists of a diagonal matrix concatenated with a matrix of zeros. Note that the eigenvalues decomposition of the curvature matrix is given by $\mathbf{H}=\mathbf{U}\mathbf{S}^2\mathbf{U}^T$, where $\mathbf{S}^2$ is a square, diagonal matrix containing the squares of the singular values of $\mathbf{G}$.

Rotating into the basis set of singular vectors, we have:

$$\begin{aligned}\varepsilon_{res}[\Lambda;\mathbf{G}] &= \left\|\mathbf{U}(\mathbf{Id}+\mathbf{S}^2\Lambda^2)^{-1}\mathbf{U}^T\mathbf{P}\right\|^2 \\ &= \mathbf{P}^T\mathbf{U}\left[\mathbf{Id}+\mathbf{S}^2\Lambda^2\right]^{-2}\mathbf{U}^T\mathbf{P} \\ &= \sum_{i=1}^{N}\left[\frac{\tilde{P}_i}{1+(S_i\Lambda)^2}\right]^2,\end{aligned}$$

where $\tilde{\mathbf{P}}=\mathbf{U}^T\mathbf{P}$.

From this expression we can see the significance of the number of degrees of freedom, $d=\mathrm{Tr}[\mathbf{G}(\mathbf{H}+\mathbf{Id}/\Lambda^2)^{-1}\mathbf{G}^T]=\sum_{i=1}^{N}1/[1+(S_i\Lambda)^{-2}]$. Steeply curved dimensions are those for which $S_i \gg 1/\Lambda$. Such dimensions do not contribute residual output error, since $(1+S_i^2\Lambda^2)^{-2} \ll 1$. On the other hand, dimensions with shallow curvature, $S_i \ll 1/\Lambda$, do contribute to the final output error. The number of such shallow dimensions thus effectively sets the number of dimensions of synaptic weight space that can be optimized during learning.

APPENDIX B: SIMULATION METHODS

Synaptic weight matrix

We created simplified models of the recurrent network of granule and Golgi cells to study the effect of network geometry. Each network included N_{Gr} granule cells and N_{Go} Golgi cells. The average number of Golgi cells making a synapse on each granule cell was 2, and the average number of Golgi cells receiving synapses from each granule cell was 5 [17]. The corresponding number of granule cells receiving and making synapses on each Golgi cell was determined from these parameters together with the ratio of the populations of the two cell types. The cells were assigned to evenly spaced locations on a square lattice. The distribution of synaptic weights was binary; that is, each pair of cells was either not connected (synaptic weight = 0), or else connected by an excitatory (synaptic weight W^{RO} for granule to Golgi synapses) or an inhibitory (weight W^{OR}) synapse. There were no granule-granule or Golgi-Golgi synapses.

The connectivity of simulated networks was controlled by two parameters, L_x and L_y. The probability of a synaptic connection from a Golgi cell located at (x_{Go}, y_{Go}) to a granule cell at (x_{Gr}, y_{Gr}) was given by a Gaussian distribution:

$$P_{Gr \to Go} = A_{GrGo} \exp\left[-(x_{Go} - x_{Gr})^2 / 2L_x^2 - (y_{Go} - y_{Gr})^2 / 2L_y^2 \right]$$

where the constant, A_{GoGr}, was determined by the total number of synapses. The pattern of granule to Golgi synaptic weights was determined by an isotropic Gaussian distribution, $P_{Go \to Gr} = A_{GoGr} \exp\left[-((x_{Go} - x_{Gr})^2 - (y_{Go} - y_{Gr})^2) / 2L_x L_y \right]$. We created synaptic weight matrices with these probability distributions using a binomial random number generator to choose whether each pair of cells was connected or not.

For Figure 6, we created an isotropic network with similar connectivity to the geometric network described above by setting $L_x' = L_y' = \sqrt{L_x L_y}$.

Simulated dynamics of the bipartite network

The activity of the cells evolved according to a linear firing model:

$$\tau \frac{d}{dt} \begin{bmatrix} v_{Gr} \\ v_{Go} \end{bmatrix} = \begin{bmatrix} 0 & W^{OR}\mathbf{S}^{OR} \\ W^{RO}\mathbf{S}^{RO} & 0 \end{bmatrix} \begin{bmatrix} v_{Gr} \\ v_{Go} \end{bmatrix} + \mathbf{I}(t).$$

Here, $\mathbf{S}^{RO}, \mathbf{S}^{OR}$ are matrices whose entries are 0 or 1, reflecting the presence or absence of a synaptic connection respectively and W^{OR}, W^{RO} are scalar parameters setting the strength of inhibitory and excitatory synapses. $\mathbf{I}(t)$ represents external input to the network. This input was provided to a randomly chosen subset of the network cells. At the start of a trial, the input was changed instantaneously to a new value reflecting the onset of the conditioned stimulus.

Modes of oscillation in a bipartite network

To determine the normal modes of the network dynamics (Figure 7D), we first created a $N_{Go} \times N_{Go}$ matrix describing the bi-synaptic connections from each Golgi cell to every other Golgi cell, via one or more granule cells. This matrix is given by:

$$W^{OO}\mathbf{S}^{OO} = (W^{OR}W^{RO})\mathbf{S}^{OR}\mathbf{S}^{RO} .$$

Note that the dynamics of Golgi cells can be expressed in a closed form, independent of the granule cells: $\tau^2 \frac{d^2}{dt^2} v_{Go} = W^{OO}\mathbf{S}^{OO} v_{Go} + W^{OR}\mathbf{S}^{OR}\mathbf{I}_{Gr}(t) + \tau \frac{d}{dt}\mathbf{I}_{Go}(t)$. The entries of the Golgi-Golgi connection matrix, $\mathbf{S}^{OO}$, correspond to the number of granule cells which connect each pair of Golgi cells, and this matrix is thus almost perfectly symmetrical: the number of granule cells which receive input from Golgi cell *i* and project to cell *j* is roughly equal to the number that receive input from *j* and project to *i*. This symmetry implies that the dynamics of granule and Golgi cells can be decomposed in terms of simple normal modes. We found the set, $\{\lambda_i\}$, of N_{Go} eigenvalues of the bi-synaptic matrix $W^{OO}\mathbf{S}^{OO}$. These eigenvalues are negative due to the symmetry of $\mathbf{S}^{OO}$ and the opposite signs of the two synaptic weight coefficients.

The network oscillation frequencies governing the dynamics of both the granule and Golgi cells are given by $\omega_i = \sqrt{-\lambda_i}$. To see this, note that the $N_{Gr} \times N_{Gr}$ granule-granule bi-synaptic connection matrix, $\mathbf{S}^{RR} = \mathbf{S}^{RO}\mathbf{S}^{OR}$, is closely related to the far smaller Golgi cell matrix. In fact, every eigenvector of $\mathbf{S}^{OO}$, for which $\mathbf{S}^{OO}v^i_{Go} = \lambda_i v^i_{Go}$, corresponds to an eigenvector, $v^i_{Gr} = \mathbf{S}^{RO}v^i_{Go}$, of the granule cell network: $\mathbf{S}^{RR}v^i_{Gr} = \mathbf{S}^{RR}(\mathbf{S}^{RO}v^i_{Go}) = \mathbf{S}^{RO}(\mathbf{S}^{OR}\mathbf{S}^{RO}v^i_{Go}) = \lambda_i(\mathbf{S}^{RO}v^i_{Go}) = \lambda_i v^i_{Gr}$. All the remaining eigenvectors of the granule cell network must have eigenvalues zero, and thus do not contribute to the dynamics.

Error landscape curvature and degrees of freedom

To compute the shape of the error landscape, we generated a set of simulated trajectories whose dynamics were determined by the network oscillation frequencies:

$$v(t_j) = \mathrm{Re}[\mathbf{UI}e^{i\omega_i t_j}] .$$

Here $\mathbf{I}$ is a random, complex valued vector determined by the difference between the initial condition and fixed point of the network dynamics. $\mathbf{U}$ is an orthogonal matrix whose columns are the network normal mode eigenvectors. The error landscape curvature was then given by

$$\mathbf{H}_{ik} = \sum_j v_i(t_j) v_k(t_j) .$$

We computed the network degrees of freedom using $\mathbf{H}$ together with a parameter, Λ, setting the dynamic range of synaptic plasticity:

$$d = \mathrm{Tr}[(\mathbf{H} + \Lambda \mathbf{Id})^{-1} \mathbf{H}] .$$

Simulated motor output

To simulate the output of a Purkinje cell in a trained subject, we first defined the desired output trajectory, $P(t)$, to be a box-car function reflecting a pause in activity starting after a fixed interestimulus interval (Figure 6A,B). We then calculated the parallel fiber synaptic weights using ridge regression, $\mathbf{J}^{*,ridge} = (\mathbf{H} + \mathbf{Id} / \Lambda^2)^{\mathbf{-1}} \mathbf{G}^T \mathbf{P}$. We used Matlab's "backslash" operator to compute the matrix inverse. The residual error was defined as the mean square difference between the desired output and the learned estimate; see the Appendix A.

FIGURES

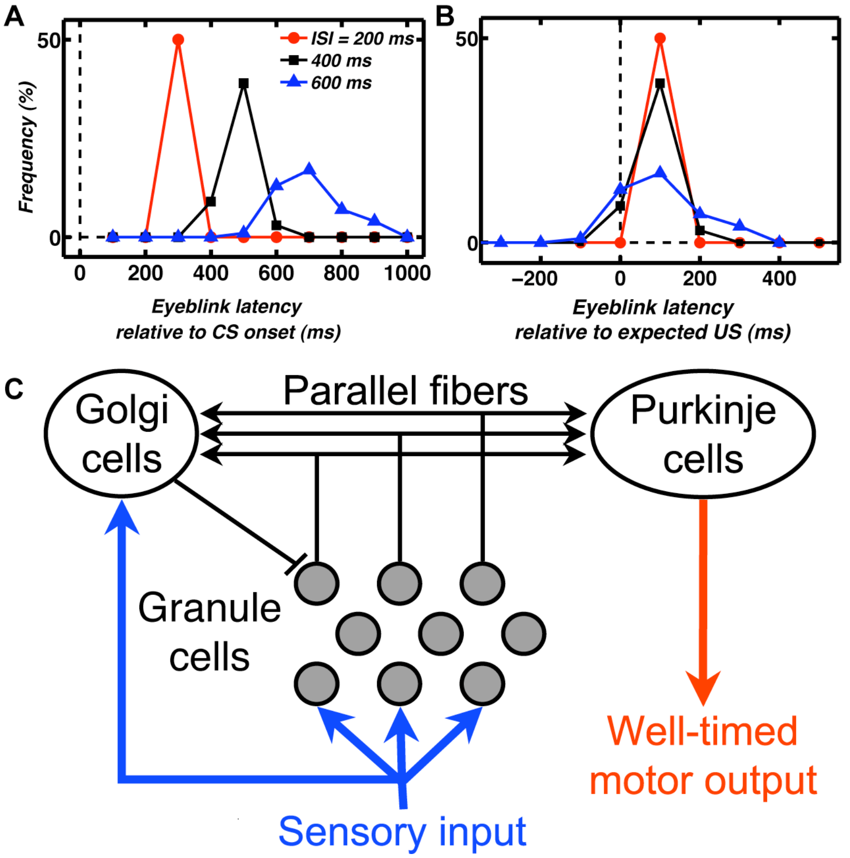

Figure 1. Cerebellar circuit diagram and timed motor output in eyeblink conditioning. (A,B) Conditioned eyeblinks in trained rabbits peak around the time of the expected onset of the US [13]. Experiments using three different interstimulus intervals show the capacity for generating multiple well-timed motor response patterns in response to the same sensory input. (C) Schematic diagram of the connections among cerebellar cell types in our model, together with sensory inputs. Granule (Gr) cells receive sensory input and send parallel fibers (PFs) to Purkinje (Pkj) and Golgi (Go) cells, as well as basket and stellate cells (not shown). Go cells make recurrent, inhibitory synapses onto Gr cells. The Pkj cells convey the output of the cerebellar cortex and drive expression of motor behaviors. Additional input arrives via climbing fiber afferents onto Pkj cells (not shown) and plays an important role during learning.

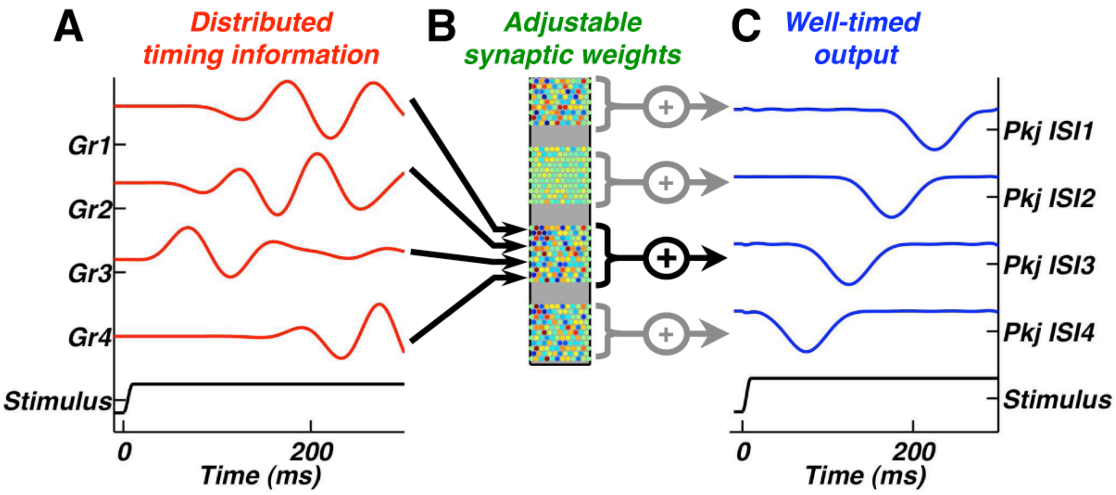

Figure 2. A distributed representation of time in the population of Gr cells would allow properly time Pkj cell responses.

(A) Sensory input (black trace) lead to irregular but reproducible temporal patterns in the activity of granule cells (red).

(B,C) Different patterns of plasticity at parallel fiber-Purkinje cell synapses (B) lead to differently timed pauses in Purkinje cell activity (C, blue traces).

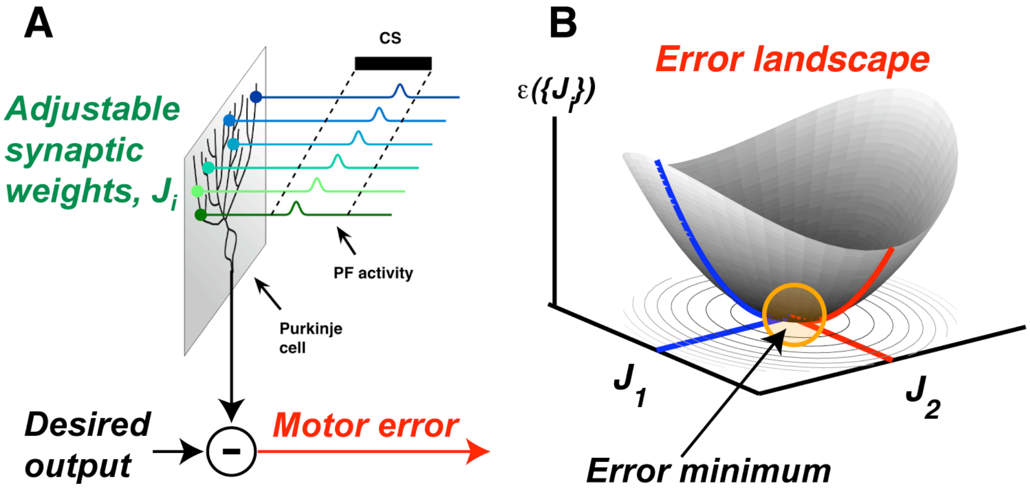

Figure 3. Statistical learning theory describes the requirements for learning timed responses.

(A) During learning, the weight of PF synapses on Pkj cells is adjusted to minimize motor error, defined as the mean square difference between actual and desired motor output signals.

(B) Learning corresponds to seeking the minimum of a parabolic "error landscape" in the space of synaptic weights. This landscape is used in statistics to estimate a function based on examples of its inputs and outputs.

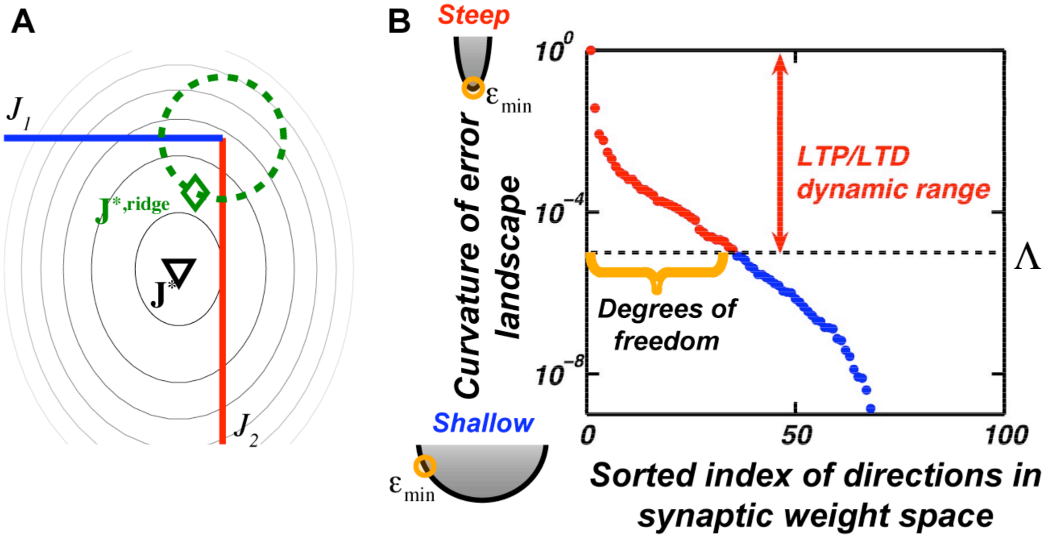

Figure 4. Dynamic range of synaptic plasticity limits learning.

(A) In this two-dimensional error landscape (gray contours), the error vanishes at the true minimum, $\mathbf{J}^*$ (triangle). The limited dynamic range of synaptic plasticity constrains $\mathbf{J}$ to the area within the green circle. The minimum error satisfying this constraint is $J^{*,ridge}$ (diamond).

(B) The curvature of the error landscape along the dimensions of synaptic weight space determines the network's ability to produce accurate motor output. The effective number of degrees of freedom is the number of orthogonal directions whose curvature is steeper than a minimum cutoff, Λ. The cutoff reflects by the dynamic range of plasticity.

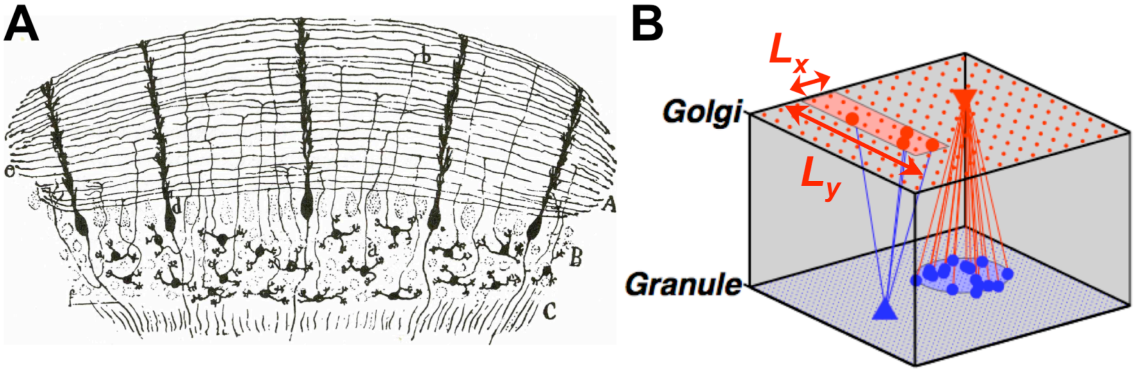

Figure 5. The length and orientation of parallel fibers determines anisotropy of the granule cell-Golgi cell network.

(A) A classic drawing of a cross-section of cerebellar cortex along the folium from Ramón y Cajal depicts the medial-lateral orientation of granule cell parallel fibers.

(B) Our network model of the recurrent network of granule (blue) and Golgi cells (red). Each granule cell may synapse on any Golgi cell whose dendrite overlaps with the range of its PF (red rectangle). Golgi cells connect to granule cells in an isotropic neighborhood determined (blue circle). Within these ranges, the particular synaptic connections are chosen randomly. The figure shows examples of connections from one granule cell to all its post-synaptic targets (blue lines), as well as from one Golgi cell to all its granule cell targets (red lines). The length scales, L_x and L_y, define the range of granule to Golgi connections, and their ratio sets network's anisotropy.

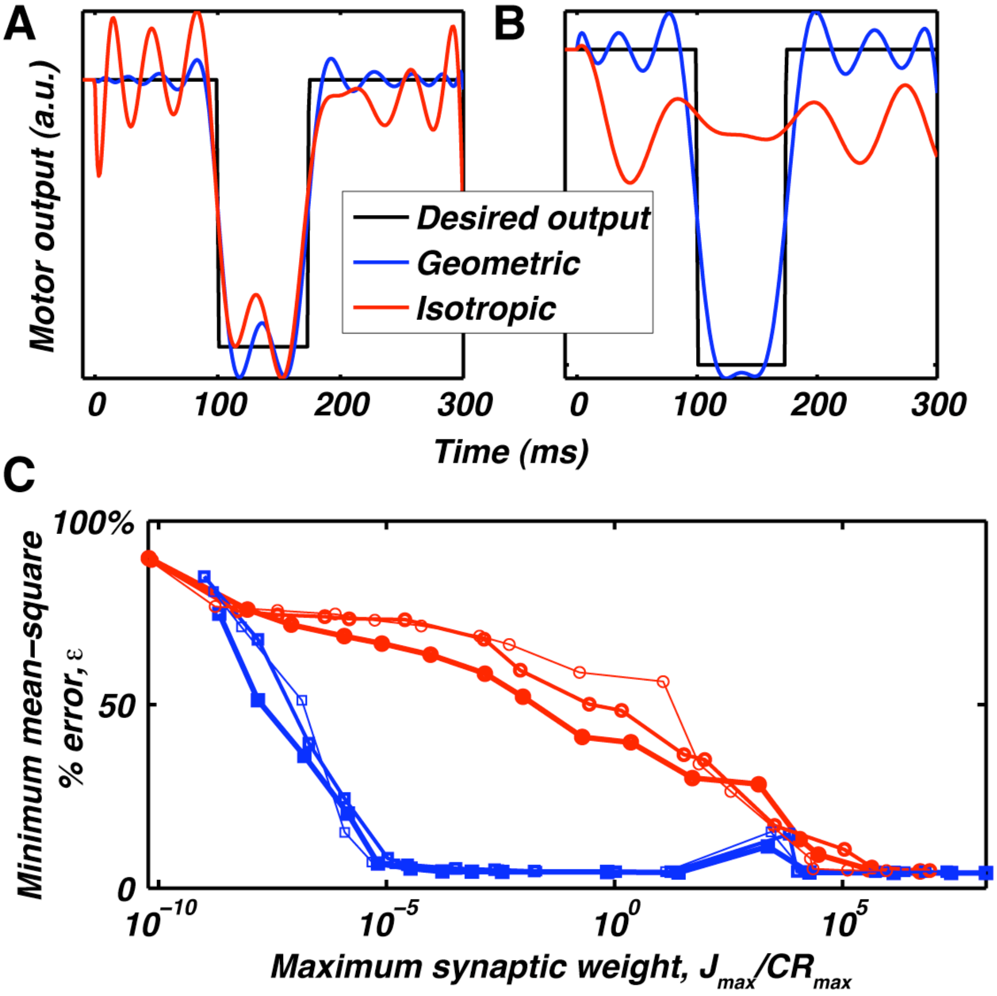

Figure 6. Motor output performance is improved by cerebellar network geometry.

(A) Output of a model Purkinje cell after training to match a desired output pattern (black trace). In this case there is no constraint imposed on the dynamic range of synaptic plasticity. Activity in granule cells from a model with cerebellum-like geometric connectivity (blue) or from a comparison network with isotropic structure (red) can be used to learn the desired output.

(B) In the presence of a realistic constraint limiting the maximum modulation of PF-Purkinje synaptic weights, $J_{\max}$, the geometric network accurately learns the desired output. The isotropic network, however, fails to learn the pattern.

(C) Motor learning is improved by geometric organization over an enormous range of the parameter, Λ, setting the dynamic range of plasticity. The isotropic network requires nearly 10 orders of magnitude greater dynamic range ($J_{\max}$) to achieve a reduction in motor error comparable with the geometrically organized network. The effect is also robust with respect to the ISI, with similar results for intervals of 30 (thinnest) to 100 (medium) or 170 ms (thickest).

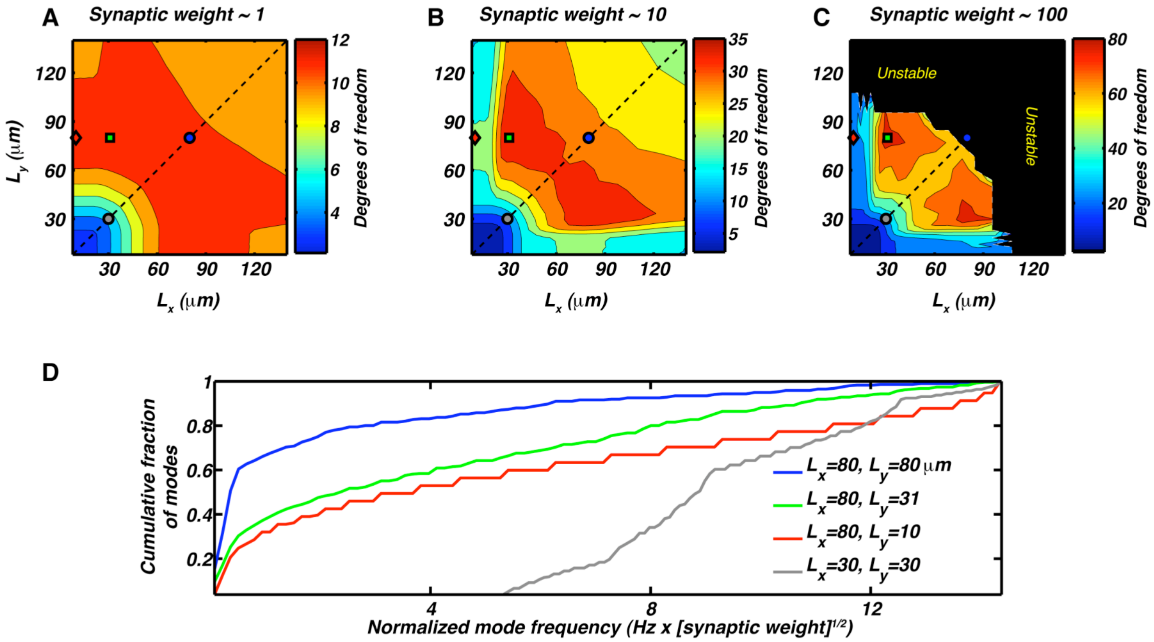

Figure 7. Anistropic network geometry can improve temporal processing for motor learning.

(A-C) Contour plots show the number of degrees of freedom for learning as a function of the length scales, L_x and L_y, for synaptic connections from Gr to Go cells. Dashed black line denotes isotropic models $(L_x = L_y)$. All synaptic weights (in arbitrary units) increase from 1 (A), to 10 (B) and 100 (C). Note that increasing synaptic weights 100 times leads to a 10-fold increase in the degrees of freedom, but also results in unstable dynamics.

(D) Cumulative density of states as a function of normalized oscillation frequency for four different geometries, corresponding to the points marked in the phase diagrams in (A-C). The oscillation timescale scales with the square root of the synaptic weight. Note that the models with an optimal number of degrees of freedom (green line and squares) achieve a broad distribution of mode frequencies but minimal redundancy between neighboring modes. These models are moderately anisotropic, with $L_x / L_y \simeq 2.7$.

Chapter 2: Lock-and-key mechanisms of cerebellar memory recall based on rebound currents

ABSTRACT

A basic question for theories of learning and memory is whether neuronal plasticity suffices to guide proper memory recall. Alternatively, information processing that is additional to readout of stored memories might occur during recall. We formulate a 'lock-and-key' hypothesis regarding cerebellum-dependent motor memory in which successful learning shapes neural activity to match a temporal filter that prevents expression of stored but inappropriate motor responses. Thus, neuronal plasticity by itself is necessary but not sufficient to modify motor behavior.

We explored this idea through computational studies of two cerebellar behaviors and examined whether deep cerebellar and vestibular nuclei neurons can filter signals from Purkinje cells that would otherwise drive inappropriate motor responses. In eyeblink conditioning, reflex acquisition requires the conditioned stimulus (CS) to precede the unconditioned stimulus (US) by >100 ms. In our biophysical models of cerebellar nuclei neurons this requirement arises through the phenomenon of post-inhibitory rebound depolarization and matches longstanding behavioral data on conditioned reflex timing and reliability**.** Although CS-US intervals <100 ms may induce Purkinje cell plasticity, cerebellar nuclei neurons only drive conditioned responses if the CS-US training interval was >100 ms. This bound reflects the minimum time for deinactivation of rebound currents such as T-type Ca^{2+}. In vestibulo-ocular reflex adaptation, hyperpolarization-activated currents in vestibular nuclei neurons may underlie analogous dependence of adaptation magnitude on the timing of visual and vestibular stimuli. Thus, the proposed lock-and-key mechanisms link channel kinetics to recall performance and yield specific predictions of how perturbations

to rebound depolarization affect motor expression.

INTRODUCTION

Research to date on the biological mechanisms of long-term memory has focused primarily on candidate mechanisms for memory formation, such as neuronal plasticity. But to what degree are the phenomenological properties of memory determined by biological mechanisms of memory recall? Studies on recall mechanisms have concerned re-consolidation processes that accompany retrieval [80, 81], network attractor theories of associative memory [82, 83], and expression of learned reflexes [17, 84-86]. However, the electrophysiological dynamics that occur during recall might have an important role in shaping qualities such as memory reliability and generalization. Thus, a basic question is whether these dynamics function primarily as a readout mechanism for retrieving stored memories or also perform additional processing of the stored information. Pattern completion is one aspect of associative memory recall for which candidate biological mechanisms have been identified [87]. Nonetheless, the existing literature on recall has generally assumed that the electrophysiological dynamics of recall should facilitate effective readout, *i.e.* retrieving the appropriate memory in response to a stimulus. The possibility that some constraints on memory expression might also be enacted at recall has not been widely considered.

Recent work on cerebellar memory systems indicates there are multiple loci of neuronal plasticity and at least two different brain areas of memory storage with distinct induction kinetics [88-93]. According to two-stage models of cerebellar learning, the numerous synapses in cerebellar cortex support flexible and rapid acquisition of new associations, while subsequent plasticity in the deep cerebellar or vestibular nuclei allows long lasting memory storage [94, 95]. Purkinje cells in the cerebellar cortex receive inputs from ~10^5 parallel fibers, and project outputs to the deep cerebellar and vestibular nuclei in a highly convergent manner, with each nuclear cell influenced indirectly by 10^7-10^8 parallel fibers [96] [94]. Given the vast number of potential network states in the cerebellar cortex, a rich set of training experiences might lead to network states that encode undesirable or inappropriate movements. The plausibility of this occurring is indicated by behavioral and computational studies that suggest the distribution of synaptic plasticity levels might evolve in a complex manner throughout learning experience, rather than purely reversing course during extinction or relearning [97, 98]. An example of an undesirable motor response is one executed in response to sensory cues that are reliably associated with rewarding or aversive stimuli but that arrive too late to be predictive of an appropriate motor action. Are there memory recall mechanisms that selectively prevent the expression of inappropriate motor responses, despite significant induction of synaptic plasticity? Or does plasticity induction always suffice to modify cerebellar-mediated motor behavior? At least for some forms of associative motor learning mediated by non-cerebellar memory systems, it has been shown that associative memory storage by itself can be insufficient to modify behavior [99].

To explore these issues, we formulated a ‘lock-and-key’ hypothesis stating that the induction of plasticity is necessary but not sufficient to modify motor behavior. There is the additional requirement that plasticity must shape the dynamics of neural activity (the ‘key’) to match a temporal filter (the ‘lock’) that selectively precludes inappropriate motor responses to sensory stimuli. We examined this hypothesis in the context of two cerebellum-dependent behaviors, classical eyeblink conditioning [100] and adaptation of the vestibulo-ocular reflex (VOR) [95, 101], for which there exist longstanding, rich behavioral data sets [20, 102]. If our hypothesis is true, what biological mechanisms might serve as the lock for these two behaviors?

This paper focuses on rebound currents in the deep cerebellar nuclei (DCN) and medial vestibular nuclei (MVN) neurons as candidate lock mechanisms, because it is well established that these currents perform significant temporal transformations of hyperpolarizing inputs, such as those from cerebellar Purkinje cells believed to trigger learned movements. Rebound channels, such as low voltage-activated (T-type) and hyperpolarization-activated cation (h) channels, are expressed at sufficient density to generate robust post-inhibitory rebound depolarizations in DCN and MVN neurons, the output neurons of cerebellar circuits [103-107]. For both behaviors studied, rebound channel kinetics emerge as crucial determinants of the minimum allowable duration between a sensory cue and a trained motor response. If the delay between the cue and a well-timed response is less than the time needed to activate rebound channels fully from the neuronal resting potential, the magnitude of the learned response declines or vanishes, thereby enacting the lock mechanism. This proposal represents a direct link from channel kinetics to learning performance and yields specific predictions of how learning performance is affected by perturbations to the rebound process.`

In eyeblink conditioning, key aspects of the behavior that remain poorly understood concern stimulus timing. Training of a reliable reflex requires the conditioned stimulus (CS), such as a tone, to start at least ~100 ms prior to the unconditioned stimulus (US), such as an air puff to the eye (Figure 1A) [20, 21]. Even after averaging data over multiple subjects, there remains a steep dependence of reflex acquisition on the CS-US training interval (Figure 1B), with the expression of conditioned blinks falling sharply for intervals <100 ms [21, 108-111]. What is the mechanistic basis for this effect? Analogous, unexplained dependencies on stimulus timing have been reported for VOR adaptation, in which the magnitude of learned eye movements depends on the timing between pulsed visual and vestibular training stimuli [102].

According to current thinking in the field an important mechanism of memory formation is the long-term depression (LTD) of cerebellar parallel fiber (PF) to Purkinje cell synapses induced by synchronous activation of PF and climbing fiber (CF) inputs to Purkinje cells [11, 12, 112, 113]. In eyeblink conditioning, it is thought that PF and CF inputs respectively convey signals regarding the CS and the US [60, 114-117], and that LTD resulting from repeated CS-US pairings leads to a conditioned reflex to the CS alone. This is proposed to occur since LTD should diminish the efficacy of CS-driven input to Purkinje cells, allowing disinhibition of deep cerebellar nuclei (DCN) neurons that receive GABAergic Purkinje cell inputs and drive conditioned reflexes [12]. In VOR adaptation, CFs and PFs respectively convey visual and vestibular information, and LTD is proposed to allow adaptive increases in VOR amplitude by reducing the strength of PF inputs signaling ipsiversive head rotation [113]. Although other cerebellar plasticity mechanisms exist [88, 90, 91], multiple strains of mice with disrupted LTD show deficits in eyeblink conditioning and VOR adaptation [118-123]. Nonetheless, accounts of cerebellar-mediated learning based solely on LTD do not easily explain the full range of behavioral data [92, 93, 98, 124-126].

One issue concerns whether Purkinje cells purely inhibit motor responses. Purkinje cells might be partly excitatory in their net effect, due to post-inhibitory depolarization in their target DCN and MVN neurons [103-106, 127]. Another issue concerns the possible role in learning of long-term potentiation (LTP) at the PF-Purkinje cell synapse. LTP and LTD induction at this synapse are spike-timing dependent [128], with LTP induced by unpaired PF or asynchronous PF-CF input [129, 130]. Maximal LTD induction seems to occur for PF activity that slightly precedes CF activity by 50-100 ms, which likely reflects the kinetics of postsynaptic Ca^{2+} dynamics [131]. LTD induction can occur with either PF or CF activity occurring first, but delays of more than ~200 ms are ineffective with either ordering [130]. It has been suggested that disinhibition of cerebellar nuclei neurons and spike-timing dependent plasticity suffice to explain the requirement in eyeblink conditioning for the CS-US interval to be >100 ms [130]. However, this has never been demonstrated explicitly using either computational modeling or experimental manipulation of behavior. A main difficulty is that the empirically determined rules for LTD induction suggest LTD should occur at short CS-US intervals that do not lead to acquisition of conditioned reflexes in behavioral experiments [130]. Furthermore, the dependence of learning performance on the CS-US interval appears much steeper than that of spike-timing dependent plasticity at the PF-Purkinje cell synapse [108-111, 130]. Thus, the degree to which conditioned reflex acquisition is shaped by physiological mechanisms other than spike-timing dependent plasticity remains an important issue for experimental investigation.

Here, we consider the novel possibility that significant shaping of learned motor expression might occur through the electrophysiological mechanisms of memory recall. In our work LTP and LTD emerge as complementary processes, both of which are important for memory formation as well as for memory clearance. This contrasts with the common view of LTD and LTP as opposing processes, one allowing memory storage and the other clearance [126, 129, 132]. Because plasticity induction is spike-timing dependent, we begin by considering the timing of sensory driven activity in the PF axons of cerebellar granule cells. Using a series of electrical compartmental models of increasing complexity, we simulate responses of DCN and MVN cells to learned sensory cues. This allows us to validate quantitatively the data from our DCN cell simulations against the classic behavioral data on eyeblink conditioning [108-111], by comparing the percentage of trials with successful responses as found experimentally to data generated by our models.

Comparison of the VOR adaptation magnitude in our modeling to that in behavioral studies suggests post-inhibitory rebounds might play a role in multiple cerebellum-dependent behaviors. Based on the results of our biophysical models we provide an algorithmic description of the ‘lock-and-key’ mechanism as a temporal filter. Learning experience that successfully modifies motor behavior shapes neural activity to match this temporal filter. Unsuccessful training can yield comparable magnitudes of synaptic plasticity, but the resulting patterns of Purkinje cell activity do not trigger learned motor responses. We have organized the following sections so that readers who wish to omit the computational details may skip the

Methods section below without loss of logical continuity.

MATERIALS AND METHODS

General simulation procedures

We created compartmental models of DCN and MVN cells in the NEURON [133] and MATLAB software environments and set model parameters using empirically determined values whenever possible. Fortunately, much is known about DCN cells from *in vitro* studies. We found that values determined from measurements in DCN and MVN cells, rather than other cell types, facilitated consistency with behavioral data. The current balance equation describing the balance of capacitive and ionic currents, $C_m(dV/dt) = -\sum I_{ionic}$, was integrated over time using the MATLAB function ode45 for deterministic one-compartment simulations, an Euler method for one-compartment simulations with stochastic synaptic inputs, or NEURON's implicit Euler method for two-compartment simulations. In all simulations, time-steps were less than or equal to 0.1 ms and the membrane capacitance, C_m, was 1 µF/cm^2.

Voltage-dependent currents obeyed equations of the form $I = \bar{g}\, x^{\psi} y (V - V_{rev})$, where $\bar{g}$ is the maximum conductance and V_{rev} is the reversal potential. Activation variables, *x*, followed first-order kinetics defined by $dx/dt = \phi_x[\alpha_x(V)(1-x) - \beta_x(V)(x)]$, where α_x and β_x are forward and backward rates and $\phi_x = Q_{10}^{\Delta T/10^\circ C}$ is a temperature factor. Q_{10} was 1.4 for T-type current and 2.3 for all other conductances. ΔT is the difference between the physiological temperature of 37°C used for all simulations and the temperature at which experimental measurements of channel kinetics were made. Inactivation variables, *y*, obeyed analogous expressions. Steady state and relaxation time constants are given in terms of α_x and β_x: $x_\infty = \alpha_x/(\alpha_x + \beta_x)$ and $\tau_x = 1\big/\left[\phi_x(\alpha_x + \beta_x)\right]$.

The firing rates of Purkinje cells were modeled to be from a cerebellar network after behavioral training. Electrophysiological data from *in vivo* recordings were used to constrain background firing rates, $r_{Pkj,b}$ = 40 Hz (Berthier and Moore 1986; Kotani et al. 2006; Jirenhed et al. 2007), and the modulation of Purkinje cell firing rates due to learning-related cerebellar plasticity; low and high Purkinje cell spikes rates following depression and potentiation of parallel fiber inputs were $r_{Pkj,d}$ = 20 Hz and $r_{Pkj,p}$ = 100 Hz, respectively (Berthier and Moore 1986; Kotani et al. 2006; Jirenhed et al. 2007).

Given these basic constraints the average Purkinje cell spike rate, $R_{Pkj}(t)$, was determined by first convolving a smooth plasticity function, $S(\Delta t)$, whose argument is the relative delay between activity in parallel fibers and climbing fibers (Figure 2A), with a boxcar function, $U(t)$, representing a US of 10 ms duration. The result of this convolution was multiplied by a smooth function, $C(t)$, representing a CS of duration no less than a minimum interval, $t_{CS\min}$ = 50 ms.

$$S(\Delta t) = A_1 + A_2\left\{T[(\Delta t - t_{LTD})/\tau] - T[(\Delta t - t_{LTD-})/\tau]\right\},$$

$$U(t) = \begin{cases} 0, & t < t_{ISI} \\ 1, & t_{ISI} \le t \le t_{ISI} + 10\text{ms} \\ 0, & t > t_{ISI} + 10\text{ms} \end{cases},$$

$$C(t) = T[(t-\tau)/\tau] - T[\left(t - \tau - \max(t_{CS\min}, t_{ISI} + 10\text{ms})\right)/\tau],$$

$$R_{Pkj}(t) = r_{Pkj,b} + C(t)\int_{-\infty}^{\infty} S(t'-t)U(t')\,dt'.$$

where t_{ISI} is the interstimulus interval between the CS and US onset times. t_{LTD-} = –10 ms and t_{LTD} = 75 ms respectively set the minimum and maximum allowable delay between CS-driven parallel fiber and US-driven climbing fiber activity for induction of LTD, and τ = 10 ms is a characteristic transition time describing the smooth temporal evolution of neural dynamics. We used the smoothing function, $T[x]$, to ensure gradual changes in spike rates:

$$T[x] = \begin{cases} 0 & x < -1 \\ (1+\cos(\pi x))/2 & -1 \le x < 0 \\ 1 & x \ge 0 \end{cases}.$$

The constants, A_1 and A_2, were set such that the minimum and maximum Purkinje cell firing rates for a long ISI were $r_{Pkj,d}$ = 20 Hz and $r_{Pkj,p}$ = 100 Hz respectively (Table 1) (Berthier and Moore 1986; Kotani et al. 2006; Jirenhed et al. 2007). These plasticity rules led to gradual transitions between distinct average firing rate values over ~20 ms intervals, approximating the observed intervals over which Purkinje cells modulate their spiking rates during expression of motor learning [134, 135].

CS-driven modulation of the rate of mossy fiber spiking was also constrained by data from *in vivo* electrophysiological recordings and was expressed as:

$$R_{MF}(t) = r_{MF,b} + (r_{MF,CS} - r_{MF,b})C(t)$$

where $r_{MF,b}$ = 10 Hz is the background firing rate for mossy fibers and $r_{MF,CS}$ = 50 Hz is the spiking rate of mossy fibers during presentation of the conditioning stimulus [136, 137].

Model 1 – A single-compartment model of DCN neurons

We modeled a DCN cell with a single electrical compartment that included leak (I_L) and T-type Ca^{2+} (I_T) currents, as well as synaptic currents due to inputs from Purkinje cells ($I_{syn,Pkj}$) and mossy fibers ($I_{syn,MF}$). Membrane voltage dynamics were determined by time integration of the current balance equation: $C_m(dV/dt) = -I_T - I_L - I_{syn,Pkj} - I_{syn,MF}$. The passive current, $I_L = g_L(V - V_L)$, was an admixture of two components: a tonic mixed-cation current that is characteristic of DCN cells and has a –30 mV reversal potential [138], and a standard leak current with –75 mV reversal potential [105, 127]. Total leak conductance, g_L, and leak reversal potential, V_L, were determined by the DCN cell's resting potential of V_{rest} = –58 mV (Aizenman and Linden 1999; Llinas and Muhlethaler 1988), and the observed membrane time constant of ~12 ms (Jahnsen 1986a; Llinas and Muhlethaler 1988):

$$g_L = C_m / \tau_m - \sum_i g_i,$$

$$V_L = V_{rest} + \sum_i I_i / g_L + \sum_{syn} I_{syn} / g_L$$

where g_i and I_i are the steady-state conductance and current density at V_{rest} for each of the active conductances, and I_{syn} are the steady-state synaptic current densities determined by V_{rest} and the background rate of spiking of each type of synaptic input.

Parameter values for Purkinje and mossy fiber synaptic inputs were constrained by physiological measurements: $V_{syn,Pkj}$ = –75 mV [105, 127], $\tau_{syn,Pkj}$ = 14 ms [139], and $V_{syn,MF}$ = 0 mV [139] (Table 1). Glutamatergic synapses in the DCN have significant AMPA and NMDA components [139]. As a simplification, $\tau_{syn,MF}$ for mossy fibers was chosen to be 23 ms by weighting the AMPA and NMDA decay time constants at –60 mV by the measured relative amplitudes of AMPA and NMDA glutamatergic input [139]. T-current was the sole voltage-dependent current: $I_T = \bar{g}_T\, nl(V - V_{Ca})$, with V_{Ca} = 140 mV [140]. T-type kinetics were adapted from a model of the α1G T-type channel [141], which is highly expressed in the DCN [142]. Steady-state values of the gating variables, n_∞ and l_∞, were modified to fit measurements of T-type currents in DCN cells [143]. Measurements of T-type currents in DCN neurons were made at room temperature, so we used a Q_{10} of 1.4 for temperature adjustment to produce rebound depolarizations at 37°C [127]. After temperature adjustment these expressions were:

$$\tau_n = 0.287 + 0.0711 \cdot \exp(-V/15.8)$$

$$\tau_l = 5.96 + 0.00677 \cdot \exp(-V/7.85)$$

$$n_\infty = \left(1.00 + \exp\left(-\left(V + 42.0\right)/4.25\right)\right)^{-1}$$

$$l_\infty = \left(1.00 + \exp\left(\left(V + 63.0\right)/3.50\right)\right)^{-1},$$

where in this and all subsequent expressions numerical parameters with dimensions of time and voltage are expressed in units of ms and mV, respectively.

Unfortunately, several DCN cellular parameters could not be tightly constrained by biophysical data. T-type channels appear to be most dense in DCN cell distal dendrites [143]. Thus, estimation of total T-type conductance from somatic recordings is difficult. However, in initial simulations we identified a broad range of T-type conductance values over which rebound depolarizations occurred, indicating that the occurrence of post-inhibitory rebound is not highly sensitive to the value of the T-type conductance density. In all subsequent single-compartment simulations (Models 1 and 2) we chose $\bar{g}_T = 0.5$ mS/cm^2, within the middle of this identified range.

The weight of Purkinje cell synaptic input, $W_{syn,Pkj}$ = 0.2 mS/cm^2, was chosen to be near the middle of a range of values capable of inducing physiological 10-15 mV changes in membrane voltage when input firing rates were modulated. The mossy fiber synaptic conductance weight, $W_{syn,MF}$ = 4 μS/cm^2, was chosen such that mossy fiber input alone was insufficient to drive the cell to rebound. Without this stipulation there would be little dependence of learned responses on Purkinje cell input, contrary to experimental findings. In Model 1 all synaptic inputs were simulated in a deterministic fashion. Each synaptic conductance density, $g_{syn}(t)$, was determined by convolving the input spike rate with an exponential function of time constant τ_{syn} and amplitude W_{syn}, which represented the conductance response (Table 1).

Model 2 – A single-compartment model of dendritic Ca^{2+} spiking

Model 2 was the same as Model 1 but with two modifications. First, we added a high-voltage activated (HVA) Ca^{2+} current, so we could study how graded T-current mediated rebound depolarizations led to the initiation of HVA Ca^{2+} spikes. The conductance model for I_{HVA} was identical to that used in [144]. Second, membrane voltage dynamics were no longer deterministic. Instead, the time-varying synaptic conductance density was determined by the stochastic arrival of action potentials at times selected independently, as governed by Poisson statistics and the instantaneous spike rate. The number of Purkinje cell ($N_{syn,Pkj}$ = 50) and mossy fiber ($N_{syn,MF}$ = 10) inputs roughly matched the ratio, R, of GABAergic to glutamatergic synapses found in the DCN [145] (Table 1). The occurrence of a pre-synaptic spike on any of these independent individual inputs led to an instantaneous jump in synaptic conductance of amplitude $W_{syn,Pkj}/N_{syn,Pkj}$ or $W_{syn,MF}/N_{syn,MF}$, which then declined exponentially with time constant $\tau_{syn,Pkj}$ or $\tau_{syn,MF}$, respectively (Table 1).

Model 3 – A two-compartment model of DCN neurons

A DCN cell model with dendritic and somatic compartments was used to test the effect of rebound conductances on Na^{+} spiking output. The potentials of dendritic (V_d) and somatic (V_s) membranes were determined by the currents flowing in each compartment [146]:

$$C_m \frac{dV_d}{dt} = -I_T - I_L - I_{HVA} - I_{SK} - I_{syn,Pkj} - I_{syn,MF} - \frac{g_c(V_d - V_s)}{1-\rho}$$

$$C_m \frac{dV_s}{dt} = -I_T - I_L - I_{HVA} - I_{SK} - I_{Kv} - I_{Na} - \frac{g_c(V_s - V_d)}{\rho}$$

where the coupling between the two compartments was specified by the conductance between compartments, g_c, and the ratio of somatic membrane surface area to total cell surface area, ρ. The passive current, I_L, was determined independently in both compartments from V_m = –58 mV and the membrane time constant τ_m = 12 ms. The somatic compartment exhibited spontaneous Na^{+} spiking, so we determined I_L by setting the I_{Na} gating variables to $m_\infty(V_{rest})$ and $h_\infty(V_{rest})$, and the I_{Kv} gating variable to $n_\infty(V_{rest})$. Stochastically arriving synaptic currents entered the dendritic compartment. For computational ease, we modeled the spike rate of Purkinje cells using a simple formula that closely approximates the rate function used in Models 1 and 2:

$$
\begin{aligned}
R_{Pkj}(t) = r_{Pkj,b} + C(t)\{r_{Pkj,p} - r_{Pkj,b} + \\
\left(r_{Pkj,d} - r_{Pkj,p}\right)\left(T\left[\frac{t - t_{ISI} + t_{LTD}}{\tau + 10\ ms} - 1\right] - T\left[\frac{t - t_{ISI} + t_{LTD-}}{\tau + 10\ ms} - 1\right]\right)\}
\end{aligned}
$$

Active somatic currents were T-type Ca^{2+}, high-voltage activated Ca^{2+} (I_{HVA}; [143]), Ca^{2+}-activated K^{+} (I_{SK}; [138]), fast Na^{+} (I_{Na}), and delayed rectifying K^{+}(I_{Kv}). The simulation used the total Ca^{2+} current, I_{Ca}, to determine the internal Ca^{2+} concentration, which controlled the gating of I_{SK}. In addition to I_L, the dendrite had T-type currents [143], I_{HVA} (at half the density as in the soma; see [143]), I_{SK}, and synaptic input. Conductance models for I_{HVA}, I_{SK}, I_{Kv}, and I_{Ca} were identical to those used in [144], including parameter values. I_{Na} was based on the model of [144]. For Na^{+} channel activation, ψ= 3 with kinetics determined by

$$\alpha_{x,Na} = 0.182\left(V+31\right)\Big/\left(1-\exp\left(-\left(V+31\right)/9\right)\right)$$

$$\beta_{x,Na} = -0.124\left(V+31\right)\Big/\left(1-\exp\left(\left(V+31\right)/9\right)\right).$$

For Na^{+} channel inactivation, ψ= 1 and

$$\tau_{y,Na} = \left[\phi_{Na}\left(0.024\left(V+48\right)\Big/\left[1-\exp\left(-\left(V+48\right)/5\right)\right] - 0.0091\left(V+73\right)\Big/\left[1-\exp\left(\left(V+73\right)/5\right)\right]\right)\right]^{-1}$$

$$y_{\infty,Na} = \left(1+\exp\left(\left(V+64\right)/6.2\right)\right)^{-1}.$$

Conductance densities and the voltage dependence of I_{Na} gating were chosen to reproduce the observed tonic DCN cell firing rate of ~25 Hz [103, 138, 147] and spike width of ~1.5 ms [105, 107]. Reversal voltages were V_{Na} = 50 mV, V_K = –90 mV, and V_{Ca} =140 mV (Jahnsen 1986b; Mainen & Sejnowski 1996).

Coupling parameters, g_c = 0.53 μS/cm^2 and ρ = 0.05, were chosen to ensure each compartment had relatively independent dynamics while still permitting dendritic voltage deflections to affect somatic spiking. Synaptic weights were set by the same criteria as for the one-compartment models, with $W_{syn,Pkj}$ = 0.2 mS/cm^2 and $W_{syn,MF}$ = 4 µS/cm^2. As in Models 1 and 2, in the dendrite $\bar{g}_T$ = 0.5 mS/cm^2 was set near the middle of a broad range of values that allowed rebound depolarization. In the soma $\bar{g}_T$ = 5 µS/cm^2, reflecting the lower density of low-voltage activated Ca^{2+} channels in this compartment [143]. The densities of Ca^{2+}-activated K$^+$ conductance, $\bar{g}_{SK}$ = 32 µS/cm^2, and of high-voltage activated Ca^{2+} conductance in the soma, $\bar{g}_{HVA}$ = 0.3 mS/cm^2, and dendrite, $\bar{g}_{HVA}$ = 0.15 mS/cm^2, had scarcely any effect on the probability of HVA Ca^{2+} spiking in response to synaptic inputs across a broad range of conductance densities and were chosen to reduce the duration of the Ca^{2+} spike to physiologically realistic values [105, 127]. Hodgkin-Huxley conductances, $\bar{g}_{Na}$ = 144 mS/cm^2 and $\bar{g}_{Kv}$ = 56 mS/cm^2, were chosen to reproduce the experimental observation of spontaneous spiking in the soma at $V_{rest} = -58$ mV.

Phase plane analysis of rebound depolarizations

To study whether a memory recall mechanism based on rebound depolarization would be robust, we reduced Model 1 to a system of two dynamical degrees of freedom amenable to graphical phase plane analysis. This involved an approximation in which the T-type activation variable was set equal to its asymptotic value, $n = n_\infty(V)$, reducing the dynamical variables to only the T-type inactivation variable, (l), and the membrane voltage, (V). Because this approximation increased the membrane excitability, resulting in larger magnitude rebounds, we decreased the density of T-type Ca^{2+} channels to $\bar{g}_T = 0.3$ mS/cm^2 as a compensatory measure. The system's dynamical trajectories within the (V, l) phase plane could then be fruitfully studied by determination of the two nullclines, on which the time derivatives vanish:

$$\frac{dl}{dt} = 0 \quad \rightarrow \quad l = l_\infty(V)$$

$$\frac{dV}{dt} = 0 \quad \rightarrow \quad l = \frac{g_L(V - V_L) + \bar{I}_{syn,Pkj}(V) + \bar{I}_{syn,MF}(V)}{\bar{g}_T\, n_\infty(V) \cdot (V_{Ca} - V)} \qquad ,$$

where the Purkinje cell and mossy fiber synaptic input currents $\bar{I}_{syn}(V) = g_{syn}(V - V_{syn})$ represent the mean synaptic currents as determined from the synaptic weights and background firing rates. Both time derivatives vanish at the intersection point of the two nullclines, so this is a fixed-point of the dynamics. Fixed-points during the neuronal resting state (stage 1), the CS-US interstimulus interval up until t_{LTD} before the expected US onset (stage 2), and the remaining portion of the CS (stage 3), were found using the MATLAB function *fzero* to solve for the intersection of the nullclines. Linear stability analysis within a neighborhood of the resting (stage 1) fixed-point at $V = -58$ mV revealed that this fixed-point is stable for $\bar{g}_T <$ 1.28 mS/cm^2. The dynamical trajectories near this fixed-point exhibit damped oscillations for $\bar{g}_T \geq 0.20$ mS/cm^2. The density of T-type channels used in our studies (Table 1) results in a stable spiral fixed-point.

A map of rebound magnitudes in the phase plane (Figure 6C) was built by numerically integrating the equations of motion using MATLAB's Runge-Kutta initial-value differential equation solver, *ode45*. A series of evenly spaced initial points was chosen along the boundary lines of the phase plane, defined by $V = -72$ mV and $l = 0$, and trajectories were integrated forward in time using the current balance equation and the constant synaptic input values of stage 3. Integration proceeded until the trajectories reached the stage 3 fixed point. Each trajectory formed a contour (level-curve) on the phase plane map with the contour amplitude given by the maximum depolarization achieved on that trajectory.

To determine the voltage threshold curve for firing all-or-none Ca^{2+} spikes (Figure 6D), we added a high-voltage activated (HVA) Ca^{2+} conductance [144] to the phase plane treatment of Model 1 in which n relaxes instantaneously to its asymptotic value, $n = n_\infty(V)$ (Figure 6 panels A-C). This yielded a deterministic version of Model 2 that produced virtually the same trajectories as Model 1 over the voltage range, $V < -35$ mV, in which the HVA channels are largely closed (Figure 6D). As before, the instantaneous activation of T-type currents led to increased membrane excitability, for which we compensated by decreasing the density of T-type Ca^{2+} channels to $\bar{g}_T = 0.3$ mS/cm^2 and the density of HVA Ca^{2+} channels to $\bar{g}_{HVA} = 0.09$ mS/cm^2. We solved for the dynamical trajectories by integrating the equations of motion forward in time starting at a series of initial points distributed along two boundary lines of the phase plane, defined by $V = -72$ mV and $l = 0$. The trajectories fell into two classes depending on whether the T-current mediated rebound was of sufficient magnitude to cross the threshold for generation of a HVA Ca^{2+} spike.

Phase plane movies showing model trajectories

Movies of deterministic (Movies 1 and 2 using Model 1) and stochastic (Movie 3 using Model 2) voltage trajectories were created in MATLAB. As in the phase plane analysis of Figure 6, Movies 1 and 2 relied on the mathematical approximation of instantaneous relaxation of the T-channel activation variable to its asymptotic value. The motion of the V nullcline was determined by solving the equation $dV/dt = 0$ for l, using the steady state values of the synaptic conductances, g_{syn}, that would be attained given constant Purkinje cell and mossy fiber spiking at rates equal to their instantaneous values. Numerical integration of the current balance equation used a maximum timestep of 0.1 ms. In Movie 3 the synaptic conductances were modulated by the independent but stochastic arrivals of spikes from 50 Purkinje cells and 10 mossy fibers.

Single-compartment model of MVN neurons

We created a simple one-compartment model of MVN cells (Table 2) obeying the current balance equation, $C_m(dV/dt) = -I_L - I_h - I_{syn,Pkj} - I_{syn,MF}$, in which the reversal potential and conductance values for the leak current, $I_L = g_L(V - V_L)$, were determined by the membrane time constant (12 ms), the resting potential (–58 mV), and $I_h = \overline{g}_h q(V - V_h)$, where $V_h = -20$ mV is the mixed-cation reversal potential [148, 149] and q is the activation variable. Synaptic inputs were modeled in a deterministic fashion, as in Model 1 above, with synaptic time constants $\tau_{syn,Pkj}$ = 8.9 ms and $\tau_{syn,MF}$ = 5.5 ms (Chun et al. 2003) and synaptic weights $W_{syn,Pkj}$ = 0.5 mS/cm^2 and $W_{syn,MF}$ = 4 µS/cm^2.

Measured h-current time constants vary broadly across cell types, but detailed measurements of h-current in MVN cells have not yet been made. The HCN2 isoform appears to be the predominant subtype of h-channel in the vestibular nuclei (Santoro et al. 2000), and the kinetics of this isoform are consistent with the activation time constant of hyperpolarization-activated rebound burst firing in the MVN, measured to be ~620 ms at 31°C (Sekirnjak and du Lac 2002). Following (Sekirnjak and du Lac 2002) we modeled I_h kinetics with a fixed time constant, τ_q. To determine the steady-state voltage dependency for HCN2 we fit measurements of total h-current obtained in a xenopus oocyte expression system [149]:

$$q_{\infty} = \left(1.00 + \exp\left[\frac{V+78}{5.53}\right]\right)^{-1}$$

$\bar{g}_h$= 3 mS/cm^2 was chosen to be within a range of values that produced rebound depolarization. We used τ_q= 400 ms for MVN simulations shown in Figures 7C and 7E, because this value approximated the time constant that generated the largest ratio of rebound amplitude between the long ISI condition and the zero ISI condition. Rebound amplitude in the zero ISI condition varied by only ~3 mV across a wide range of time constants, 50 ms < τ_q < 1000 ms.

Linear-Nonlinear (LN) model of lock-and-key mechanism

For our algorithmic description of memory retrieval we generated a set of 'key' activity patterns, $K(t)$, using the CS-driven waveforms for the instantaneous Purkinje cell spike rates arising for ISI values ranging from 0 to 200 ms. The Purkinje cell spiking rates were the same as for biophysical Model 1. We created a linear filter,

$$F(t) = \frac{1}{Z}\left[-\left(T\left[\frac{t-\tau}{\tau}\right] - T\left[\frac{t-t_F-\tau}{\tau}\right]\right) + \frac{2}{3}\left(T\left[\frac{t-t_F-\tau}{\tau}\right] - T\left[\frac{t-2t_F-\tau}{\tau}\right]\right)\right],$$

where Z is a normalization constant chosen to be the maximum absolute value of the linear response, $T\left[t\right]$ is the smooth transition function (see *General stimulation procedures*), t_F = 20 ms and τ = 10 ms. The filtered key activity was determined by the convolution:

$$x(t) = \int_0^{\infty} K(t')F(t-t')\,dt'$$

Finally, this signal was passed through an exponential nonlinearity, $M(t) = G[x(t)] = \exp[h * x(t)]$, where h = 12 was a gain factor. The response amplitude for a given ISI value (Figure 8E) was determined by the peak value of $M(t)$ normalized by the amplitude attained for a long ISI of 200 ms.

Table 1. Parameters for compartmental simulations of DCN neurons.

The symbol, description, value, and citations for each parameter in Models 1-3 of DCN cells.

Parameter	Description	Value	Model	References
C_m	Membrane capacitance	1 μF/cm^2	1,2,3	
V_{rest}	Resting membrane potential	-58 mV	1,2,3	[105, 107]
V_{Ca}	Ca^{2+} reversal potential	140 mV	1,2,3	[140]
$V_{syn,Pkj}$	GABAergic reversal potential, determined by Cl$^-$ gradient	-75 mV	1,2,3	[105, 127]
$V_{syn,MF}$	Glutamatergic reversal potential	0 mV	1,2,3	[139]
V_{Na}	Na$^+$ reversal potential	50 mV	3	[150]
V_K	K$^+$ reversal potential	-90 mV	3	[127, 150]
$r_{Pkj,b}$	Background Purkinje cell spike rate	40 Hz	1,2,3	[134]
$r_{Pkj,d}$	Reduced Purkinje cell spike rate due to LTD of parallel fiber inputs	20 Hz	1,2,3	[134]
$r_{Pkj,p}$	Elevated Purkinje cell spike rate due to LTP of parallel fiber inputs	100 Hz	1,2,3	[134]
$r_{MF,b}$	Background mossy fiber spike rate	10 Hz	1,2,3	[136, 137]
$r_{MF,CS}$	CS-driven mossy fiber spike rate	50 Hz	1,2,3	[136, 137]
$\bar{g}_T$	Maximum T-type Ca^{2+} conductance in Models 1 and 2	0.5 mS/cm^2	1,2	
$\bar{g}_T$	Maximum T-type Ca^{2+} conductance in phase plane modeling	0.3 mS/cm^2	1,2	
$\bar{g}_T$	Maximum T-type Ca^{2+} conductance in Model 3, somatic compartment	5 μS/cm^2	3	
$\bar{g}_T$	Maximum T-type Ca^{2+} conductance in Model 3, dendritic compartment	0.5 mS/cm^2	3	
$W_{syn,Pkj}$	Maximum total conductance of Purkinje cell synapses	0.2 mS/cm^2	1,2,3	
$W_{syn,MF}$	Maximum total conductance of mossy fiber synapses	4 μS/cm^2	1,2,3	
$\bar{g}_{HVA}$	Maximum high-voltage activated Ca^{2+} conductance in Model 2	0.15 mS/cm^2	2	
$\bar{g}_{HVA}$	Maximum high-voltage activated Ca^{2+} conductance in Model 2 phase-plane	0.09 mS/cm^2	2	
$\bar{g}_{HVA}$	Maximum high-voltage activated Ca^{2+} conductance in Model 3, somatic compartment	0.3 mS/cm^2	3	

$\bar{g}_{HVA}$	Maximum high-voltage activated Ca^{2+} conductance in Model 3, dendritic compartment	0.15 mS/cm^2	3	
$\bar{g}_{SK}$	Maximum Ca^{2+}-dependent K^+ conductance	32 μS/cm^2	3	
$\bar{g}_{Na}$	Maximum Hodgkin-Huxley type fast Na^+ conductance	144 mS/cm^2	3	
$\bar{g}_{Kv}$	Maximum Hodgkin-Huxley type K^+ conductance	56 mS/cm^2	3	
τ_m	Membrane time constant	12 ms	1,2,3	[103, 105]
$\tau_{syn,Pkj}$	GABAergic synaptic time constant	14 ms	1,2,3	[139]
$\tau_{syn,MF}$	Glutamatergic synaptic time constant	23 ms	1,2,3	[139]
$N_{syn,Pkj}$	Number of Purkinje cell inputs	50	2,3	
$N_{syn,MF}$	Number of mossy fiber inputs	10	2,3	
R	Ratio of Purkinje to mossy fiber inputs	5	2,3	[145]
g_c	Inter-compartmental coupling	0.53 μS/cm^2	3	[144, 146]
ρ	Percent of membrane surface area occupied by somatic compartment	5%	3	

Table 2. Parameters for compartmental simulations of MVN neurons.

The symbol, description, value, and literature citations for each of the parameters used in simulations of MVN cells.

Parameter	Description	Value	References
C_m	Membrane capacitance	1 μF/cm^2	
V_{rest}	Resting membrane potential	-58 mV	[151]
$V_{syn,Pkj}$	GABAergic reversal potential, determined by Cl$^-$ gradient	-75 mV	[140]
$V_{syn,MF}$	Glutamatergic reversal potential	0 mV	[140, 152]
V_h	I_h mixed-cation reversal potential	-20 mV	[148, 149]
$r_{Pkj,b}$	Background Purkinje cell spike rate	40 Hz	[134]
$r_{Pkj,d}$	Reduced Purkinje cell spike rate due to LTD of parallel fiber inputs	20 Hz	[134]
$r_{Pkj,p}$	Elevated Purkinje cell spike rate due to LTP of parallel fiber inputs	100 Hz	[134]
$r_{MF,b}$	Background mossy fiber spike rate	10 Hz	[136, 137]
$r_{MF,CS}$	Mossy fiber spike rate during head rotation	50 Hz	[136, 137]
$\bar{g}_h$	Maximum h-type cation conductance	3 mS/cm^2	
$W_{syn,Pkj}$	Maximum total conductance of Purkinje cell synapses	0.5 mS/cm^2	
$W_{syn,MF}$	Maximum total conductance of mossy fiber synapses	4 μS/cm^2	
τ_m	Membrane time constant	12 ms	[153, 154]
$\tau_{syn,Pkj}$	GABAergic synaptic time constant	8.9 ms	[152]
$\tau_{syn,MF}$	Glutamatergic synaptic time constant	5.5 ms	[152]
τ_q	h-current activation time constant	400 ms	

RESULTS

A theoretical framework for cerebellum-dependent learning and memory

Cerebellar granule cells number in the ten billions but individually appear to be rarely active, producing only a few spikes at a time in response to mossy fiber input [70]. Such transient activation implies that after behavioral training and plasticity induction at PF-Purkinje cell synapses, presentation of a learned sensory cue should drive a biphasic modulation of population Purkinje cell activity (Figure 2). For example, in classical conditioning transient CS-driven granule cell activity that is concurrent with US-driven CF activity will lead to LTD at PF-Purkinje cell synapses (Figure 2A). CS-driven granule cell activity that is asynchronous with CF activity will lead to LTP. During subsequent CS input, the net effect of LTD and LTP induction at distinct PF-Purkinje cell synapses will be biphasic modulation of the aggregate Purkinje cell activity received by a DCN neuron (Figures 2B and 2C). Similarly, overlapping pulses of vestibular and visual input in VOR adaptation will also lead to biphasic modulation of Purkinje cell activity. This general pattern of modulation does not hinge on the details of granule cell coding but is contingent on there being spike-timing dependent bi-directional plasticity and subsets of granule cells in which sensory driven activity lasts for only portions of the sensory cue duration [16, 17, 85].

We explored the conditions under which biphasic activation of Purkinje cells leads to reliable post-inhibitory rebound depolarization of their target neurons that drive learned motor responses. In classical conditioning, whether Purkinje cell spiking first rises and then falls in response to a learned CS, or *vice versa*, depends on whether the CS and US were paired with a 'forward' (CS-US) or 'backward' (US-CS) ordering. The two patterns of aggregate Purkinje cell activity should be quite distinct in their propensity to induce DCN cell rebounds. A rise and then fall of Purkinje cell spiking appears well suited to induce rebounds by causing a hyperpolarization and then a depolarization in DCN target cells. The DCN cell resting potential is about –58 mV [107], at which T-channels are largely inactivated (Figure 3). The initial hyperpolarization allows T-channels to deinactivate and the ensuing depolarization allows them to activate. The opposite pattern of Purkinje cell spiking resulting from backward training should be a poor initiator of DCN cell rebounds, because the initial depolarization will heighten T-channel inactivation and should largely preclude rebounds. To test these ideas, we performed compartmental modeling of DCN cells to explore if such a disparity in rebound generation could account for the observed differences in behavioral responses following backward versus forward classical conditioning. For our modeling, we described the timing dependence of LTP and LTD induction on the interval between paired activation of PF and CF afferents as a smooth function that permits LTD for PF activity anticipating CF activity by up to a time t_{LTD} ~75 ms (Figure 2A) (see Materials & Methods). This timing dependence mimics that of the experimental data [130]. The maximal levels of LTD and LTP induction in our models did not depend on the CS-US training interval. By comparison, the durations of each phase of the biphasic Purkinje cell activity did vary with the CS-US interval. This distinction allowed us to focus initially on the signal processing performed by the DCN cells rather than on effects that depend on plasticity amplitude. We subsequently explored how changes in plasticity amplitude, as quantified through the resulting changes in Purkinje cell spike rates, affect a rebound-based mechanism for memory recall in the DCN cells.

Memory recall in a one-compartment model DCN neuron

We studied whether forward and backward patterns of biphasic Purkinje cell spiking could lead to distinct patterns of rebound activity in DCN cells after presentation of a classically conditioned stimulus. We created a series of compartmental DCN cell models that received inputs from both Purkinje cells and mossy fibers, and we interpreted the resulting rebounds as the initiators of conditioned motor responses. The simplest model (Model 1) had one electrical compartment, lacked fast spiking capability, and had only leak, T-type, and synaptic conductances (Figure 3A). This allowed us to focus initially on rebound generation, apart from issues studied later concerning membrane potential noise and downstream readout. Kinetic parameters for T-currents were obtained from *in vitro* measurements in DCN cells [141, 143]. Deinactivation can occur within ~20-100 ms of hyperpolarization from the resting potential, and activation can then occur within a few milliseconds during subsequent depolarization (Figures 3D and 3E). Conductance densities were set to reproduce the observed resting potential of -58 mV and membrane time constant of ~12 ms [103, 105, 107].

We compared the model's responses to forward and backward patterns of biphasic Purkinje cell input. In our initial studies, the forward CS-US interstimulus interval (ISI) was at least 200 ms, more than sufficient delay for reliable conditioning in rabbits (Figure 1B) [21]. Mossy fiber excitation rose during the entire CS but was insufficient to drive a rebound during baseline or elevated Purkinje cell spiking. This is consistent with data supporting a key role for Purkinje cells in generating properly timed reflexes via the suppression of early, mossy fiber driven responses to the CS, which can be unveiled by blocking Purkinje cell inputs to the DCN [92, 155]. We found that biphasic Purkinje cell input shaped by forward training led to rebounds that initiated as Purkinje cell spiking transitioned from an elevated to a diminished rate, about t_{LTD} prior to the expected US onset (Figure 4A, red traces). Hence, rebounds could drive blinks that anticipate the US. We then tested the effect of varying the ISI value. With backward training there was insufficient deinactivation of T-currents to generate rebounds (Figure 4A, blue traces). With positive ISI values <100 ms, rebounds occurred but with diminished amplitude, since there was insufficient time for T-channel deinactivation during the brief increase in Purkinje cell spiking (Figure 4A, orange trace). Thus, rebound generation occurred selectively for sufficiently positive ISI values and anticipated US arrival.

We also explored the dependence of rebound generation on the graded magnitude of LTP and LTD at the PF-Purkinje cell synapse, as quantified through the resulting elevation and diminution in Purkinje cell spike rates, respectively (Figures 4B and 4C). Rebound generation in the DCN cell required biphasic Purkinje cell spiking, with both a sufficient elevation and subsequent decline in spiking needed for large amplitude rebounds (~50 mV). Ample levels of both LTP and LTD would thus be needed to induce sufficient biphasic variation in Purkinje cell spiking. These findings held across a broad range of T-channel densities and open the possibility that DCN cell T-currents help shape the differences in conditioned reflex expression following backward and forward training [21].

Readout mechanisms of rebound depolarization and correspondence to conditioned behavior

If rebounds induce learned motor action, how do DCN cells convey rebound magnitudes via the rate of Na^+ spikes sent to pre-motor areas? The graded amplitude of pure T-current mediated rebounds indicates these low-voltage activated events are not stereotyped Ca^{2+} spikes. Real DCN neurons do exhibit Ca^{2+} spikes, mediated by high-voltage activated Ca^{2+}-channels, and as in other cell types dendritic Ca^{2+} spikes may be good triggers of somatic Na^+ spike bursts [105, 127]. We reasoned that the amplitude of T-current mediated rebounds should set the likelihood of crossing the voltage threshold for Ca^{2+} spike generation, with membrane potential fluctuations influencing the degree of variability. Smaller amplitude rebounds that occur with shorter ISI values would be less likely to cross the Ca^{2+} spike threshold. Within this framework we interpret a Ca^{2+} spike as the initiator of signals sent downstream to drive a conditioned motor response.

To test whether this readout mechanism would be able to convert the amplitude of rebound depolarization into the probability of Ca^{2+} spike generation, we examined an enhanced one-compartment model that included high-voltage activated Ca^{2+}-channels (Model 2, Figure 3B) and membrane potential fluctuations due to stochastic arrival of synaptic inputs (Materials & Methods). This contrasts with Model 1, in which both synaptic inputs and membrane voltage followed deterministic time courses. In Model 2 a biphasic pattern of Purkinje cell input resulting from forward training with a long ISI value led reliably to a T-current mediated rebound of sufficient magnitude to trigger a Ca^{2+} spike. These Ca^{2+} spikes were properly timed, prior to the expected US. Backward training led to small rebounds and virtually no Ca^{2+} spiking. Forward training with a short ISI value led to unreliable Ca^{2+} spiking, with the amplitude of the T-current mediated rebound being sufficient to trigger a Ca^{2+} spike on some trials but not others (Figure 5A). Thus, as the ISI value varied, the amplitude of T-current driven rebounds set the probability of crossing the Ca^{2+} spike threshold (Figure 5B, closed green triangles). Of prime interest, the shape of the curve describing this response probability as a function of the ISI closely resembles that obtained in rabbit eyeblink conditioning studies (Figure 5B, open red squares, diamonds, and downward-facing triangles), validating the plausibility of a rebound based recall mechanism. The sum of t_{LTD} and the T-channel inactivation time constant determine the temporal offset of the curve from the origin. It follows that experimental manipulations lengthening the time needed for T-channel deinactivation during the ISI are predicted by our theory to cause a rightward shift of the behavioral data curve (Discussion).

We examined readout issues in greater depth using a two-compartment model DCN cell (Model 3, Figure 3C) that included a dendrite and soma, as well as channels mediating dendritic Ca^{2+} and somatic Na^{+} spikes (Materials & Methods). The somatic and dendritic compartments were only weakly coupled, which was intended to mimic the electrotonic isolation between the cell body and the long distal dendrites of DCN cells where T-channels appear to be most dense, >100 μm from the cell body [143]. This is consistent with the observation that Purkinje cell input triggers DCN cell rebounds much more effectively than somatic hyperpolarization of comparable magnitude [107]. Synaptic inputs in Model 3 arrived stochastically, inducing membrane potential fluctuations. As in real DCN cells, a tonic cation current induced a basal rate of somatic spiking at ~25 Hz [103, 138, 147]. Simulations revealed that a dendritic rebound induces a Ca^{2+} spike, which in turn drives a corresponding increase in the rate of somatic Na^{+} spikes (Figures 5C-E). This increase represents a plausible signal from the DCN cell to downstream pathways for driving learned motor output (Figure 1A). Forward training with an ISI >100 ms virtually always led to such a spike burst. Na^{+} spike bursts occurred with lower probability under the same conditions that failed to produce large amplitude rebounds in Model 1, such as backward training or forward training with a short ISI. Across ISI values the probability of a Na^{+} spike burst closely matched the behavioral dependence of conditioned blinking on the ISI value as observed in rabbits (Figure 5B, closed blue circles).

Phase plane analysis of rebound generation as a robust mechanism for recall

To explore further the basic dynamics and robustness of rebound mechanisms, we studied DCN neuronal dynamics using a phase plane analysis of Model 1. Prior applications of such analysis to other neuron types have provided considerable insight into Ca^{2+} spike generation, spike bursting, and transitions between 'up' and 'down' activity states [156-159]. As is common in phase plane analysis, we focused on the slow dynamical variables that set the relevant timescale. Here, these variables are membrane voltage, V, and the T-type channel inactivation variable, l. The latter has a voltage-dependent time constant of ~10-100 ms (Figure 3E), close to the minimum ISI for reliable memory retrieval (Figure 1B). By comparison, the time constant for T-current activation is ~1-10 ms, considerably faster than motor memory recall and rebound depolarization. Because of this separation of timescales we approximated T-type activation as occurring instantaneously and thus restricted to the (V, l) plane. Rebounds may then be viewed as trajectories in this two-dimensional (2-D) phase plane (Figure 6).

Phase plane analysis of Model 1 revealed the key ingredients for rebounds. The analysis can best be understood by breaking a CS presentation into three stages: the initial resting condition, the ISI, and the remainder of the CS following the ISI (Figure 6A). During each stage, the system has a unique attractive fixed-point at the intersection of the V and l nullclines, the curves on which the time derivatives dV/dt and dl/dt respectively vanish (Figure 6B). The three fixed-points and the ISI value are the chief determinants of the dynamics. At rest (stage 1), the system resides at a fixed-point location at which the T-current is mainly inactivated (open black circle in Figure 6B and Movies 1 and 2). At CS onset and during the ISI (stage 2), mossy fiber and Purkinje cell input to the DCN cell shift the fixed-point location to a potential at which T deinactivates (open green triangle in Figure 6B and Movies 1 and 2). The system approaches the stage 2 fixed-point during the ISI, starting from the resting position (Figure 6B; Movies 1 and 2). The ISI value determines the duration and proximity of the system's approach. At ~t_{LTD} prior to the moment of the expected US (stage 3), Purkinje cell activity declines and the fixed-point shifts to a third location that is depolarized relative to rest (open red square in Figure 6B and Movies 1 and 2). This initiates a rebound that is well timed for driving an anticipatory reflex. More precisely, there is a family of trajectories that undergo rebound depolarization during stage 3, with rebound amplitude a strict function of the (V, l) values attained by the end of stage 2. A 2-D color map of rebound amplitude as a function of (V, l) reveals the basis for the sharp dependence on the ISI value and the stage 2 and 3 fixed-point locations (Figure 6C). In turn, these fixed-point locations depend critically on the degree of biphasic Purkinje cell spiking and thus on the levels of LTP and LTD attained during training.

Stage 3 rebound trajectories with the greatest depolarization initiate in a neighborhood of the (V, l) plane that may be viewed as a memory recall 'reliability zone' from which a large rebound will occur without fail (Figure 6C, red shaded region). The level of LTP and the peak Purkinje cell spiking rate are important, because they determine the proximity of the stage 2 fixed-point to the reliability zone. However, even with sufficient LTP if the ISI is too brief the system does not have time to reach the reliability zone during stage 2, leading to a small or no rebound (Figure 6C, blue shaded region; Movie 2). The rebound amplitude also hinges on the location of the stage 3 fixed-point, due to the dependence of T-channel activation on the reduction in Purkinje spike rate and the level of LTD.

To understand the implications of these observations for a readout mechanism based on Ca^{2+} spike generation (Figure 5), using Model 2 we determined the set of stage 3 trajectories in the (V, l) plane that lead to a Ca^{2+} spike (see Materials & Methods). Large amplitude rebounds that initiated within the reliability zone passed furthest above the spiking threshold (Figure 6D). Rebounds that initiated elsewhere either failed to reach or just crossed threshold. In the presence of membrane potential noise, this implies that if the system reaches the reliability zone the probability of Ca^{2+} spike generation is high. Much as in our two-compartment simulations, this probability falls markedly as the ISI is shortened (Movie 3 and

Figure 5B). The Ca^{2+} spike voltage threshold does not vary much across a wide range of HVA Ca^{2+} channel density (data not shown), indicating Ca^{2+} spiking is a robust readout of whether the system has entered the recall reliability zone. Thus, the phase plane analysis illuminates key features of a rebound-based memory recall mechanism, including conditions for reliable recall.

Role of post-inhibitory rebounds in VOR gain adaptation.

Because cerebellar circuitry is highly conserved, rebound depolarization might serve multiple forms of cerebellar memory recall. For example, floccular Purkinje cells involved in horizontal VOR adaptation project to target cells in the vestibular nuclei that also exhibit significant rebound depolarization *in vitro* mediated by hyperpolarization-activated currents [160, 161]. These currents require further characterization and are expressed to varying degrees across MVN cell types, but as a group the MVN neurons receiving input from the floccular Purkinje cells exhibit exceptionally pronounced rebound burst spiking [160]. The currents involved seem to include the h-type cation current and probably some amount of Na^{+} and T-type Ca^{2+} currents [106, 161, 162]. Regardless of the current identities, the empirically determined time constant (~620 ms) describing the duration of hyperpolarization needed for maximal rebound burst firing is considerably longer than that for DCN cells (Figure 2E) [106]. Might rebound depolarization and the need for a long period of hyperpolarization underlie some of the temporal asymmetries seen in behavioral studies of VOR adaptation?

Well known primate behavioral studies have shown that the amplitude of learned VOR responses depends on the relative timing of vestibular and visual stimuli in a manner resembling the dependence on CS-US timing in classical conditioning. Raymond and Lisberger repeatedly paired a vestibular stimulus, a 600 ms pulse of head rotation, with a brief visual stimulus consisting of moving dots [102]. The visual motion stimulus was presented at one of three different ISI values, a zero ISI condition analogous to backward conditioning (Figure 7C, left panel), a short forward ISI of 225 ms (Figure 7C, middle panel), and a long forward ISI of 450 ms (Figure 7C, right panel). A learned VOR response developed in all cases, but the response amplitude grew as the ISI lengthened. Such dependence on the ISI may be analogous to that seen in eyeblink conditioning. Could rebounds underlie this effect? The timescale of the behavioral effect is similar to that of h-current activation.

To study the issue we created a simple, one-compartment model of an MVN cell in which h-currents mediated rebounds (Figure 7B). The model is analogous to Model 1 of a DCN cell in that the model has only one compartment and lacks the channels responsible for the fast spontaneous spiking that MVN cells exhibit. The single compartment thus better mimics a dendrite than a soma. As before, we interpreted the rebounds as signals driving learned motor responses. MVN simulations used a fixed time constant, τ_q, for I_h activation, as in the MVN cell model of [106].

We used a stimulus protocol based on the Raymond-Lisberger experiments and found the largest rebounds arise when the visual stimulus occurs during the latter portion of the vestibular impulse (Figure 7E). Longer ISI values allow more time for h-currents to activate at hyperpolarized voltages, heightening rebound depolarization. Rebound amplitude also depends on biphasic Purkinje cell spiking and thus on the levels of LTP and LTD induced during training (data not shown). By varying τ_q over a range of values, we found that a value of ~400 ms generated the largest ratio of rebound amplitudes between the long ISI condition and the zero ISI condition (Figures 7D and 7E). This value of ~400 ms for τ_q at physiological temperature appears consistent with the empirical value of ~620 ms measured at 31-33°C by [106]. Thus, the amplitude dependence of learned eye movements on the ISI value might stem from variable levels of current flow through hyperpolarization-activated conductances such as h. However, the component of the learned response that is independent of the ISI value is unlikely to be driven by rebounds and is beyond the scope of our present model, which seeks only to account for the ISI-dependent component.

DISCUSSION

We have presented a lock-and-key hypothesis on how the expression of memory responses may undergo filtering via neurophysiological mechanisms active during memory retrieval. This hypothesis and our computational work exploring a candidate rebound-based lock-and-key mechanism were prompted by data suggesting that backward-ordered classical eyeblink conditioning as well as non-cerebellar forms of aversion conditioning can lead to latent memory storage or changes in neural activity, despite a lack of conditioned responses [99, 163]. The complex manner in which plasticity might evolve across a large network of synaptic connections throughout learning experience also suggests some constraints on motor memory expression might be implemented via neurophysiological mechanisms of recall [97].

We explored these ideas through computational studies of two cerebellar behaviors by examining whether DCN and MVN cells can filter signals from Purkinje cells to influence response timing and prevent certain motor responses. Biophysical models of these two cell types that incorporate rebound channels lead to consistent explanations for behavioral data on cerebellar motor learning. These models make direct links between ion channel kinetics and memory expression, and particularly for eyeblink conditioning yield specific predictions of how learning performance varies as a function of the relative timing of paired training stimuli. Within our lock-and-key framework for these models, subjects undergo both cerebellar LTP and LTD regardless of whether the training stimuli were presented in forward or backward order. As a result, the learned sensory cue drives biphasic Purkinje cell activity. Yet, this biphasic activity triggers rebound depolarization in the DCN cells and drives well-timed classically conditioned reflexes only if the training ISI was sufficiently positive. In this way inappropriate motor responses to conditioned stimuli that do not precede the US sufficiently are avoided. Phase plane analysis reveals the basic ingredients for reliable reflex expression, including ample levels of both LTD and LTP. In MVN cells, rebound currents may underlie the variation of VOR adaptation magnitude with the relative timing of visual and vestibular training stimuli. Such effects hinge on the observed capabilities of both DCN and MVN cells for rebound depolarization.

Electrophysiological properties of DCN neurons are consistent with the rebound theory

In vitro studies of DCN cells have found that rebounds occur in both cerebellar slice and isolated cerebellum-brain stem preparations [103, 105, 107, 127]. MVN neurons also undergo rebounds *in vitro*, but there is more uncertainty about the channels involved [106, 164]. There is also indirect physiological and pharmacological evidence DCN cells rebound *in vivo* [147, 165], including for a class of neurons with blink-related activity [166]. Input from a single Purkinje cell induces a large conductance change in the DCN cell [167], which is sufficient to allow a modest post-inhibitory rebound and increase in Na^+ spike rate (*M. Molineux, personal communication*). Multiple Purkinje cells might drive larger rebounds and spike bursts in concert, such as through coordinated Purkinje cell spiking [168, 169]. The anatomical convergence of many hundreds of Purkinje cells onto each DCN neuron implies that the aggregate activity of a population of Purkinje cells influences DCN cell activity.

The *in vivo* extracellular recordings performed to date of DCN neurons during classical conditioning do not provide strong evidence either for or against our rebound theory. Single-unit and multi-unit recordings both reveal an increase in DCN spiking rate that precedes motor output [9, 170-173]. By comparison, evidence for a pause in spiking during early portions of the ISI is limited. Berthier and Moore reported some cells with reduced spiking at the beginning of the CS, although this pattern is not apparent in all single unit recordings [170, 171]. Multi-unit recordings of DCN cell activity do not exhibit a pause, but these recordings may not provide sufficient sensitivity to reveal a partial reduction in spiking within a sub-population of recorded neurons [9, 172, 173]. Irrespective of these results, the rebound model does not make a strong prediction concerning DCN firing during early portions of the CS. During the early portion of the ISI, DCN neuron spiking might remain virtually unchanged despite increased Purkinje input, due to the dendritic location of most Purkinje synapses and T-type channels (Figure 5). Technically difficult *in vivo* intracellular recordings would be required to determine how sub-threshold responses in DCN cells develop during conditioning.

Experimental predictions for studies of cerebellum-dependent motor learning

The lock-and-key hypothesis leads to a clear prediction that is testable independent of whether rebound depolarization provides a lock mechanism: Classical conditioning with a short (<100 ms) or backward CS-US interval should lead to plasticity in the cerebellar cortex despite the lack of reflex acquisition. Several experimental tests of this prediction are possible. Second-order classical conditioning, which can induce expression of previously latent first-order conditioning, might be useful for demonstrating explicitly that a memory of backward conditioning is formed [99]. Alternatively, studies of backward eyeblink conditioning using human brain imaging techniques might be capable of revealing plasticity related effects (Cheng et al., *Soc Neurosci Abstr*, Program No. 204.13, 2007). Electrophysiological recordings of Purkinje cell spiking during reflex conditioning in decerebrate ferrets and guinea pigs have revealed changes in spiking patterns in response to forward training [174, 175], and thus might be used to examine the effects of backward or short interval conditioning. Recordings from rabbit Purkinje cells, but not DCN cells, were reported to show changes in activity patterns following backward US-CS pairings [163]. This lends support to the lock-and-key hypothesis, but more experimental data is still needed.

Several other testable predictions emerge from our modeling of rebound dynamics during memory recall. Removal or blockade of DCN cell rebound conductances should hinder expression of conditioned blinks. To test this idea, T-channel blockers might be used to prevent rebounds in trained subjects [176, 177]. Failure to impair conditioned blinks would cast serious doubt on our proposed rebound mechanism. A related test might be performed in trained animals during recordings of DCN cells that drive conditioned responses. Transient depolarization of these cells during the ISI in trained animals should prevent or diminish blink-related spiking activity by thwarting T-channel deinactivation. Although perturbation of one cell seems unlikely to disrupt the blink itself, stimulation of many DCN cells might have such an effect.

A corollary to this logic concerns inhibition of olivary cells by projections from GABAergic DCN cells that also receive Purkinje cell inputs and exhibit rebounds [178, 179]. These connections appear critical for extinction of conditioned reflexes following unpaired CS presentations [180]. Thus, blockade of rebound channels should hinder both expression and extinction of conditioned reflexes. A caveat is that rebound channel blockers or electrical stimulation applied to the DCN might alter climbing fiber input to the cerebellum through nucleo-olivary inhibitory feedback [181, 182]. To dissociate the role of rebounds in excitatory versus inhibitory DCN neurons might require genetic tools for cell-type specific manipulation, such as optogenetic techniques [183].

Another prediction of the rebound theory is that blink-related Purkinje cells should exhibit, in the aggregate, biphasic patterns of CS-driven activity in conditioned subjects. This contrasts with Albus's proposal that GABAergic Purkinje cells should drive DCN cells through disinhibition [12]. The recordings made to date of the spiking activity of individual Purkinje neurons in trained animals have revealed a diversity of spiking patterns in response to the CS [134, 163, 174, 175, 184-186]. Although multiple regions of cerebellar cortex have been implicated in eyeblink conditioning, most studies of Purkinje cell activity in classical conditioning have focused on a subset of these areas.

Recordings of Purkinje cell activity in trained rabbits from the cerebellar anterior lobe [184] and cerebellar lobule HVI [134, 163], both areas implicated in eyeblink conditioning, have revealed a mixture of activity increases and decreases in response to the CS. Excitatory responses might occur earlier in the ISI, with activity decreases occurring later [184]. Recordings from lobule HVI in decerebrate guinea pigs and ferrets have also revealed a mixture of responses [174, 185]. One particularly careful study, in which Purkinje cells were first identified as being responsive to the US prior to training, reported mainly decreases in Purkinje cell spiking in response to the CS after training [174]. However, this study also mentions that after training some cells undergo an increase in firing during the first 50-100 ms of the CS, followed by an abrupt drop in spiking. Overall, Purkinje cell recordings in trained animals are mainly consistent with the rebound theory but do not provide sufficient evidence to validate the prediction that aggregate Purkinje cell activity should be biphasic. Given the central role in the theory of such biphasic aggregate responses, it seems important that future studies testing this prediction should record from multiple Purkinje cells concurrently using multi-electrode techniques.

If rebounds are instrumental in driving classically conditioned responses, the timing of training stimuli needed for successful conditioning should depend on rebound channel kinetics. Because the minimum ISI that leads to rebounds is set by the sum of the inactivation time constant and t_{LTD}, an increase in the inactivation time constant should increase the minimum ISI for successful training. Likewise, slowing the kinetics of hyperpolarization-activated currents in MVN cells should increase the interval required between vestibular and visual inputs needed to generate the greatest changes in VOR gain.

Further predictions of the rebound theory concern classically conditioned subjects in which cerebellar LTD or LTP is impaired. Elimination of both LTD and LTP would prevent biphasic patterns of Purkinje cell spiking, precluding rebounds. Our simulations also suggest that if either LTP or LTD is partially impaired, rebound amplitude decreases and sporadically triggers readout by Ca^{2+} or Na^{+} spikes (data not shown). However, rebound driven spiking remains well timed when it occurs. Thus, some animals partially deficient in LTD or LTP might exhibit conditioned responses sporadically or of diminished amplitude, but with proper timing. Greater disruptions of cerebellar plasticity would disrupt response timing due to insufficient biphasic modulation of Purkinje cell spiking.

To date, several groups have studied eyeblink conditioning in mice with disrupted cerebellar LTD. Four strains of such mice exhibited blinks that occurred sporadically or with altered amplitude, but with the same distribution of time courses as wild-type mice [119, 121-123]. It was reported that mice expressing a protein kinase C inhibitor that impaired LTD had ill-timed residual blinks [120]. However, questions have been raised about this study regarding the degree to which motor learning was cerebellar in origin [120, 174, 187]. It is possible PKC inhibition disrupts more than just LTD. Although further study is needed, the former four strains seem consistent with the rebound theory and suggest normal LTD levels might be unnecessary for normally timed responses.

Comparison to other work on the role of rebounds in cerebellar function

Several other authors have considered a potential role for DCN cell rebounds in cerebellar network function. Inspired by the observation that Purkinje cell activity effectively triggers rebounds *in vitro* [107], Kistler and de Zeeuw created computational models emphasizing a potential role for rebounds in reverberatory olivo-cerebellar network activity through which Purkinje cells can affect their own subsequent activity and timed motor responses [188]. In experiments on classically conditioned decerebrate cats, Hesslow noted that brief electrical stimulation of the cerebellar cortex leads to a delayed activation of muscle fibers as assessed by electromyography [165, 189]. This finding was explained as arising from post-inhibitory rebound depolarization in the DCN and supports a basic assumption underlying our models, namely that rebound driven DCN cell activity can trigger motor action.

Aizenman and Linden have found that DCN cell rebound depolarization and the associated spike burst are important for determining the polarity of gain changes at the Purkinje cell to DCN cell synapse (Aizenman and Linden, 1998). Mauk and collaborators have performed extensive network modeling studies of cerebellar learning and have suggested that the rise in intracellular Ca^{2+} due to DCN cell rebounds might be a potent trigger for plasticity at the mossy fiber to DCN cell synapse [85, 124]. A recent experimental study has demonstrated this effect, confirming the importance of post-inhibitory rebound current for potentiation of mossy fiber to DCN cell synapses [190]. However, in the original simulations the inefficacy of brief ISI values for classical conditioning arises from a minimum time interval for synchronization of granule cell activity following the CS, rather than from a rebound based mechanism [85]. For ISI values shorter than the minimum for granule cell synchronization, plasticity in the cerebellar cortex is precluded. By comparison, our lock-and-key hypothesis is based on the supposition of at least some plasticity occurring with brief forward or backward ISI values.

A non-synaptic form of plasticity capable of increasing the intrinsic excitability of DCN and MVN neurons appears suited to raise the propensity for rebounds over the course of training. Bursts of synaptic input can induce DCN and MVN cells to display long-lasting gains in excitability due to increases in the maximum rebound current [162, 191, 192]. A biphasic burst-and-pause pattern of Purkinje cell activity may be optimally suited to induce this increase [124, 190, 191]. Thus, rebound generation might be facilitated during training, perhaps reducing the number of Purkinje inputs needed to drive motor action. Overall, multiple cerebellar plasticity mechanisms in addition to those invoked here at the PF-Purkinje synapse are likely to occur during motor training and to shape both cerebellar dynamics and motor learning in a cooperative fashion [85, 88, 90, 91, 124, 193, 194]. Collectively, the existing experimental data are not only consistent with our computational models but also reach well beyond, suggesting that rebound currents have multiple functions additional to the temporal filtering role explored here.

A role for rebounds in temporal shaping of learned responses

A longstanding debate has concerned the issue of whether cerebellar-mediated memory storage occurs within the cerebellar cortex, or within the deep cerebellar or vestibular nuclei [95, 113]. Consistent with the identification of multiple plasticity mechanisms *in vitro*, recent *in vivo* studies indicate both the cortex and the nuclei have a role in memory storage, but with distinct kinetics for plasticity induction and with distinct roles in setting the timing of motor expression [92, 124, 155, 195, 196]. In both eyeblink conditioning and ocular reflex adaptation, shortly after the start of training motor memory is susceptible to extinction and is impaired by lesion or pharmacological disconnection of the cerebellar cortex. After a few days, motor memories are more resistant to extinction and can persist without Purkinje cell involvement [92, 155, 195, 196].

However, even long after training, Purkinje cells appear important for setting the proper timing of learned responses. In eyeblink conditioning, disconnection of Purkinje cells' projections to the DCN leads to ill-timed, short-latency blink responses to the CS [92, 155]. These short-latency blinks probably arise because without inhibitory input from Purkinje cells, DCN cells may be driven strongly by mossy fiber inputs starting from the onset of the CS. For VOR adaptation, the relative contributions of the flocullar Purkinje cells and the MVN neurons in shaping the temporal character of the learned motor response remain less clear. Our interpretation of the available data is that for VOR adaptation, a baseline component of the learned motor response that is not as sensitive to the relative timing of vestibular and visual training stimuli arises from mossy fiber driven activity in the MVN (Figure 7C). However, there also appears to be another component that is sensitive to the relative timing of the training stimuli [102], and it is this second component that we propose has the proper form and kinetics to be explained by a rebound based mechanism using a slow, hyperpolarization-activated current (Figure 7E), such as that identified by [106]. Unlike eyeblink conditioning, for which there is no learned response with short ISI values, there is VOR adaptation following short interval training, but at reduced amplitude relative to adaptation at longer ISI values [102]. Thus, rebound mechanisms cannot account quantitatively for the full magnitude of VOR adaptation.

Generalizations of the proposed rebound mechanism

Our simulations involved specific choices of cellular parameters, but the ideas presented here on rebounds may have explanatory power surpassing that of our detailed models. One aspect of generality concerns the channels that mediate rebounds. It was first thought mainly T-type channels drive DCN cell rebounds, but new evidence also suggests a role for Na^+ currents (Sangrey and Jaeger, *Soc Neurosci Abstr*, Program No. 179.10, 2005). Further, rebound amplitude or kinetics may vary between individual DCN neurons due to differential expression of T-channel isoforms [197]. Identities of the rebound channels in MVN cells are still in question, but in addition to the h-type current there is evidence Na^+ and low-voltage activated Ca^{2+} currents play a role [106, 161]. Further characterization of rebound channels may be important for understanding the precise timing relations seen in behavioral studies. However, the basic idea that rebounds shape temporally asymmetric behavioral responses is general and may apply to other cerebellum-dependent learning paradigms. For example, learned aspects of smooth visual pursuit exhibit temporal dependencies on the ~200 ms scale that might reflect requirements for rebound generation in cells of the cerebellar caudal fastigial nucleus [198].

Role of synaptic plasticity in memory formation and recall.

A prevalent view holds that LTD and LTP are opposing mechanisms, with one encoding memories and the other erasing [126, 129, 132]. We are proposing a different view, in which both cerebellar LTD and LTP are needed for reliable recall. The lock mechanism based on rebound depolarization enforces the requirement for forward training, but nevertheless both LTD and LTP are induced during forward and backward training.

In our lock-and-key framework, backward conditioning and forward training at short intervals induces synaptic plasticity at parallel fiber to Purkinje cell synapses, despite the absence of behavioral output. In this respect, our work follows several prior models of cerebellar-mediated learning in which neuronal plasticity can occur without any change in motor behavior. Models of learning with two or more stages have been proposed for eyeblink conditioning [85], VOR adaptation [95] and other motor behaviors (Smith et al. 2006). These models involve an intermediate stage of learning during which plasticity has occurred but learning is not yet expressed. More generally, any model of learning with two or more serial stages of plasticity inherently implies the possible existence of plasticity without expression of learning. We also present a two-stage model, but here the second stage of processing is implemented as temporal filtering occurring at memory recall rather than as plasticity occurring during memory formation. As with prior two-stage models of plasticity, our model also predicts the possibility of plasticity without behavior modification.

The idea of rebound generation as a temporal filter pertains to the ongoing debate about mechanisms that shape the dependence of classically conditioned responses on the CS-US ISI (Figure 1B). Ideas from recent studies on synaptic plasticity and animal behavior have suggested conflicting explanations. Some authors have suggested spike-timing dependent plasticity can account for the variation in performance with the ISI value [130]. A timing-dependent plasticity rule in which PF activity preceding CF activity is optimal for LTD induction does create a distinction between backward and forward training (Figure 2). However, the recent and older experimental data show that LTD can occur with forward or backward timing protocols (Figure 2A) [112, 130, 199], so another mechanism auxiliary to LTD seems needed to create conditioned reflexes selectively following forward but not backward training. Our model incorporates spike-timing dependent plasticity, and t_{LTD} is a key factor that causes the conditioned response to precede the expected US. However, it is rebound generation that precludes learned responses following backward training. In a sense, a memory of backward training is formed in the resulting patterns of LTD and LTP, but this memory is not retrieved in response to the CS.

This statement of the lock-and-key hypothesis fits well with studies of classical aversion conditioning, which have shown backward training leads to an associative memory of the US-CS pairing [200]. CS presentation alone does not yield conditioned motor responses, but the associative memory can be demonstrated through second-order conditioning [99].

Thus, at least for this form of aversion conditioning, mediated outside the cerebellum, the lock-and-key hypothesis appears to be correct. What mechanism prevents the stored memory from yielding a response to the CS? For aversion conditioning the answer remains unknown, but our study points to rebound depolarization as a candidate mechanism by which such behavioral filtering might occur for cerebellum-dependent behaviors.

An algorithmic lock-and-key description of memory retrieval

In sensory neuroscience, filters have been fruitfully used to describe receptive fields and spiking behavior in an algorithmic manner, apart from physiological details. Our study suggests an algorithmic lock-and-key description of cerebellar memory recall (Figure 8). The 'lock' resides in the DCN or MVN as a temporal filter that requires specific 'key' input activity to drive rebound spiking and a motor response. The key is the biphasic pattern of Purkinje cell activity that is shaped by training and driven by a learned sensory cue. Synaptic plasticity does not lead to learned responses unless key activity is shaped to match the temporal filter of the lock (Figures 8A and 8B). Backward training induces plasticity, but the resulting key activity does not fit the lock, precluding conditioned responses.

We found that a simple linear-nonlinear (L-N) filter model, inspired by those used to describe visual receptive fields [201, 202], can accurately predict the responses of our biophysical models and the behavioral dependence of learned responses on ISI (Figures 8C-E). Thus, from an algorithmic standpoint memory retrieval may involve a temporal filter that excludes certain behaviors and allows others.

The lock-and-key model of recall might be adaptable to other non-declarative forms of memory involving feed-forward networks. Striatal memory for motor sequences shows promising similarities, since both Purkinje cells and striatal medium spiny neurons receive diverse sensory information, undergo bi-directional plasticity, and are GABAergic projection neurons that trigger motor sequences [203]. The similarity between these two GABAergic cell types extends also to their respective targets, for there is one class of striatal target neuron in the globus pallidum that is spontaneously active, has a baseline membrane potential of about –60 mV, and reliably undergoes post-inhibitory rebound depolarization in response to electrical stimulation of striatal input fibers [204]. By analogy, cortico-striatal plasticity alone may not represent a memory, since select patterns of striatal input activity to the globus pallidum may be needed to unlock motor expression. These ideas challenge classical theories of memory storage that focus almost exclusively on neuronal plasticity but avoid mention of how retrieval dynamics may help shape memory expression.

Neuroscientists have long sought the location and substance of memories. This quest led to the notion of the engram, the physical unit of memory, and to the later idea that changes in synaptic strength are a candidate substrate for memory [205, 206]. Cognitive scientists have countered that synaptic plasticity alone does not account for the complex phenomenology of memory recall [207, 208]. In this opposing view, a memory is a physical dynamic that occurs exclusively at recall. Plasticity may help shape this dynamic, but is not by itself a memory. Our computational models provide concrete examples in which it is difficult to identify any physically localized engram. Would the engram include the synapses that have undergone plasticity, the key activity that triggers a memory, the activity unlocked at a rebound, or all of the above? By itself plasticity seems a poor candidate for a complete engram, since it is insufficient to allow recall. Questions regarding the physical substance of memory may be misleading, neglecting that recall occurs due to a sequence of events culminating in a specific form of neural dynamics. This caveat regarding the substance of memory in our models might also apply broadly to multiple memory systems.

FIGURES

Figure 1. Neural pathways and stimulus timing requirements for eyeblink conditioning. (A) Neural pathways involved in delay eyeblink conditioning. Cerebellar climbing fibers (CFs) originate in the inferior olive (IO) and convey activity driven by the unconditioned stimulus (US). Mossy fibers (MF) originate in the pons and convey activity driven by the

conditioned stimulus (CS). The Golgi (Go) and granule (Gr) cell network processes the CS-driven signals. Purkinje (Pkj) cells receive synaptic inputs from parallel fiber (PF) axons of Gr cells. Pkj cells send GABAergic projections to neurons in the deep cerebellar nuclei (DCN) that drive conditioned motor responses via the red nucleus (RN).

(B) The reliability of conditioned responses to a CS in trained rabbits, as a function of the CS-US interstimulus interval used in training. Data were collected from classic studies of ([108] solid black line and black squares), ([109] dotted blue line and blue diamonds), ([110] solid red line and red triangles), and ([111] dotted green line and green circles).

Figure 2. Cerebellar memory formation based on temporally sparse granule cell coding and bi-directional plasticity at the PF-Purkinje cell synapse.

(A) The relative timing of PF and CF activation sets the propensity towards LTD or LTP. Maximal LTD induction arises when PF activity precedes CF activity by up to a time, t_{LTD}, of ~75 ms, but LTD can also occur when CF activity slightly precedes PF activity [129, 130].

(B) In classical eyeblink conditioning, individual PFs are assumed to exhibit elevated activity during only a brief portion of the CS. By the plasticity rule in A, some PF inputs will be strengthened and others depressed, depending on the relative timing of PF and US-driven CF activity. DCN cells receive input from populations of Purkinje cells whose activity reflects aggregate input from CS-activated PFs.

(C) Repeated CS-US training (upper panels) leads to biphasic CS-driven Purkinje cell spiking due to the bi-directional plasticity shown in B. In subjects that received forward training (lower left), spiking rises and then falls relative to baseline (red curve). In subjects that received backward training (lower right), spiking falls and then rises (blue curve). The red arrows (lower left) correspond to t_{LTD}.

Figure 3. Compartmental modeling of T-type Ca^{2+} current rebounds in DCN cells. Compartmental simulations of a Purkinje target neuron in the DCN involved three models of increasing complexity.

(A) Model 1 has one electrical compartment, contains T (g_T) and leak (g_L) conductances, and receives glutamatergic mossy fiber and GABAergic Purkinje cell inputs. Membrane voltage follows a deterministic time course.

(B) Model 2 adds high-voltage activated Ca^{2+} (g_{HVA}) channels. Synaptic inputs arrive stochastically, leading to membrane potential fluctuations and non-deterministic dynamics.

(C) Model 3 has dendritic and somatic compartments, coupled by a conductance, g_c. Synaptic inputs are localized to the dendrite, approximating empirical findings. The soma has fast Na^+ (g_{Na}) and delayed rectifier K^+ (g_{Kv}) conductances. Both compartments have leak, T, SK, and HVA Ca^{2+} conductances. Synaptic inputs arrive stochastically, leading to non-deterministic dynamics.

(D) Voltage dependence of the activation (dashed red curve) and inactivation (solid blue curve) gating variables for the T-type conductance in DCN neurons. At the resting potential (about –58 mV, dashed vertical line), T-currents are largely inactivated. Hyperpolarization deinactivates T-currents, allowing activation during subsequent depolarization.

(E) Voltage dependence of the T-channel activation (dashed red curve) and inactivation (solid blue curve) time constants. Parameter dependencies in D and E are based on [141, 143].

Figure 4. DCN cell rebounds require a minimum CS-US interstimulus interval and sufficient expression of cerebellar LTP and LTD.

(A) The time course of CS-driven depolarization in Model 1 (Figure 3A). If prior training involved a sufficiently positive ISI, the CS-driven rebound is of large amplitude and occurs at a time ~t_{LTD} before the expected US (red traces). If training involved an insufficient ISI value, CS-driven rebounds do not occur (blue traces). For short ISI values, rebounds are diminished in amplitude (orange trace). The color bar indicates the ISI values, which are also marked above the graph with the color corresponding arrowheads for each voltage trace. Rebounds occur prior to the expected US, indicating anticipatory responses.

(B) Rebound amplitude varies with the degree to which the CS drives biphasic Purkinje cell activity. This, in turn, depends on having sufficient expression of both PF-Purkinje cell LTP and LTD (Figure 2). Driving a large amplitude rebound in the DCN cell requires that during the first phase of biphasic activity the Purkinje cell spiking rate rises well above the spontaneous frequency of 40 Hz. The three voltage traces (blue, cyan, red traces) in panel B_1 occurred with the color corresponding, Purkinje cell peak spiking rates shown in B_2. Lower peak spiking rates reflect lower expression levels of LTP. The arrowhead indicates the ISI value of 200 ms.

(C) Driving a large amplitude rebound in the DCN cell also requires that during the second phase of biphasic activity the Purkinje cell spiking frequency drops below the 40 Hz spontaneous rate. The three voltage traces in panel C_1 (blue, cyan, red) were created using the color corresponding, Purkinje cell minimum spiking rates shown in C_2. The higher rates reflect lesser degrees of LTD. The arrowhead indicates the ISI value of 200 ms.

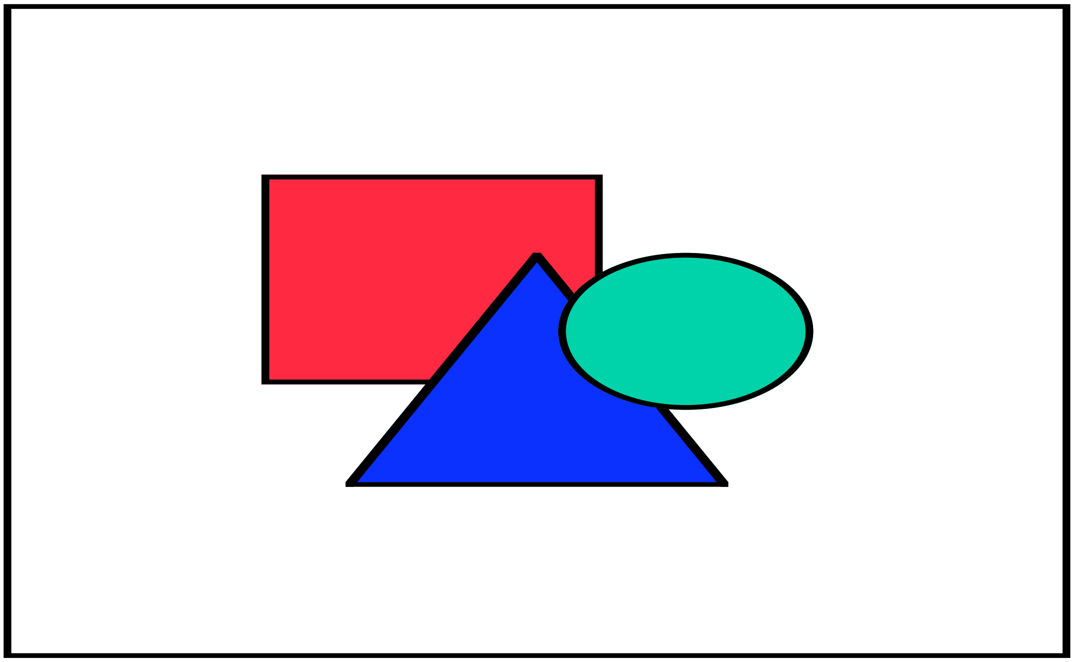

Figure 5. Readout of rebounds via Ca^{2+} spikes leads to a dependence of the response reliability on the CS-US interstimulus interval.

(A) Sample voltage traces during CS presentation in Model 2 (Figure 3B) in the presence of membrane potential fluctuations from noisy synaptic inputs. At an intermediate ISI of 100 ms, a T-current mediated rebound depolarization triggers a Ca^{2+} spike during one trial (dashed red line) but not another (solid blue line).

(B) The reliability of learned responses in Model 2 (closed green triangles) and Model 3 (closed blue circles), defined as the probability of generating a dendritic Ca^{2+} spike in response to a test CS, plotted as a function of the ISI. Classic data on the reliability of conditioned blinks in trained rabbits are re-plotted from Figure 1 (open red symbols) [108-110], showing the similarity to the model data. A t_{LTD} of 75 ms was used for the model data, which is consistent with empirical data indicating t_{LTD} is in the ~50-200 ms range [130].

(C, D) Example voltage traces from the dendritic and somatic compartments of Model 3 (Figure 3C) during CS presentation with an ISI of 200 ms. A T-mediated rebound depolarization leads to a high-voltage activated dendritic Ca^{2+} spike (C) that drives a rise in the somatic Na^{+} spike rate (D).

(E) The corresponding time courses of the activation (n, solid red curve) and inactivation (l, dashed blue curve) gating variables during the Ca^{2+} spike.

Figure 6. Phase plane analysis of CS-driven rebounds.

(A) Membrane voltage time course (blue curve) in response to a CS that initiates at time t=0 in Model 1, under the approximation of instantaneous relaxation of the T-channel activation variable to its asymptotic value. The rebound peaks at a time ~40 ms prior to the expected US at 200 ms after CS onset. Dashed vertical lines delineate three stages of the phase plane trajectory in B.

(B) The state trajectory (blue curve) in the 2-D phase plane defined by the voltage (V) and T-type inactivation variable (l), corresponding to the voltage trace in A. The open black circle marks the fixed-point in the resting state (stage 1). The open green triangle marks the fixed-point from CS onset until ~t_{LTD} prior to the expected US (stage 2). The open red square marks the fixed-point during the remainder of the CS (stage 3). According to longstanding convention, channels are completely inactivated when l equals zero [209].

(C) A color plot conveying the amplitude of the rebound that occurs during stage 3 for the state trajectory passing through each point in the phase plane of B and converging towards the stage 3 fixed-point (open red square). Warmer hues indicate the larger rebounds (color bar) that initiate if during stage 2 the system has successfully entered the ‘memory reliability zone’ near the stage 2 fixed-point (open green triangle). White curves are example state trajectories.

(D) The addition of high-voltage activated (HVA) Ca^{2+} channels to the phase plane analysis of C reveals those stage 3 trajectories that lead to a Ca^{2+} spike (red trajectories) and those that do not (blue trajectories). All of the trajectories closely concur with those in Model 1 (panel C) in the voltage range $V < -35$ mV over which the HVA Ca^{2+} channels are largely closed. The red trajectories, which initiate within the reliability zone near the stage 2 fixed-point (green triangle), cross the Ca^{2+} spike threshold and allow successful readout of the rebound (Figures 5A and 5B). Horizontal dotted lines indicate the resting potential of –58 mV in A-D. Solid and dashed black curves in B, C, and D are nullclines during the resting state for the l and V variables, respectively, on which the time derivatives dl/dt and dV/dt respectively vanish during stage 1.

Figure 7. Vestibular nuclei cell rebounds lead to temporally asymmetric VOR adaptation.

(A) Vestibulo-cerebellar pathways for VOR horizontal gain adaptation involve Purkinje cells (Pkj) that project to target neurons within the medial vestibular nucleus (MVN). Neurons in the MVN project to brainstem motor nuclei (MN) that drive eye movement. Slip of the visual scene on the retina is conveyed to the cerebellum via climbing fibers (CFs). Information about head velocity arrives via mossy fibers (MFs) originating in the vestibular ganglia (VG), is processed within the Golgi (Go) and granule (Gr) cell network, and reaches Purkinje cells by way of parallel fibers (PFs). Conjunctive arrival of CF and PF signals is thought to induce synaptic plasticity at the PF-Pkj synapse that underlies gain adaptation.

(B) A one-compartment model of an MVN Purkinje target neuron that contains h- (g_h) and leak (g_L) conductances and receives glutamatergic mossy fiber and GABAergic Purkinje cell input. Membrane voltage follows a deterministic time course.

(C) Primate behavioral data from well-known studies in which pulses of head rotation (top panel, black bars) were paired during training with moving dot visual stimuli (top panel, grey bars) at three distinct interstimulus intervals. During later testing with pulsed head rotations in the dark, the learned component of VOR expression increased markedly with greater ISI values (bottom panel, green, blue, and red curves) [102].

(D) Relative rebound amplitude as a function of the h-current activation time constant, τ_q. The plot shows the maximum depolarization from the resting potential following training with zero (green), short (blue), and long (red) ISIs, normalized for each value of τ_q by the maximum depolarization of the zero-ISI trajectory (green). Dashed black line indicates τ_q of 400 ms used for simulations shown in panel E.

(E) Voltage traces (top panels) and state trajectories (bottom panels) from the model MVN cell in response to a test pulse of head rotation following training with the three different ISI values shown in C. The three state trajectories (bottom panels) traverse the 2-D phase plane defined by the voltage, V, and the activation level of the h-current relative to that at rest, h. Horizontal dashed lines in the top panels indicate the resting potential of –58 mV. The solid and dashed black curves in the bottom panels are the nullclines during the resting state for V and h, respectively, on which their respective time derivatives vanish. Vertical dashed lines in panels C and E mark the period of head rotation.

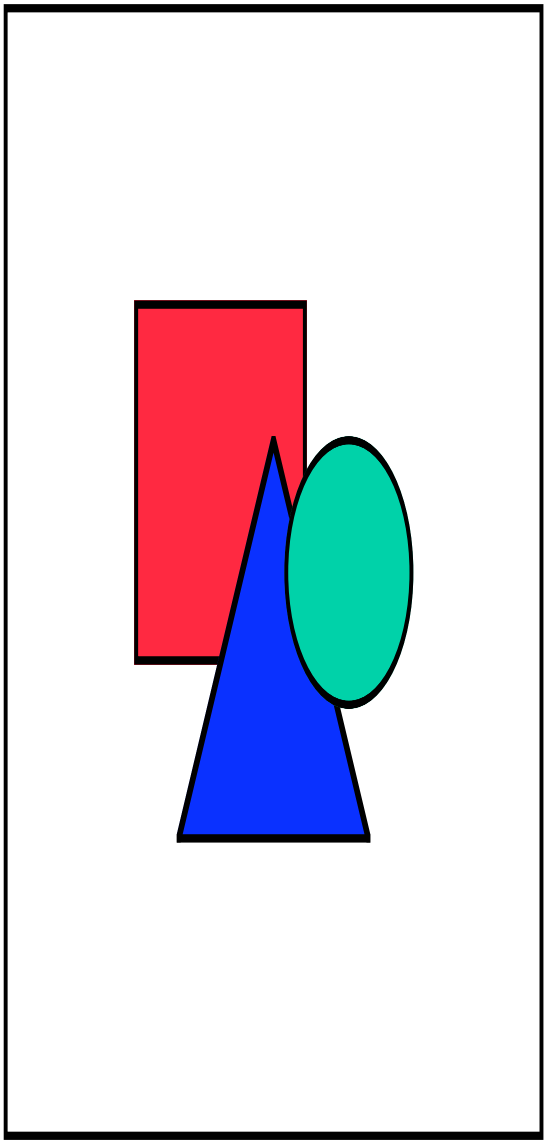

Figure 8. A lock-and-key description of memory recall.

(A) Schematic of memory formation within the lock-and-key description. Paired PF and CF activity induce synaptic plasticity in the cerebellar cortex. This shapes Purkinje cell ‘key’ activity in response to subsequent presentations of the learned sensory input.

(B) Schematic of memory recall, in which the lock resides in the DCN. Key activity driven by learned sensory input is sent to the DCN via Purkinje cell axons. Not all keys will be successful at driving DCN cell activity and learned motor responses. The lock prevents inappropriate motor responses by filtering out non-matching key activity. Partially matching key activity leads to unreliable recall or responses of diminished amplitude.

(C) An implementation of a lock-and-key mechanism using a linear-nonlinear (L-N) filter model, similar to those used to describe sensory receptive fields. Key activity, K (t), that is input to the lock passes through a linear filter, F (t), and then a nonlinear threshold function, G (x).

(D) The L-N model enables responses, equal to G (K (t) $*$ F (t)), to be made selectively to only those keys, K (t), that are shaped by training with sufficiently positive interstimulus interval (ISI) values.

(E) Response of the L-N model as a function of ISI in eyeblink conditioning. The results mimic those of DCN cell simulations (Figures 4-6) and the classic rabbit behavioral data (Figures 1B and 5B).

MOVIES

Movies 1, 2 and 3 are available from the Journal of Neurophysiology at http://jn.physiology.org/

Chapter 3: High-speed, miniaturized fluorescence microscopy in freely moving mice

ABSTRACT

A central goal in biomedicine is to explain organismic behavior in terms of causal cellular processes. However, concurrent observation of mammalian behavior and underlying cellular dynamics has been a longstanding challenge. We describe miniaturized (1.1 g) epi-fluorescence microscopy for cellular-level brain imaging in freely moving mice and applications to imaging microcirculation and neuronal Ca^{2+}-dynamics.

Since the invention of compound optical microscopy, an ongoing pursuit has been the development of increasingly versatile intravital microscopes for observing live cell dynamics, ideally in behaving organisms. In mammals, achieving this ideal has been elusive due to the difficulties of microscopy in active animals. Among mammalian species, the mouse is the preeminent genetic model for targeted manipulation of cellular properties. Thus, innovation of microscopes for use in active mice holds particular importance due to potential for combining genetic manipulations with cellular-level observations during animal behavior.

Prior investigations have exploited flexible fiber-optics to study behaving mammals, including mice [210-212], but these involved monitoring of bulk parameters rather than imaging [211, 212], or epi-fluorescence [210, 213] or reflectance [214] imaging at resolutions too coarse to resolve individual cells. One pioneering study demonstrated fiber-scanning two-photon microscopy in freely moving rats [215]; however, except when the animals were still, imaging suffered from motion artifacts and did not allow measurements of cellular dynamics during active behavior. These artifacts arise due to slow (~2 Hz) rates of frame-acquisition during fiber-scanning, movement-induced distortions to scanning patterns, and brain displacements relative to the focal plane. Head-fixation of alert animals reduces motion artifacts and permits use of conventional instrumentation but precludes many animal behaviors [28] (Nimmerjahn *et al.*, Society for Neuroscience abstracts, 2007).

We describe an epi-fluorescence microscope (1.1 g mass) that can be carried by an adult mouse and allows high-speed cellular imaging (≤ 100 Hz) during active behavior (Figure 1a,b). Penetration depths into tissue using epi-fluorescence are inferior to those of two-photon imaging, but we reasoned that the faster frame rates, lack of scanning, and greater depth of field of epi-fluorescence imaging would reduce motion artifacts. Epi-fluorescence microscopy can also address broader fields of view, since miniaturized scanners for portable two-photon imaging generally have limited range [216]. Further, some questions about deep tissues are addressable via epi-fluorescence, by labeling neurons with deep somata but superficial dendrites [213], or insertion into tissue of micro-optical objectives [216].

Our microscope has fiber- and micro-optics for light transmission and miniaturized gears for focusing (Figure 1 and Figure 5a). There are three microlenses, a 1-mm-diameter gradient refractive index (GRIN) objective, a 2-mm-diameter GRIN focusing lens, and a 2-mm-diameter coupling lens that projects an image of the specimen onto a dense bundle of optical fibers (Figure 1c). We designed two objective versions, a longer, needle-like objective (6.2 mm long) for imaging in deep tissues (Figure 1a), and a shorter (1.4 mm) version for superficial tissues.

Light propagates bi-directionally in the fiber bundle, which delivers illumination and transmits images (240-370 μm diameter; ~2.8-3.9 μm lateral resolution) to a camera (Figure 1d). To relieve torsional strain within the bundle as the mouse behaves, the bundle rotates freely within a commutator, allowing the animal to move without significant resistance. Thus, captured image sequences exhibit rotations that depend on the mouse's actions. We used an encoder to track the bundle's rotations (Figure 1d). Based on the encoder data and an image registration algorithm, we computationally realigned stable records of brain dynamics (Figure 1e and Figure 5b).

As an initial test we studied microcirculation in the neocortex (n = 4 mice) and CA1 hippocampal area (n = 8), using the short and long objective, respectively. Prior studies visualized microcirculation in these tissues in anesthetized animals [217, 218]. After fixation of the microscope onto the cranium and intravascular injection of fluorescein-dextran dye, we began imaging while animals were anesthetized. Erythrocytes appeared dark against the brightly labeled blood plasma [217, 218]. Image-acquisition at rates ≤ 100 Hz permitted observations of single erythrocytes. We then allowed the mice to awake and resume active behavior. We again performed high-speed imaging, often with the mice running vigorously. After data realignment, erythrocyte flow was readily evident (Figure 2a). Few motion artifacts were apparent, generally < 2 μm of sporadic lateral displacement (Figure 5c). The microscope has ~10 μm depth of field and axial motion artifacts were normally not apparent. Thus, our device allows physiological dynamics to be imaged at higher speeds and seems to provide superior robustness to motion artifacts as compared to fiber-scanning two-photon microscopy [215].

Detailed analysis of microcirculation focused on a subset of data from neocortex (n = 3 mice) and hippocampus (n = 3) (Figure 2b-j). By using temporal cross-correlation analysis (Supplemental Methods) we found broadly distributed flow speeds for capillaries (< 10 μm diameter) and venules and arterioles (Figure 2b-j). For individual mice we calculated flow speed maps (Figure 2f,i). The range of speeds we found in behaving mice lies within reported ranges from anesthetized rodents [217, 218]. We also studied relationships between flow speed and vessel diameter. There were noticeable differences between mice, but larger vessels generally exhibited higher speeds in neocortex (Figure 2g) (correlation coefficient, $\langle r \rangle = 0.46 \pm 0.05$; mean ± s.d.; n = 3 mice) and hippocampus (Figure 2j) ($\langle r \rangle = 0.67 \pm 0.10$; n = 3). Given these findings, miniaturized microscopy should be applicable to studies of cerebrovascular diseases and interactions between brain activity, behavior, and hemodynamics [218].

We next studied Ca^{2+}-dynamics of cerebellar Purkinje cells during active locomotor behavior. Purkinje cells' dendritic trees lie within parasagittal planes and extend into superficial portions of the cerebellar molecular layer accessible by epi-fluorescence imaging. Stereotyped Ca^{2+}-action potentials are known to pervade the Purkinje cell dendritic tree during complex (Na^{+} and Ca^{2+}) spikes, are believed critical for motor coordination, are driven reliably by climbing fiber input from the inferior olive, and have been visualized previously by *in vivo* Ca^{2+}-imaging [219, 220]. To label Purkinje cells with fluorescent Ca^{2+}-indicator, we injected membrane-permeant Oregon Green-488-BAPTA-1-acetoxymethyl into the cerebellar vermis prior to mounting the microscope with the shorter objective installed.

While mice were anesthetized we first characterized Ca^{2+}-spiking within up to tens of dendritic trees concurrently. Ca^{2+}-spikes appeared stereotyped, with a brief rise in relative fluorescence intensity (0.5%-1.5% $\Delta F/F$) followed by an exponential decline over ~150 ms reflecting the dye's kinetics of Ca^{2+} unbinding. Spikes occurred within a parallel series of stripe-shaped regions (~100-250 μm long, ~7-11 μm wide) that were approximately aligned to the rostral-caudal axis and consistent with the breadth of Purkinje cell arborizations [220].

This striped appearance occurs because the dendrites extend perpendicular to the field of view. The striped profiles matched those we and others have observed by *in vivo* two-photon imaging [219] and verified correspond to complex spikes by simultaneous electrophysiological recordings (Supplemental Methods) [29].

After terminating anesthesia we studied Ca^{2+}-spiking during animals' immediate recovery (Figure 3a-d) and subsequent active behavior (Figure 3e-h and Figure 6) ($n = 6$ mice). To account for slight spatial heterogeneities in illumination causing variations in fluorescence intensity as the bundle rotates, relative fluorescence traces from behaving mice were normalized by the mean illumination for a given bundle orientation and denoted with primes ($\Delta F'/F'$). Ca^{2+}-spikes were again apparent in striped domains (Figure 3a,b,e,f), allowing spike identification by an algorithm that accounts for dye kinetics by a temporal deconvolution prior to application of the detection threshold (Supplemental Methods). Average spike waveforms were closely similar across anesthetized, alert, and actively moving animals (Figure 3d,h).

To determine spike rates, we assessed false positive detections by analyzing signals from control regions with the same forms as Purkinje cells but orthogonal, medial-lateral orientations. These control regions yielded false positive spike rates that depended on the detection threshold (Supplemental Methods). After setting the threshold such that false positive rates were fixed at an insignificant value (0.025 Hz), we examined Ca^{2+}-spiking rates of Purkinje cells. Mean rates in awake mice were 0.57 ± 0.39 Hz (mean ± s.d.; $n = 42$ cells from 3 mice), within reported ranges [221]. Our percentage of spikes detected in error (2-10%) was within the error ranges (0-30%) achieved by extracellular electrophysiological recordings in live animals [222].

We next examined whether Purkinje cells in the vermis exhibit significant changes in net rates of Ca^{2+}-spiking during active locomotion. Previous electrophysiological studies have examined this question in decerebrated cats, but the data have often been ambiguous [25]. We sorted our recordings into periods when the awake mice were active or resting by using the encoder trace as an indicator of animal movement. Periods when the mouse rotated the fiber bundle at > 200 deg/s were designated as movement episodes. Analysis of Ca^{2+}-spiking rates showed nearly all Purkinje cells exhibited moderately greater spiking during motor activity compared to rest, creating highly significant differences between spike rate distributions in the active and resting states ($P < 10^{-7}$; one-sided Wilcoxon signed rank test) (Figure 4, Supplemental Methods).

We also compared correlations in Ca^{2+}-spiking between pairs of Purkinje cells. Prior studies in rats have reported increases in correlated Purkinje cell activity during tongue-licking [46]. These studies used electrode arrays > 2 mm in extent with 250 μm between electrodes. Our observations are of local networks of multiple neurons within a single field 240-370 μm across. Calculation of the mean correlation coefficient across all simultaneously recorded cell pairs [46] (Supplemental Methods) revealed modest increases in correlated Ca^{2+}-spiking during active movement as compared to rest ($P < 10^{-4}$; one-sided Wilcoxon signed rank test) (Figure 4). We more frequently saw synchronous Ca^{2+}-activity within neighboring, but not distally separated, dendritic trees (Figure 6). This is consistent with the notion Purkinje cells are organized into micro-zones of correlated activity [220].

In summary, we created a microscope allowing cellular-level imaging in freely behaving mice. As compared to fiber-optic two-photon microscopes [215], which offer optical sectioning and lower background fluorescence, our microscope provides broader fields of view, faster frame rates, and decreased susceptibility to motion artifacts, but with the ~50-100 μm depth limitations inherent to epi-fluorescence imaging [216]. Our microscope provides superior resolution to that attained by fiber bundles used without micro-optic objectives [210]. This latter approach, unlike ours, has not resolved cells. The GRIN microlenses we employed are similar to those for fluorescence microendoscopy [216, 217], but modified to accommodate our microscope's focusing capability. Overall, our microscope is well suited for studying cellular processes in accessible tissues, such as the Purkinje cell Ca^{2+}-dynamics examined here, within actively behaving wild-type or genetically manipulated mice or other laboratory animals.

Other cells that might be studied with our microscope include olfactory bulb neurons or neocortical neurons with superficial dendrites. Deeper cells might be accessed by insertion of the microscope's longer objective into tissue. Feasible improvements include addition of a second fluorescence channel or electrodes for auxiliary electrophysiological recordings. Given the growing use of animal disease models, including numerous mouse models of cerebellar ataxia and Purkinje cell degeneration, miniaturized microscopy should permit studies in behaving mice of neurological, behavioral, and cerebrovascular disorders.

FIGURES

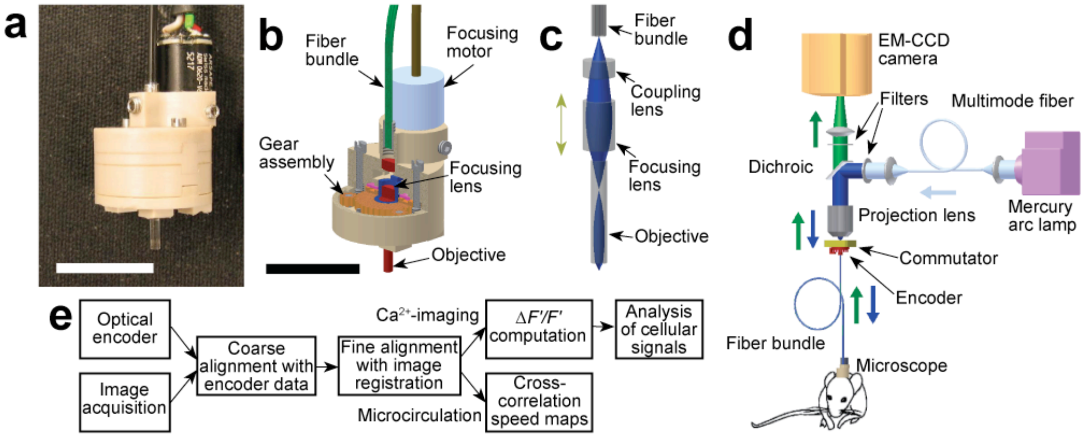

Figure 1. Miniaturized fluorescence microscopy for high-speed brain imaging in freely behaving mice.

(a) Miniaturized fluorescence microscope.

(b) Cutaway view of a computer-aided design (CAD) of the microscope.

(c) CAD drawing of the microscope's optical components. An image of the specimen is projected onto the fiber bundle using three microlenses, one of which is a focusing lens. Yellow arrows show the 1.1-mm movement range of this lens. A bundle of rays (blue) is shown passing through the optics on axis.

(d) Illumination and light collection pathways. Illumination from a Hg-arc lamp passes through a multi-mode fiber, reflects off a dichroic mirror, and is coupled into the fiber bundle. The bundle delivers illumination to the microscope on the mouse and returns the fluorescence image, which is focused onto a high-speed camera. A commutator allows the bundle to rotate as the mouse moves. An encoder tracks these rotations for offline image stabilization.

(e) Flow chart showing procedures for image alignment and analysis.

Scale bars in a,b are 1 cm.

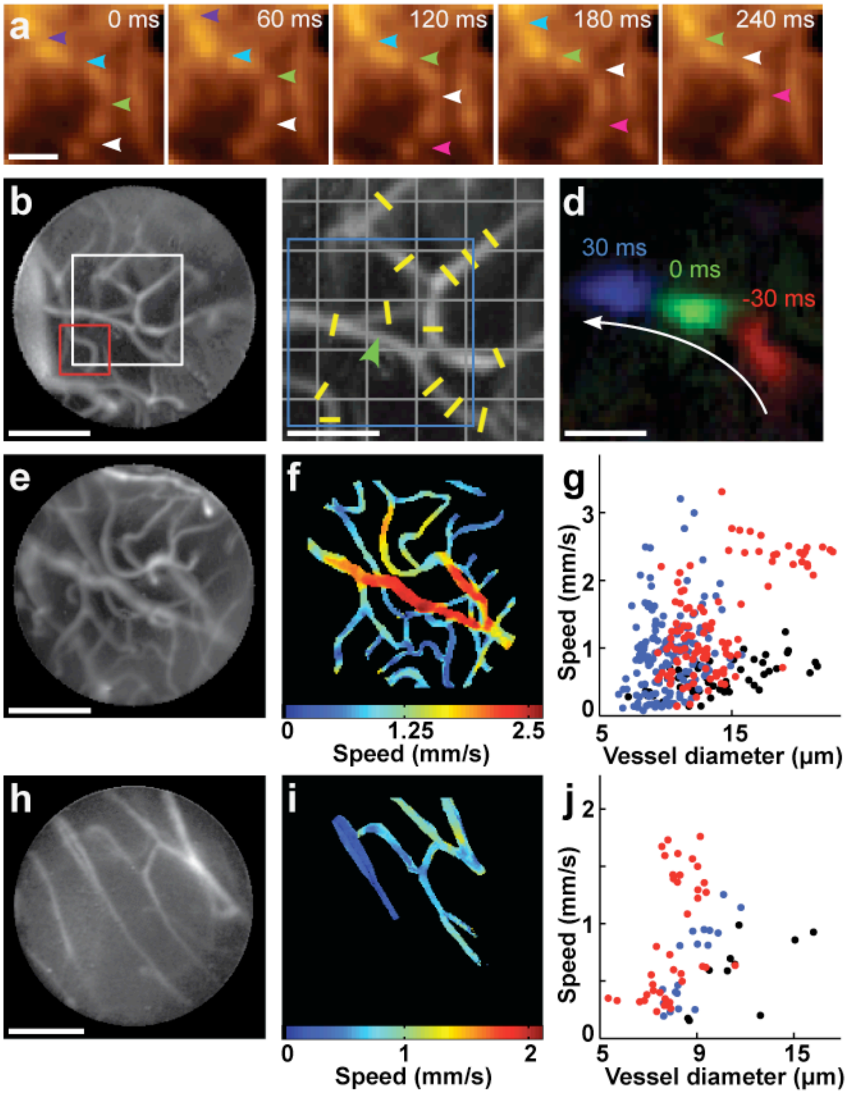

Figure 2. High-speed imaging of cerebral microcirculation in freely behaving mice.

(a) Raw images of neocortical microcirculation in a behaving mouse, cropped to within the red box in (b). Arrowheads mark progress of single erythrocytes.

(b) Image of neocortical vasculature, computed as the standard deviation of images acquired over a 10 s interval. Red and white boxes enclose regions shown in (a) and (c), respectively.

(c) Methods for vessel diameter and erythrocyte speed determinations for (g) and (j). Images were divided into a grid and no more than one vessel was sampled per square. Yellow lines indicate cross-sectional vessel widths. Arrowhead points to where erythrocyte speed was computed in (d), which shows the area inside the blue box.

(d) Three cross-correlation maps computed at -30, 0, and +30 ms time delays, revealing average erythrocyte progression over 7 images acquired at 100 Hz.

(e) Image of neocortical vasculature in an awake mouse, obtained as in (b).

(f) Map of erythrocyte speed across image (e) when the mouse was freely moving.

(g) Average erythrocyte speed *vs*. vessel diameter in the neocortex of behaving mice. Each data point color represents a different mouse.

(h) Image of hippocampal vasculature in an awake mouse, obtained as in (b).

(i) Map of erythrocyte speed across image (h) when the mouse was freely moving.

(j) Average erythrocyte speed *vs*. vessel diameter in the hippocampus of behaving mice. Each data point color represents a different mouse.

Neocortical images acquired at 100 Hz with 3 × 3 pixel-binning on the camera and 450-550 μW illumination. Hippocampal images acquired at 75 Hz using 2 × 2 binning and 250-1020 μW. Scale bars: 20 μm (a, d), 100 μm (b, e, h), 40 μm (c).

Figure 3. High-speed imaging of cerebellar Purkinje cell dendritic Ca^{2+}-spiking

(a) 12 dendritic tree segments identified in an unrestrained mouse sitting quietly while recovering from anesthesia. Filled areas mark two cells studied in (c) and (d).

(b) Average of 8 frames in which the red filled cell of (a) exhibited a Ca^{2+}-spike.

(c) $\Delta F/F$ traces averaged over the color corresponding filled areas in (a).

(d) Average of 350 ms windows surrounding 20 (red) and 22 (blue) Ca^{2+}-spikes extracted by temporal deconvolution and threshold detection from the color corresponding traces in (c). The windows were triggered to begin 90 ms before the onset of the spike.

(e) 17 dendritic trees segments identified in a freely behaving mouse. Areas filled red and blue mark cells studied in (g) and (h).

(f) Average of 4 frames in which the blue-filled cell in (e) exhibited a Ca^{2+}-spike.

(g) $\Delta F'/F'$ (blue and red traces) averaged over the color corresponding filled areas in (e).

(h) Average of 350 ms windows surrounding 37 (red) and 23 (blue) spikes from the color corresponding traces in (g).

All images acquired with 3×3 binning of the camera pixels. Illumination power at the specimen was ~56 µW (a-d), and ~420 µW (g-j). Imaging frame rates were 62.5 Hz (c-d), and 100 Hz (g-j). Distance scale bars are 100 µm.

Figure 4. Comparisons of Purkinje cell Ca^{2+}-spiking during rest and motor behavior.

(a) For each of 42 cells, a comparison of Ca^{2+}-spiking rates during periods when the mouse was moving versus periods of rest as determined from the rotational velocity of the animal's head. Dashed line demarcates equal rates under both conditions.

(b) Comparisons of Ca^{2+}-spiking rates (mean ± s.e.) (left) and correlation coefficients for pairwise synchronous Ca^{2+}-activity (right), during periods of rest and active movement, for the 42 cells of (a). Modest but highly significant differences exist between the two behavioral states, for spiking rates and correlation coefficients ($P < 10^{-7}$ and $P < 10^{-4}$, respectively; one-tailed Wilcoxon signed rank test).

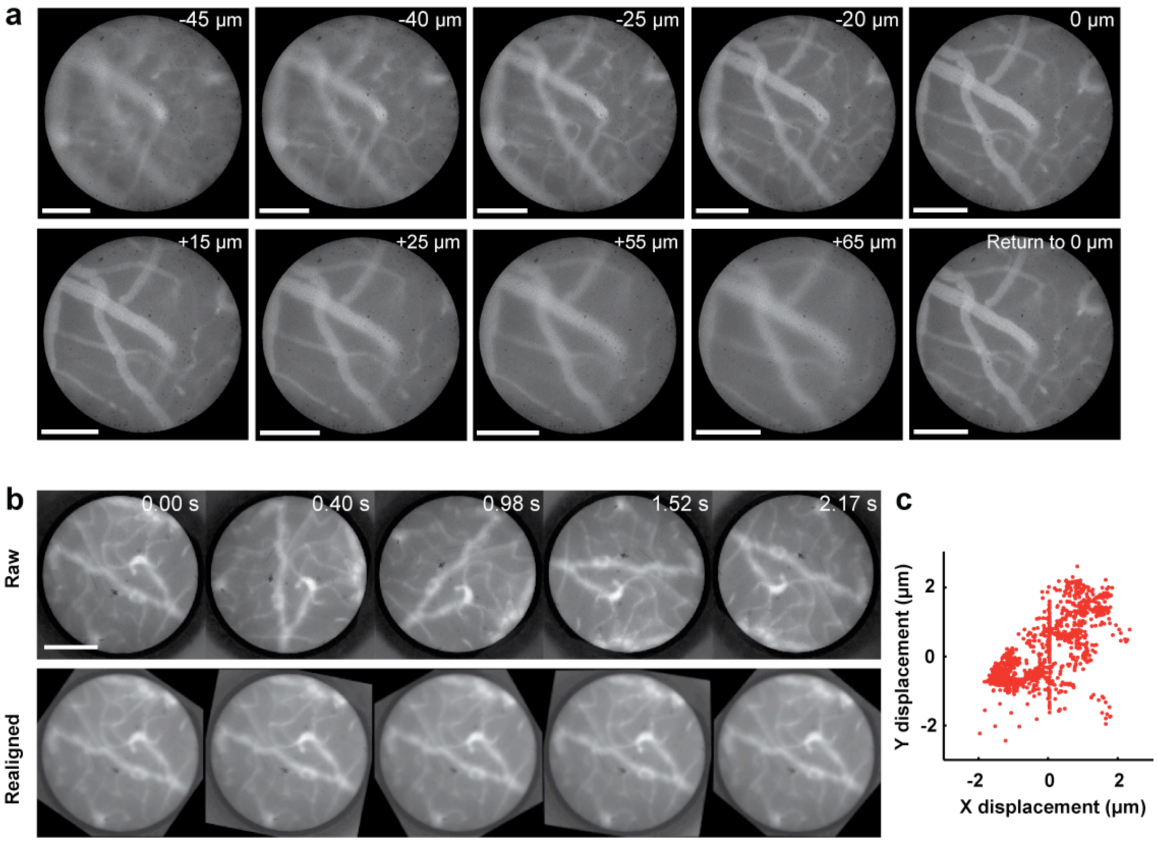

Figure 5. Aspects of basic usage of the 1.1 g epi-fluorescence microscope.

(a) Adjustment of the internal focusing lens provides focusing capability. Ten images of neocortical microvasculature acquired with the microscope's internal focusing lens placed at nine different positions within its 1.1-mm-range of movement. Over the entire focusing range the focal distance varies by approximately 110 µm in tissue. The axial positions marked in each frame denote the shift in focal plane relative to the starting position (last panel in each row). The field of view size also changes substantially, over ~240-370 µm in diameter. Since epi-fluorescence microscopy does not provide 3D optical sectioning, out of focal plane blood vessels appear blurry but do not fully disappear from view. Scale bar: 100 µm.

(b) Computational realignment of image sequences correcting for rotation of the fiber bundle. A sequence of raw image frames of neocortical microvasculature acquired at the indicated times (top), and the corresponding frames after realignment (bottom). Scale bar: 100 µm.

(c) Motion artifacts in actively moving mice are minimal after computational realignment of image sequences. Plot of the movement artifacts remaining in the two lateral dimensions after correction for bundle rotation in a behaving mouse. Each data point shows the lateral displacements of an individual image, measured as the displacement of the image content relative to the fiber bundle perimeter, from a recording of hippocampal microcirculation. The particular experiment used for this analysis had a level of motion artifact that was small in absolute terms (< 2 μm) but nonetheless atypically large relative to our usual experience. In nearly all experiments, the lateral jitter of the images within the bundle was sub-pixel in magnitude.

Figure 6. Purkinje cell dendritic Ca^{2+} activity recorded in a freely moving mouse.

(a) A trace of the rotational velocity of the mouse's head, as recorded via the optical encoder that senses rotation of the fiber bundle. Shading emphasizes when the mouse was turning its head or body.

(b) $\Delta F'/F'$ traces from the 17 Purkinje cells of Figure 3e. The traces of this figure cover the same time period shown in Figure 3, and the blue and red traces are those of Figure 3g.

MOVIES

Movies 1, 2 and 3 are available from the Nature Methods at http://www.nature.com/nmeth

SUPPLEMENTAL METHODS

Miniaturized microscope

We designed a miniature fluorescence microscope < 1.1 g in mass using ZEMAX software for optical design and Autodesk Inventor for mechanical design. The microscope housing was fabricated in Polyetheretherketone (PEEK), which is lightweight and offers exceptional stiffness and machinability. The microscope has three microlenses (Figure 1c). The first is a 2-mm-diameter coupling lens (Thorlabs, 350430-A) with a focal length of 5 mm. The second, a focusing lens, is a 2-mm-diameter Ag-ion doped gradient refractive index (GRIN) element of ~0.15 pitch length (Grintech GmbH). This lens resides in a threaded shaft movable through a 1.1 mm range, corresponding to 70-100 µm of focal adjustment at the specimen at a working distance of 300-400 µm. Using a custom planetary gearbox (8:33 gear ratio) we controlled the movements of this lens manually with a gear-tipped handtool or via an optional stepper motor (ADM 0620, MicroMo). The motor can be mounted in the housing, increasing the microscope's mass to ~2.8 g. We generally found it satisfactory to adjust the focus prior to, rather than during, imaging in the freely moving condition. Thus, we usually omitted the motor. Adjustment of the focusing lens alters the field of view diameter between approximately 240-370 µm in tissue.

The microscope allows two options for the third lens, in both cases an objective. For imaging superficial structures such as neocortex and cerebellar cortex we used a shorter objective, a 1-mm diameter Ag-ion doped GRIN lens of 0.15 pitch and 0.47 NA (Grintech GmbH). For imaging deeper structures such as hippocampus we used a longer GRIN lens of similar specifications but with an extra 0.5 pitch in length. This longer objective protrudes ~4 mm beyond the microscope housing. The microscope's internal optical magnification is approximately 1.9 –2.7×, depending on the focal position.

Optical illumination and detection pathways

Illumination from a mercury (Hg) arc lamp passes through a 1-mm-diameter multimode fiber (BFH37-1000, Thorlabs), two collimating lenses (f = 35 mm and f = 50 mm), and an excitation filter (HQ480/40x or D485/25x, Chroma). After reflecting off a dichroic mirror (FF506-Di02, Semrock), illumination is focused by the projection lens (Fluar 10× 0.5 NA, Zeiss) onto a 1.5-m-long fiber bundle (IGN-08/30, Sumitomo, NA 0.35), consisting of

30,000 fibers each with a 2.2 µm core and spaced ~3.7 µm apart. The portion of the bundle used for imaging is 0.72 mm in diameter, the outer diameter is 0.96 mm, and the bending radius is 40 mm.

The proximal end of the fiber bundle is fixed within a custom commutator that allows the bundle to rotate with only a few microns lateral precession. Below the commutator the bundle passes through an optical encoder (E4P-300-059-H, US Digital) that provides digital quadrature signals (300 cycles/revolution) to a data acquisition board (PCI-6259, National Instruments) for tracking the bundle's rotary motion. The bundle is looped to 20 cm diameter and the distal end is attached via a metal sleeve to the microscope housing. Fluorescence images collected by the microscope are focused onto the distal face of the bundle, which routes the array of pixel intensities to the proximal face. After passing back through the projection lens, the dichroic, and an emission filter (ET535/50m, Chroma), images are focused by a tube lens (f = 150 mm) onto an EM-CCD camera (iXon^{EM}+, Andor). The projection and tube lenses provide a combined 9.1 × magnification, such that the entire face of the fiber bundle is imaged onto the camera. When there is no binning of camera pixels this ensures there is no degradation of spatial resolution by the camera, which has 512 × 512 pixels, each 16 µm wide. Binning pixels degrades the resolution beyond the lateral resolution limit of ~2.8-3.9 µm imposed by the spacing between fibers and the microscope's magnification.

Computational image realignment

As the mouse torques its head or body the fiber bundle rotates within the commutator, relieving torsional strain within the bundle. The acquired images thus appear to rotate over time as the mouse behaves. To analyze cellular dynamics the raw image data stacks must be rotationally realigned. We did this using custom software written in Matlab (Mathworks).

In the first stage of realignment a binary image of the bundle, which is approximately circular, is created for each frame of the video data. After finding the bundle's centroid, we translate the image so the bundle is centered. These translations correct for precession of the bundle and are typically ~3 camera pixels in magnitude. The encoder data is then used to rotationally realign images frame-by-frame, using the first image as a reference. Due to slight inaccuracies in the encoder data, we use a second stage of realignment in which an edge detection filter is applied to the raw images to highlight small particles that inevitably accumulate on the proximal fiber bundle face. These particles can be removed, but we leave a few to guide realignment. The filtered images are rotated by linear interpolation through a sequence of angles in increments of ~2 deg. Each rotated frame is compared to a reference frame by computing the cross-correlation of the two. Each frame is then rotated by the angle that yields the greatest correlation with the reference frame. This process is repeated for the entire video 3-4 times using progressively smaller angular steps. A final computation that we performed before all subsequent analyses of biological dynamics normalizes each image by its mean intensity across the field of view. This corrects for momentary fluctuations in illumination power or slow decreases in signal due to photobleaching.

For rotational realignment to work well, it is important that image acquisition be sufficiently fast so that particles serving as fiducial markers are not blurry due to the mouse's motion. At a 75 Hz frame rate a small percentage of the images are blurry during active mouse behavior, but at 100 Hz almost all frames are sharp. To acquire images at 100 Hz our particular camera requires at least 3×3 pixel binning.

Analysis of cerebral microcirculation

Following intravascular injection of fluorescent dye, individual erythrocytes appear dark in relief [223]. Because signals from those pixels lying within vessels exhibit significant temporal fluctuations due to erythrocyte movements, computation of each pixel's standard deviation over a brief stretch of time creates an image in which the vasculature is highlighted against the background (Figure 2b,e,h). To compute a blood vessel's diameter, the cross-sectional intensity profile was fit to a Gaussian and the vessel diameter was defined as the $1/e^2$ full width.

To compute the average speed of erythrocytes passing through a chosen reference pixel in a stack of images (typically 10 s of data acquired at 100-Hz or 75-Hz frame rate), we computed the temporal cross-correlation between the fluorescence intensities at that pixel and at neighboring pixels for relative temporal delays, τ, of -1, 0, and $+1$ image frames. After setting to zero the values of all pixels lying outside vessels or exhibiting cross-correlations with the reference pixel below a minimum threshold value, we computed the centroid of each of the three cross-correlograms. The average speed of erythrocytes passing through the reference pixel was determined as the average of the distances between the $\tau = -1$ and the $\tau = 0$ centroids and the $\tau = 0$ and $\tau = +1$ centroids, divided by the time between image frames (see

Figure 2d, for which delays of $\tau = +3$ and $\tau = -3$ frames were used for clarity). To create maps of mean flow speed, cross-correlograms were computed for all pixels manually determined to be within a vessel.

For speed-diameter plots, erythrocyte speed was computed for pixel located at the peak of the Gaussian fit to the cross-sectional profile, which was usually close to the center of the vessel. To ensure unbiased sampling of vessels, the field of view was divided into a grid with squares of width 20-25 μm, approximately the widths of the largest vessels to be analyzed. No more than one vessel within each square was chosen for analysis. No analysis was performed if there was either no vessel or only a vessel junction located within the square. Pixels for which the Gaussian fit did not meet a minimum goodness-of-fit criterion or for which cross-correlogram amplitudes were below a minimum threshold were not included in Figure 2g,j.

Detection of Ca^{2+} spikes

When analyzing data from anesthetized or unrestrained mice that were not moving actively, we first computed $\Delta F(t)/F$ for each pixel, where F represents the pixel's mean level of fluorescence averaged over all image frames, and $\Delta F(t)$ is the difference between the pixel's momentary intensity level and F [32]. For active mice we corrected for slight angle-dependent non-uniformities in illumination by computing what we call $\Delta F'(t)/F'$. All frames in an image stack were grouped into angular bins based on the orientation of the fiber bundle at the time of data acquisition, as determined from realignment. For every pixel in the realigned images, the mean intensity as a function of the bundle orientation, F′, was found by averaging over all frames within individual angular bins. $\Delta F'(t)$ represents the difference between the instantaneous intensity at each pixel and the value of F′ appropriate for the bundle's instantaneous orientation. We then normalized by F′ to find $\Delta F'(t)/F'$. A potential drawback of this approach is that it relies on sufficient visitation of each orientation for best estimation of F′.

To detect Ca^{2+}spikes in Purkinje cell dendrites, we first manually drew regions of interest (ROIs) enclosing Purkinje cell dendritic trees using image sequences obtained when the mouse was anesthetized. We created control ROIs of the same size and shape as Purkinje cell ROIs by rotating the latter by 90° within the field of view. For both control and Purkinje cell ROIs we averaged $\Delta F(t)/F$ or $\Delta F'(t)/F'$ traces for all pixels within a given ROI. We performed algorithmic spike detection [224] on these traces by taking a temporal deconvolution with a decaying exponential filter (150 ms time constant) and then marking as spikes all instances of positive-going threshold crossings. To set the threshold for spike detection, we chose a threshold value such that the rate of spikes detected in error from control ROIs was both fixed (0.025 Hz) and much lower than Purkinje cell spike rates. We verified that spike rate determinations were unchanged to within our measurement accuracy when we used control ROIs obtained by rotating Purkinje cell ROIs by -90°, so that noise estimates came from different regions still perpendicular to Purkinje cell dendrites.

For comparisons of Purkinje cell Ca^{2+} activity between periods of animal movement and periods without substantial movement, we used the optical encoder signal as an indicator of the mouse's activity level. Image frames acquired while the rotational speed of the fiber bundle was > 200 deg/s were deemed to occur during animal movement. Image frames acquired while the rotational speed was < 200 deg/s were deemed to occur during inactivity. Ca^{2+} spikes were then sorted into these two behavioral states according to the image frame of their occurrence. Ca^{2+} spiking rates for each cell were determined as the total number of spikes the cell generated while the animal was in a given behavioral state (active or resting), divided by the total time the animal spent in that state.

To compute pairwise correlation coefficients, we first smoothed and down-sampled both the Purkinje cell Ca^{2+} spike trains and the optical encoder signal into time bins of 50 ms. Periods of animal activity and inactivity were discriminated from the resulting optical encoder trace by again designating 200 deg/s as the minimum rotational speed of the fiber bundle during animal movement. To assess levels of synchronous Ca^{2+} activation the correlation coefficient was computed for all pairs of cells and for each of the two behavioral states. As in prior work[46], the correlation coefficient for two cells was defined as the covariance in the activity between the two spike trains divided by the product of the standard deviations in activity for each of the individual spike trains. This yields a correlation coefficient that is bounded within [-1,1].

As a control, we computed in analogous fashion the activity rates and correlation coeficients for the control ROIs rotated by 90 deg from the Purkinje cell ROIs. Not surprisingly, we found there were slight increases in both the rates of spikes detected in error from the control ROIs and in the associated correlation coefficients during periods of animal movement, as compared to rest. However, these false positive increases were approximately ten-fold smaller in value than the increases detected in the real ROIs. Further, statistical comparison of the rate and correlation coefficient increases that occurred during movement for the real ROIs, versus those for the control ROIs, showed that the real and false positive increases came from highly distinct distributions ($P < 10^{-9}$ and $P < 10^{-25}$, for spike rates and correlation coefficients, respectively; Kolmogorov-Smirnov test). Thus, the increases during animal movement we found for both the Purkinje cell Ca^{2+} spiking rates and correlation coefficients cannot be accounted for by the much smaller increases in false positive detection.

Statistical analysis

Statistical analysis was performed in Matlab (Mathworks) and relied on non-parametric tests to avoid assumptions regarding normality. Comparisons of the distributions of Ca^{2+} spiking rates and pairwise correlation coefficients between active and inactive behavioral states were performed using the one-sided Wilcoxon signed rank test (Figure 4b). Comparisons of activity rate and correlation coefficient increases between the real and control ROIs were performed using the Kolmogorov-Smirnov test.

Animals

We used mice (5-17 weeks) of various genetic backgrounds (mainly C57BL/6, BALB/c, and CD-1). The Stanford APLAC approved all procedures involving animals.

Surgical preparation and habituation.

Mice were anesthetized with either gaseous isoflurane (1.5-2.5%; mixed with 2 L/min oxygen) or intraperitoneal injection of ketamine and xylazine (80-100 mg and 16-20 mg per kg body weight, respectively). Body temperature was monitored using a thermistor placed underneath the animal's belly and maintained at 36-37°C using a heating blanket. Depth of anesthesia was assessed by monitoring pinch withdrawal, eyelid reflex, corneal reflex, respiration rate, and vibrissae movements. The animal's skull was exposed and cleaned above the site of brain imaging.

For experiments involving hippocampal imaging, two skull screws were used to maintain attachment of dental acrylic cement to the skull. A craniotomy 2.5-2.7 mm in diameter was performed at 2.0-2.5 mm posterior to bregma and 2.0-2.5 mm lateral. The exposed cortex was kept superfused with warm Ringer's solution. The dura was removed carefully with forceps. Neocortical matter above the hippocampus was removed by aspiration. The exposed hippocampus was irrigated with Ringer's solution. A windowed-capillary (~3 mm length, 2.5 mm outer diameter, #0 end cover glass window) was inserted into the craniotomy. In some animals, low melting point agarose (1.5%) sealed the craniotomy and mechanically stabilized the capillary over the hippocampus. Alternatively, a custom silicon washer served the same purpose. A final sealing of dental acrylic was applied over exposed areas of the skull. Animals recovered from surgery for at least 1 week prior to imaging.

For experiments involving cerebellar cortex or neocortex, a custom metal plate allowing cranial access and repeated head fixation was attached to the skull with dental acrylic cement. No skull screws were used. Animals recovered from this procedure for several days. For habituation, a training dummy similar in mass to our microscope was attached to the animal's head plate during that period. On the imaging day, a cranial window (1.5-3.5 mm in diameter) was opened 1.5 mm posterior to bregma and 1.0 mm lateral (neocortex) or -6.0 to -6.5 mm posterior to bregma and 0.5 mm lateral (cerebellum). During this procedure, exposed cortex was superfused with warm artificial cerebral spinal fluid (ACSF; 125 mM NaCl, 5 mM KCl, 10 mM D-Glucose, 10 mM HEPES, 2 mM $CaCl_2$, 2 mM $MgSO_4$; pH adjusted to 7.4 with NaOH). The dura was removed carefully. For Ca^{2+}-imaging studies, the fluorescent indicator was injected into the brain (see Fluorescence Labeling).

To dampen brain motion the craniotomy was filled with agarose (1.5%; Type III-A, high EEO; Sigma) in ACSF and covered by a coverslip (Erie Scientific, 7 mm diameter circle, #1 thickness). The coverslip was kept in place by a custom metal ring; the optical window was sealed by coating coverslip edges with a thin layer of dental acrylic. Due to the short length of the GRIN objective used and the limited focusing range of our microscope, we flattened and thinned the sealing layer of dental acrylic using a high-speed drill and by scraping with a microsurgical blade.

Imaging sessions

Mice were initially kept under isoflurane anesthesia (1.5-2.5%; mixed with 2 L/min oxygen) and maintained at 36-37°C using a heating blanket. The miniature microscope was positioned above the optical window and lowered towards the brain using a translation stage until a fluorescence image of labeled surface structures was obtained at a low illumination power (30-150 μW). After locating a suitable recording site and focal depth, we turned off the illumination, fixed the head mount to the skull and/or to the metal head plate using Cerebond™ adhesive (myNeuroLab.com), and applied dental acrylic. Without Cerebond™, surface structures often shifted several tens of microns out of focus during the final hardening stage of the dental acrylic (~15-20 minutes after application), possibly due to contraction of the cement. Cerebond™ minimized this effect. When the motor is omitted the entire mass mounted on the mouse's head is ~1.8 g, including ~0.7 g for the head plate and adhesives. After the dental acrylic had hardened, mice were removed from isoflurane, placed in a round enclosure (20 cm diameter) and given at least several minutes to recover from anesthesia. We recorded mouse behavior using a video camera. In most experiments, after a session of active behavior mice were anesthetized again with isoflurane for further brain imaging.

Fluorescence labeling

For studies of microcirculation, we labeled the blood plasma by tail vein injection of 0.15-0.25 ml FITC-labeled dextran (Sigma, 2000 kDa, 10 mg/ml) [223]. For cerebellar Ca^{2+}-imaging studies, we used targeted bulk-loading of tissue with the membrane-permeant fluorescent Ca^{2+}-indicator Oregon Green 488 1,2-bis (2-aminophenoxy) ethane-N,N,N',N'-tetraacetate-1 acetoxymethyl (OGB-1-AM; Molecular Probes), performed as described[32, 225, 226]. Final concentrations of OGB-1-AM and DMSO in our pipette solution were 500 μM and 2%, respectively. To obtain a homogenous and widespread stain of the Purkinje cell and molecular layer we performed pressure ejections (3-3.5 psi pressure) using a picrospritzer (Picospritzer III, General Valve) of 1-2 min duration at different tissue depths and locations. Typically, 5-6 injections were made in a grid pattern with ~200 μm spacing.

Combined imaging and electrophysiological recordings of Purkinje cell complex spikes

Studies were performed using C57BL/6 mice under isoflurane anesthesia. Ca^{2+}-transients in distal Purkinje cell dendrites were visualized using a custom-built two-photon microscope equipped with a 40 × 0.8 NA water-immersion objective (Zeiss). Extracellular Purkinje cell activity was recorded using beveled borosilicate glass electrodes (A-M Systems, Inc.; OD, 1.0 mm; ID, 0.58 mm; internal solution, 2 M NaCl and 200-500 μM sulforhodamine 101; pipette resistance, 5-8 MΩ). Electrodes were positioned under visual control near OGB-1-AM labeled Purkinje cell somata. Purkinje cells were identified by the brief pause in simple spike activity following each complex spike. The raw electrode signal was amplified, filtered (A-M Systems, Model 1800 microelectrode AC amplifier; low cut-off, 300 Hz; high cut-off, 5 kHz), digitized (20 kHz) and stored on disk for off-line analysis. Two-photon imaging data of dendritic calcium activity was recorded simultaneously (frame rate, 20.35 Hz) in superficial regions of the molecular layer (depth, 20-50 µm) located directly above the microelectrode tip. The correspondence between dendritic Ca^{2+} spikes and electrophysiological complex spike activity was verified through a spike-triggered average analysis written in Matlab and agrees with prior combined optical and electrophysiological studies of Purkinje cells [219].

Chapter 4: Motor behavior activates Bergmann glial networks

ABSTRACT

Although it is firmly established neuronal activity is a prime determinant of animal behavior, relationships between astrocytic excitation and animal behavior have remained opaque. Cerebellar Bergmann glia are radial astrocytes that are implicated in motor behavior and exhibit Ca^{2+}-excitation *in vitro*. However, Ca^{2+}-excitation in these cells has not previously been studied in live animals. Using two-photon microscopy we found that Bergmann glia exhibit three forms of Ca^{2+}-excitation in awake behaving mice. Two of these are ongoing within the cerebellar vermis even when animals are resting. During locomotor performance concerted Ca^{2+}-excitation arises in networks of at least hundreds of Bergmann glia extending across several hundred microns or more. Concerted Ca^{2+}-excitation was abolished by anesthesia or blockade of either neural activity or glutamatergic transmission. Thus, large networks of radial astrocytes can be activated by specific animal behaviors and undergo excitation of sufficient magnitude to potentially initiate macroscopic changes in brain dynamics.

INTRODUCTION

Neurons have long been studied in live animals by electrophysiological techniques, but astrocytes' lack of electrical excitability has made these cells' relationships to animal behavior more challenging to uncover. Nonetheless, since the discovery of astrocytic Ca^{2+}-excitation *in vitro*, multiple processes that involve astrocytic Ca^{2+}-signaling have emerged. In adults these include modulation of neuronal activity and regulation of cerebral blood flow [36, 227]. Recently, a few *in vivo* Ca^{2+}-imaging studies have begun to reveal roles played by astrocytic Ca^{2+}-excitation in live animals. These pioneering studies have provided initial glimpses of how astrocytes respond to sensory stimulation and modulate microcirculation

[33, 34, 228, 229]. However, the current *in vivo* imaging data generally share two limitations.

First, results were obtained in anesthetized animals, and it is unexplored whether there are specific animal behaviors that elicit astrocytic Ca^{2+}-excitation. Anesthesia precludes animal behavior, and neuronal responses to sensory stimuli can be quite distinct in the anesthetized and awake brain [39, 230, 231]. The strong possibility exists astrocytes will show equally divergent dynamics across brain states. Since astrocytic dynamics may reflect aspects of neuronal activity [31, 33, 34, 228], neurons' dependence on anesthesia makes this possibility all the more probable.

A second limitation is that *in vivo* studies have typically not distinguished among astrocytic sub-types. It is widely understood that different neuron types, even in the same brain area, can show grossly dissimilar activity patterns. Like neurons, astrocytes display considerable non-uniformity, so astrocytic dynamics might also vary between brain areas and layers [227, 232, 233]. Throughout the adult brain, a basic distinction exists between the radial astrocytes, which include the Bergmann glia of the cerebellar cortex and the Müller cells of the retina, and the 'star-shaped' astrocytes of the neocortex and other areas. These two classes exhibit markedly different morphologies and patterns of Ca^{2+}-excitation as seen *in vitro* [227, 233], but Ca^{2+}-excitation in radial astrocytes has not previously been studied in live animals.

Motivated both by the recent imaging progress and the above limitations, we studied Ca^{2+}-excitation within radial astrocyte networks in awake behaving mice. We focused on a basic behavior, locomotion, towards ascertaining whether the Bergmann glia have behaviorally triggered dynamics. These cells lie in areas crucial for motor coordination, are implicated in motor behavior [234], form gap-junction coupled networks, and can be identified by their palisade arrangement and status as the sole astrocytes in the cerebellar molecular and Purkinje cell layers. We developed methods for two-photon imaging in the cerebellum of awake, head-restrained mice allowed to move on an exercise ball and studied how wakefulness and motor behavior affect Ca^{2+}-excitation as it normally occurs in Bergmann glial networks.

We found that Bergmann glia of the cerebellar vermis, an area implicated in locomotor coordination, exhibit three forms of Ca^{2+}-transients in awake mice. One of these, which we named 'Ca^{2+} flares', was triggered during locomotion and extended over macroscopic domains at least hundreds of microns across. Flares involved large networks of astrocytes containing at minimum hundreds of Bergmann glial fibers. The other two forms of

Ca^{2+}-excitation, which we named ‘sparkles’ and ‘bursts’, persisted in awake but resting animals. Sparkles appeared restricted to individual fibers. Bursts were radial waves expanding in three dimensions across the fibers of ~10-40 Bergmann glia. Anesthesia markedly disrupted Bergmann glial Ca^{2+}-excitation, showing that studies in awake animals are essential. The dependence of flares on locomotion emphasizes the need to study behaving subjects.

RESULTS

Imaging Bergmann Glia in Awake Mice During Rest and Locomotion

Our first step was to show we could visualize glial Ca^{2+} transients in behaving mice without undue levels of motion artifact. To achieve this we used two-photon microscopy in awake, head-restrained mice allowed to run voluntarily on an exercise ball (Figure 1A). After habituating a mouse to periods of head-restraint, we created an optical window in the cranium, injected the membrane-permeant fluorescent Ca^{2+}-indicator Oregon-Green-488-BAPTA-1-acetoxymethyl (OGB-1-AM) into cerebellar cortex [235], and performed Ca^{2+}-imaging across areas 200-500 μm wide with the cranium held secure. With this approach we studied lobules V and VI of the vermis (Figure 1B), areas implicated in coordinating hind and forelimbs [236, 237]. Fluorescently labeled Bergmann glia were identified by the distinct palisade arrangement of their fibers in the molecular layer, the size and location of their somata in the Purkinje cell layer, their slow Ca^{2+}-transients, and by counterstaining with the red fluorescent astrocyte marker sulforhodamine101 (SR101) [238-241]. SR101 images revealed both Bergmann glial somata and fiber palisades (Figures 1C and S1). Since SR101 is a non-functional morphological marker, we also used it to characterize the motion artifacts in our image data.

When the mice were not driving the exercise ball, motion artifacts were essentially negligible (± 0.4-0.6 μm, lateral s.d.; ± 1.5 μm, axial). During mouse locomotion there were small, transient lateral displacements (± 0.7-0.8 μm; Figure 1D) that were correctible by offline image registration. Artifacts due to axial displacements (± 3.4 μm) were largely mitigated by the coinciding alignments of the optical axis and Bergmann glial fibers in the dorsal-ventral direction. Crucially, within SR101 images we saw no characteristic phenomena resembling the forms of Ca^{2+}-excitation seen repeatedly in OGB-1-AM images from all 33 alert mice studied. Having achieved a suitable imaging assay, we next characterized these forms of Ca^{2+}-excitation in detail.

Ca^{2+} Sparkles are Ongoing in Awake Mice and Appear in Individual Glial Fibers

When mice were on the exercise ball but voluntarily sitting still, we observed two forms of spontaneous Ca^{2+} excitation in Bergmann glia. The first of these we termed Ca^{2+} sparkles due to their highly localized appearance.

Sparkles occurred in the main Bergmann glial fibers of the palisades that were co-labelled by SR101 (Figures 2A-D). Sparkles were modest in amplitude, with relative fluorescence increases ($\Delta F(t)/F$) sometimes just above background fluctuations. To determine the rate and time course of sparkles we performed a conservative analysis by focusing on events with >10% increase in $\Delta F(t)/F$, which could be readily identified. Within this event subset (13.2 ± 4.1% peak $\Delta F(t)/F$; mean ± s.d.; $n = 4285$ events), sparkles had variable duration (3.3 ± 3.9 s) and occurred sparsely (4.5 ± 2.6 Hz/mm^2; $n = 15$ animals) (Figures 2B and 2D). By considering the density of glial fibers and the proportion of the molecular layer sampled in one of our recordings, we estimated individual glial fibers have mean sparkle rates between ~0.01-0.1 min^{-1}. This is comparable to spontaneous Ca^{2+}-excitation rates for astrocytes in other brain areas [31, 232].

Based on sparkles' localization and the anatomical association between neuronal synapses and fine protrusions from Bergmann glial fibers, sparkles might represent activation of Bergmann glial microdomains in response to synaptic activation [241]. Consistent with this idea, sparkle rates were reduced by local blockade of either neuronal activity using the Na^+-channel blocker tetrodotoxin (TTX) (likelihood ratio test; $p < 10^{-6}$ in each of $n = 4$ mice) or glutamatergic transmission using the broad-spectrum glutamate receptor antagonist γ-D-glutamylglycine (γDGG) ($p < 0.002$ in each of $n = 3$ mice). Pyridoxalphosphate-6-azephenyl-2',4'-disulfonic acid (PPADS), a non-selective P2 purinergic blocker, had no significant effects on sparkle rates. However, in mice anesthetized by isoflurane, sparkle rates fell markedly ($p < 10^{-5}$ in each of $n = 3$ mice).

Ca^{2+} Bursts Involve up to ~40 Bergmann Glial Cells in Alert Mice

Like sparkles, bursts occurred spontaneously in mice resting voluntarily on the exercise ball. Bursts appeared as radially expanding waves that were not restricted to SR101-

labeled palisades and spread over approximately ellipsoidal regions with greater medio-lateral (55 ± 37 μm) than rostro-caudal (44 ± 33 μm) extent ($n = 63$ events; Wilcoxon signed rank test, $p < 0.001$) (Figures 2E and 2H). By using three-dimensional (3D) two-photon imaging to monitor multiple tissue depths with near simultaneity [242], we found bursts originated across all depths of the molecular layer and had an axial extent of 56 ± 23 μm (Figures 2E and 2G). Typical burst volumes (86,000 ± 61,000 μm^3; $n = 10$) implicate involvement of many fibers, from on average ~10 but up to ~36 Bergmann glia based on these cells' somatic density [238, 239]. Peak $\Delta F(t)/F$ increases for bursts (13 ± 4.8 %; $n = 96$ events in 11 subjects) were similar to those for sparkles, but the two phenomena were unmistakably distinct since bursts occupied much larger volumes. Like sparkles, bursts were infrequent (126 ± 173 mHz/mm^2), but rates varied substantially between mice (range: 3.2 – 566 mHz/mm^2).

Bursts displayed a time course (Figure 2F) similar to those of Bergmann glial Ca^{2+}-transients previously recorded in brain slices in response to parallel fiber stimulation [243, 244]. The similarity suggests both phenomena might involve ATP-triggered release of Ca^{2+} from intracellular stores, as found *in vitro*. Consistent with this, bursts were eliminated by PPADS ($p < 0.01$ in each of $n = 4$ mice). However, bursts persisted after application of either TTX, γDGG, or isoflurane anesthesia, none of which produced reliable changes in burst frequency.

The burst time course conformed closely ($r = 0.9 \pm 0.1$, s.d.; $n = 96$ events) to a convolution of two exponentials, with one time constant each for both the rise (1.5 ± 0.7 s) and decay (1.8 ± 1.2 s) of the Ca^{2+} signal (Supplemental Methods). The rise and decay time constants for each burst were highly correlated, with 81% of events showing a negligible (<1%) difference between the two. This form most simply arises from two biochemical events occurring in succession at approximately equal rates. One of these could be the diffusion of a small molecule trigger such as ATP, a possibility consistent with bursts' expansion from a central initiation point. In the presence of γDGG, bursts' decay times significantly lengthened (9.6 ± 10 s; $p < 0.001$ Wilcoxon rank sum) as compared to all other pharmacological conditions (Figure 2F). Rise times under γDGG (2.2 ± 1.2 s) were not significantly affected. Given that bursts persist under γDGG, glutamate might play a role in bursts' termination but perhaps not in their initiation.

Motor Behavior Activates Ca^{2+} Flares Across Bergmann Glial Networks

When mice ran voluntarily on the exercise ball, Ca^{2+}-excitation arose in large networks of Bergmann glia. These events, which we termed flares, occurred almost exclusively at locomotor onset (n = 187 events) and appeared to be partly composed of a large number of sparkles. Flares often extended across our entire field of view (200-500 µm), and typically involved hundreds of Bergmann glial fibers (Figure 3).

Because astrocytes have been reported to show Ca^{2+} responses to mechanical stress, we were initially concerned flares might be artifacts due to transient mechanical perturbation to the brain at running onset. However, multiple lines of evidence and control experiments were inconsistent with this possibility, leading us to conclude flares are part of normal brain physiology. First, flares show a completely different time course and are an order of magnitude more rapid than the Ca^{2+} signals reported to occur in response to mechanical stress [245]. Second, flares occurred in our movie data even when there was no perceptible movement of the brain. Third, in cases when there were measurable, micron-scale brain displacements at motor onset, we found no correlation between the magnitude of tissue displacements and either the occurrence or amplitude of Ca^{2+} flares. Fourth, the occurrences of flares were independent of mechanical factors such as the craniotomy diameter (0.5-3.0 mm), whether the dura mater was removed or the craniotomy covered with agarose and a cover slip. Fifth, we never saw flares in control experiments with awake mice in which we applied pressure on the animal's head holder or optical window. This was so even when the applied pressure caused far greater tissue displacements than those occurring during running-evoked flares in the same mice. Sixth, flares were abolished by blockade of neuronal Na^{+} channels or glutamatergic transmission, showing that flares require neuronal activity and do not arise from direct mechanical activation of individual Bergmann glia. Seventh, we saw during flares no signs of spreading neuronal depression, a known effect of mechanical injury to the brain (Charles et al., 1991), as assessed by imaging neuronal Ca^{2+} activity in Purkinje cells and molecular layer interneurons. Thus, we found no support for the idea flares were mechanically induced.

Among flares' interesting properties was their mode of initiation. Flares did not seem to involve a stereotyped manner of propagation, but usually arose similarly across multiple locomotor events within individual fields of view. At flare onset we typically saw a smaller number of fibers within the palisades exhibiting Ca^{2+} increases that resembled sparkles. The number of sparkling fibers rose over 4.0 ± 1.5 s until reaching peak activation (Figure 3D). The rate and extent of sparkling varied significantly between flares. Only in 6.9 ± 5.9% (*n*=87) of cases did flares of significant extent and sparkling rates (>100 Hz/mm^2) occur without measurable rotation of the exercise ball.

We initially wondered if isolated sparkles might be spontaneously occurring, 'miniature' flares, perhaps analogous to the miniature post-synaptic potentials that represent spontaneous neurotransmission. Flares showed a similar pharmacological profile as sparkles and were eliminated by TTX or γDGG (Figure 4A and B) even while locomotor activity persisted. In isoflurane-anesthetized mice we never saw flares, even if we moved or electrically stimulated the animals' limbs. PPADS did not affect flares (Figures 4A and 4B). However, by comparing the SR101 and OGB-1-AM data channels we saw that Ca^{2+} increases during flares occurred both in the palisades and surrounding areas where finer fibers lie (Figures 3D and 3E). This does not eliminate the possibility flares might be seeded by or partly composed of sparkles, but does show flares are not simply composites of multiple sparkles (Discussion).

We next characterized the spatial magnitude of flares, because events involving widespread activation might be more likely to have a role in animal behavior or large-scale brain dynamics. 3D imaging showed flares occurred over at least ~$0.5 \times 0.5 \times 0.1$ mm^3 (our maximum sample volume) and reached significant depths of the molecular layer (Figure 3E), sometimes involving Bergmann glial cell bodies. Thus, flares are macroscopic and cover several cerebellar microzones, divisions of the cerebellar cortex that map to parts of the body [246]. Given this size and their behaviorally triggered character, flares seem well poised to influence neuronal or cerebral blood flow dynamics associated with motor behavior (Discussion).

Ca^{2+} Flares Have Refractory Dynamics That Depend on Running History

As an initial investigation of flares' potential relevance to locomotor behavior, we analyzed the temporal relationships between the two. We found a consistent order of events and a refractory character to flares' occurrences, which restrict how flares might relate to the modulation of cerebral blood flow during behavior (Discussion).

Flares nearly always occurred after movement of the exercise ball and never before (Figures 4B-D). Flares also had limited duration, regardless of how long the mouse's locomotor episode lasted (Figure 4C). If >10 s had elapsed since termination of the prior episode, flares reached peak fluorescence 4.4 ± 2.5 s (*n* = 93 events) after re-initiation of movement. Ca^{2+} increases during flares were spread over 9.7 ± 3.1 s (*n* = 110 events) (Figure 4C). Thus, if locomotion was brief (<4 s) flares outlasted the movement, but *vice versa* if motor activity persisted (>16 s). Flares were also refractory and arose more reliably at motor onset when there had been no prior recent movement.

Thus, flares tended to accompany the initial but not the later movements within series of running bouts occurring in quick succession (Figure 4B). The percentage of cases in which flares occurred at motor onset rose from ~7% when the animal had moved in the previous 4 s (*n* = 41), to over 93% if >32 s had elapsed since termination of the prior movement (*n* = 75) (Figure 4D). These refractory time scales are reminiscent of the refractory character of cerebral blood flow up-regulation [247, 248], suggesting a potential link between the two (Discussion).

DISCUSSION

We conducted the initial study of radial astrocyte Ca^{2+}-excitation in the live adult brain. In Bergmann glia we found three types of Ca^{2+}-excitation: flares, bursts, and sparkles. The modes of initiation and pharmacological data suggest sparkles and flares could share some common mechanisms distinct from those of bursts. Sparkles and flares were affected by isoflurane, as well as blockades of either neuronal activity or glutamatergic transmission, but not of purinergic transmission. Bursts require purinergic signaling but neither Na^{+} channel activation nor glutamatergic transmission. Strikingly, at their initiation flares appear to be composed of multiple sparkles. Thus, it is tempting to speculate sparkles could be seed or miniature flares.

This speculation must be tempered on several grounds. Sparkles occur in animals at rest and persist at low rates under TTX or isoflurane, whereas the neural activity patterns needed to initiate flares appear specific to those for motor behavior. Further, flares activate both the main glial fibers in palisades and surrounding fine fibers. Our recordings do not have the resolution to eliminate the possibility sparkles activate fine fibers closely adjacent to the palisades, but the data show conclusively the tighter spatial association between sparkles and palisade fibers. Flares activate the palisade fibers along large portions of their axial extent, but sparkles have more limited axial size.

Regardless of whether sparkles are precisely limited to the palisades, it remains plausible a smaller number of localized events, sparkles, might together initiate a composite macroscopic event, a flare. This scenario prompts many questions for future study, including how many sparkles constitute a flare and under what conditions can sparkles become flares? Overall, network Ca^{2+} activation during flares contrasts with that of bursts, which appear to have a single initiation site, use different receptor pathways, and cover only tens of cells.

Isoflurane Anesthesia Disrupts Bergmann Glial Ca^{2+}-excitation

Recent studies have explored Ca^{2+}-excitation in neocortical astrocytes of anesthetized animals [31, 33, 34, 228, 229, 232, 240]. Our studies focused on Bergmann glia, but a comparison of our recordings from awake mice to those from the same fields of view while the mice were under isoflurane anesthesia provides reason to be cautious in interpreting any results on astrocytic dynamics from anesthetized subjects. Isoflurane exerts direct and indirect effects on astrocytes, influencing glutamate uptake and gap-junction-mediated signaling [249, 250]. We found isoflurane disrupts Bergmann glial Ca^{2+}-excitation by eliminating flares and reducing sparkle rates. These results suggest sparkles and flares are each promoted by specific forms of neural activity. Hence, wakefulness and animal behavior seem to be key influences on the excitation of glial networks.

Ca^{2+} Excitation Differs in Bergmann Glia and Protoplasmic Astrocytes

A few studies in anesthetized animals have found spontaneous and sensory-evoked Ca^{2+} transients in protoplasmic astrocytes of the neocortex The spontaneous transients were temporally sparse, generally confined to individual astrocytic processes, and rarely involved cell bodies [31, 33, 232]. Sensory-evoked transients were also restricted to individual astrocytes, but in some cells arose in both processes and somata [33, 34]. Sensory-evoked excitation needed glutamatergic transmission [33, 228] but spontaneous transients were unaffected by TTX and PPADS [232]. In contrast, we found Bergmann glia fibers show both highly localized and two forms of spreading Ca^{2+} transients. In Bergmann glial somata, Ca^{2+} transients were less frequent ($\sim 8.9 \pm 3.7$ mHz/mm^2) than in fibers but involved large $\Delta F(t)/F$ increases (22 ± 12 % ; $n = 12$), were localized to one or a few neighboring somata in resting animals, and sometimes spread among several cell bodies during flares.

A recent study in the neocortical S1 hindlimb area of awake mice reported subpopulations of astrocytes that show Ca^{2+} activity correlated with running behavior [251]. These Ca^{2+} increases lasted several seconds, arose ~1-2 s after running onset, and were reported to be qualitatively similar to those seen in anesthetized animals. In comparison, locomotor-evoked cerebellar flares had no analogs in anesthetized mice, variable delays relative to running onset and relatively consistent durations.

Our results highlight the need to compare astrocytic physiology across brain regions, between astrocytic sub-types, and under anesthesia to during animal behavior. Unlike protoplasmic astrocytes, Bergmann glia have a radial, not a star-shaped morphology, which confers distinct cellular organization and functional properties [227, 233]. Our results also show the utility of 3D imaging, which allowed us to reconstruct burst volumes and assess the axial extents of sparkles and flares. Comparisons of 3D data between areas and layers will help reveal the diversities of roles astrocytes might play in brain dynamics.

Possible Functions of Sparkles and Bursts

In vitro studies have shown Bergmann glia can respond to neuronal activity with intracellular Ca^{2+} transients and that Ca^{2+} signaling plays a key role in regulating structural and functional interactions between these cells and cerebellar neurons [233, 252]. It has been unclear if these responses evoked *in vitro* have analogs *in vivo* and if Bergmann glia exhibit spontaneous Ca^{2+} transients in live animals. We found Bergmann glia exhibit at least two types of spontaneous Ca^{2+} excitation, sparkles and bursts.

Given their spatial localization, sparkles might represent activation of Bergmann glial microdomains in response to synaptic activity [241]. Since sparkles persist at reduced rates under TTX, γDGG, and isoflurane, they might partly be of intrinsic origin [232, 253]. The low rate of sparkles we report is comparable to rates of spontaneous Ca^{2+} activity seen *in vivo* in neocortical astrocytes [31, 232]. Due to our conservative detection threshold our sparkle rate estimate should be viewed as a lower bound (Supplemental Methods).

Bursts occur both within and outside the main glial fibers of the palisades and cover fibers from typically ~10 but sometimes up to a few tens of Bergmann glia. The radially expanding waveform of bursts suggests a diffusible messenger could mediate these events and indicates bursts involve finer glial processes. Bursts' ellipsoidal volumes suggest the diffusion of any such messenger is anisotropic, which is plausible given the ordered, anisotropic architecture of cerebellar cortex.

ATP is a likely candidate as a messenger, since blockade of P2 purinergic receptors eliminates bursts. The burst time course closely resembles that of large amplitude Ca^{2+} transients evoked in Bergmann glia upon parallel fiber stimulation in brain slices. These evoked transients also involved purinergic receptor activation and release of Ca^{2+} from intracellular stores [243, 244]. Given these similarities, we interpret the ongoing bursts we saw *in vivo* as the physiological analogs of the evoked transients recorded *in vitro.* Using 3D imaging we found bursts spread over an approximately ellipsoidal volume of ~86,000 mm^3, yielding estimates that bursts traverse 55,000-70,000 parallel fiber-Purkinje cell synapses from ~3 adjoining Purkinje cells [254, 255]. However, *in vitro* studies to date have not found any influence of electrically evoked bursts upon parallel fiber to Purkinje cell synaptic transmission [243]. We saw considerable variability in burst rates between individual mice, and future work must identify the origins of this variability and the role of bursts in brain function.

Relationship of Flares to Motor Behavior and Hemodynamics

Prior work *in vitro* and in anesthetized animals has shown the involvement of astrocytic Ca^{2+} signaling in hemodynamic modulation [34, 228, 229, 256-258]. This suggests a possible link between flares and blood flow regulation. Thus, quantifying the kinetic relationships between locomotion, hemodynamics, and flares is essential for logically constraining the roles flares could play. Time scales of cerebral hemodynamic responses have been well reported [247, 248, 259], but we also confirmed basic kinetic features of the responses that occur in our mouse locomotor assay by using laser Doppler measurements of blood perfusion levels in the vermis (Figure S2, Supplemental Methods). Taken with our kinetic studies of flares and motor behavior, the data permit several conclusions about flares' possible role in functional hyperemia.

First, flares' ~5 s rise time following locomotor onset is consistent with a possible role for flares in initiation of blood flow responses. In our locomotor assay flares and local blood perfusion levels rose over a very similar time course (Figure S2). Second, flares' limited duration contrasts with the well documented finding that during motor activity the resulting hemodynamic up-regulation persists throughout the behavior [260]. Laser Doppler data confirmed this persistence in our mouse locomotor assay (Figure S2). Thus, flares do not have a role in sustaining hemodynamic responses, even if flares aid their initiation. Third, flares' refractory dynamics are reminiscent of the refractory character of hemodynamic regulation [247, 248]. Taken together, our data imply a potentially intricate relationship between flares and animal behavior.

One possibility is flares occur in direct response to the neural activity underlying locomotion. Since Bergmann glia, like other cortical astrocytes, contact neuronal synapses and blood vessels, flares might mediate neural-driven blood flow increases. Another possibility is blood flow increases, not neural activity, trigger flares. Thus, flares might provide retrograde signaling from blood vessels to neuropil [261]. A third possibility is flares are part of a dynamical interplay between neurons and astrocytes that mediates animal behavior. For example, flares might cause Bergmann glia to release neuro-active molecules, in turn affecting the neural activity underlying locomotion. Other possibilities involve combinations of these three. To dissect these issues, microscopy in behaving mice must be combined with targeted manipulations of Bergmann glia, such as by genetic or optogenetic means.

FIGURES

Figure 1. Two-Photon Microscopy in Behaving Mice

(A) Two-photon microscopy in awake, head-restrained mice using an exercise ball. Mouse locomotion was tracked by video and an optical encoder on the ball.

(B) Areas of lobules V, VI of the cerebellar vermis examined in this study (pink). Scale bar: 1 mm.

(C) Bergmann glial processes and somata in the molecular (upper) and Purkinje cell layers (lower), respectively, were labeled with the Ca^{2+}-indicator OGB-1-AM and identified based on morphology and counterstaining with SR101, an astrocyte marker. Scale bars: 25 µm.

(D) Probability densities of image displacements in dorso-ventral, *z* (blue), medio-lateral, *x* (green), and rostro-caudal, *y* (red), dimensions. Insets show example traces of displacements in *x*-*y* and *x*-*z* planes during locomotion. Scale bars: Upper, 50 mm/s; Lower, 10 s and 5 µm.

Figure 2. Ca^{2+}-Excitation is Ongoing in Awake Resting Mice

Ca^{2+}-sparkles (A-D, red traces) appear in individual Bergmann glial fibers. Ca^{2+}-bursts (A-B, E-H, green traces) spread as 3D radial waves over multiple Bergmann glia.

(A) Cellular structures in the molecular layer labeled with OGB-1-AM. Ca^{2+} dynamics were analyzed in regions of interest where fibers within palisades (yellow) had been identified by SR101 co-staining. A selection of fibers showing asynchronous sparkles and a burst are marked red and green, respectively. Scale bar: 50 μm.

(B) Traces of relative increases in OGB-1-AM fluorescence, Δ*F* (*t*)/*F*, from the 14 regions marked red in (A) and sorted by the onset time of Ca^{2+}-excitation. Scale bars: 5 s, 10% Δ*F* (*t*)/*F*.

(C) Image frames showing a Ca^{2+} sparkle. Scale bar: 10 μm.

(D) Overlay of average (red) and individual sparkle time courses (black). Scale bars: 5 s, 10% Δ*F* (*t*)/*F*.

(E) Images from a 3D recording showing Δ*F* (*t*)/*F* signals at different depths below the pia. Scale bar: 50 μm.

(F) Top, average (green) and individual (black) burst time courses in awake mice. Bottom, averaged and normalized traces under four conditions: isoflurane-anesthetized (blue), awake resting (green), TTX- (red) and γDGG-treated (orange). Scale bar: 5 s.

(G) Top, locations and volumetric profiles of bursts detected by 3D imaging during an 8.2 min period. Colored outlines show extent of Ca^{2+} spread for each burst. Bottom, z-projection image.

(H) Bursts have greater extent parallel to the long axis of the cerebellar folium, which lies perpendicular to the parasagittal planes of cerebellar cortex. Colored data points denote individual bursts recorded under the conditions noted in (F).

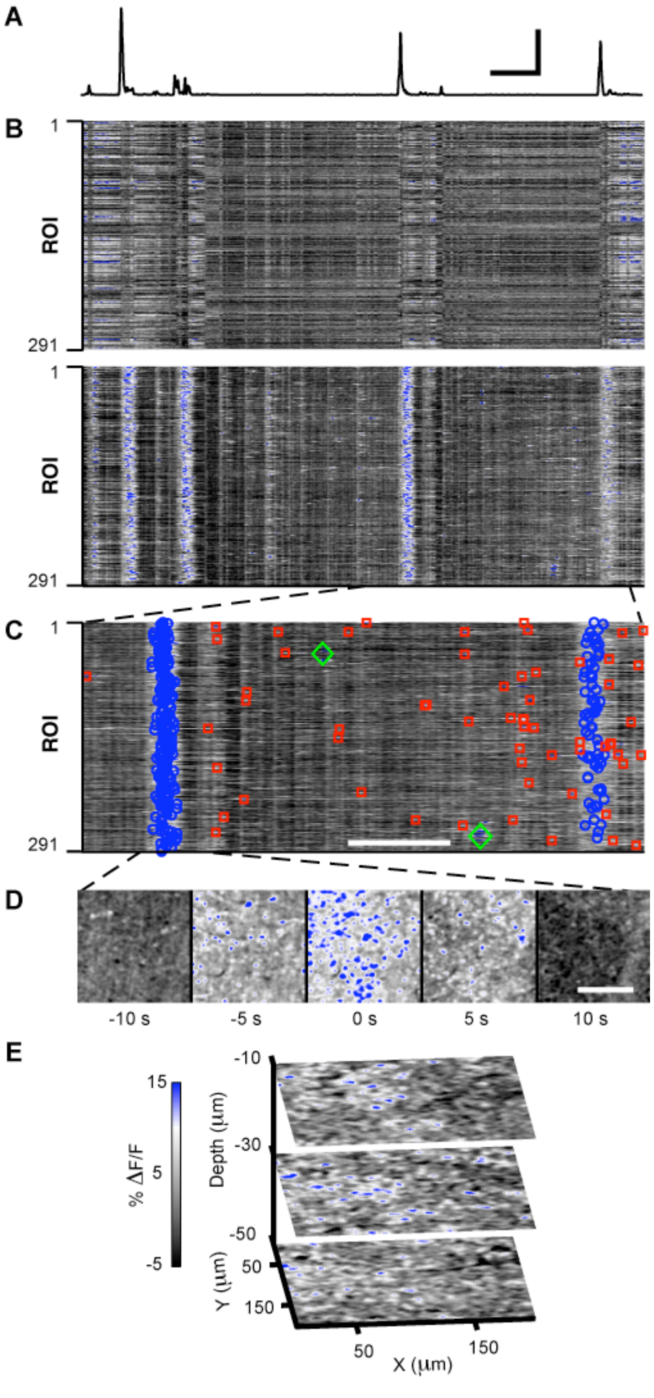

Figure 3. Motor Behavior Evokes Ca^{2+}-Excitation Across Bergmann Glial Networks

(A) Encoder trace of rotational speed of the exercise ball, showing multiple running bouts during a 9.8 min period. Scale bars: 50 s and 20 mm/s; the former also applies to (B).

(B) Top, SR101 $\Delta F\ (t)/F$ traces from 291 glial fiber regions during the same time interval as in

(A). Fibers were sorted by x coordinate. Bottom, the same regions in OGB-1-AM images reliably exhibit concerted $\Delta F(t)/F$ increases during running. Signals of similar magnitude and time course are absent in the SR101 data.

(C) OGB-1-AM $\Delta F(t)/F$ traces over the interval indicated in (B) by dashed lines. Flares are marked blue and are interleaved in time with sparkles (red) and bursts (green). Scale bar: 50 s.

(D) Image frames of OGB-1-AM $\Delta F(t)/F$ increases during the flare indicated by dashed lines in (C). Scale bar: 100 µm.

(E) Images from a 3D recording of $\Delta F(t)/F$ increases during a flare.

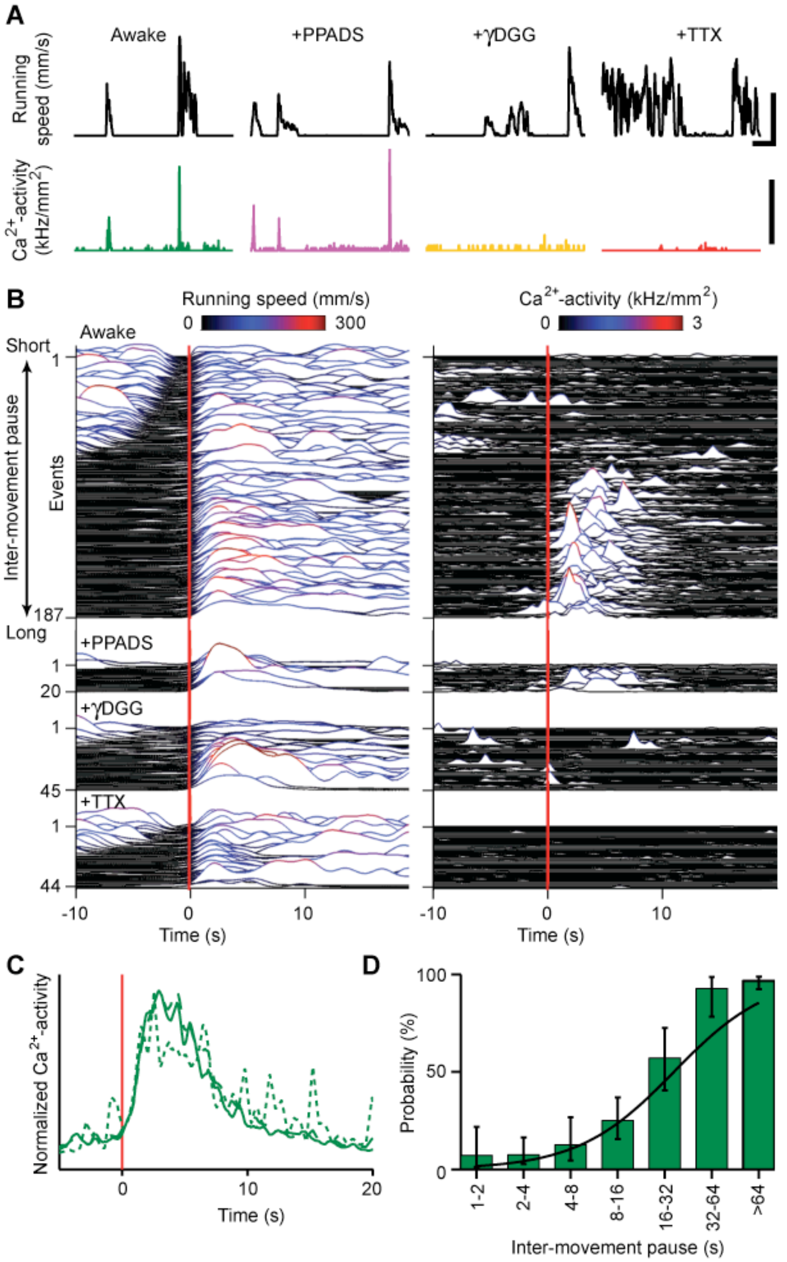

Figure 4. Locomotor Evoked Ca^{2+}-Excitation

(A) Traces revealing Ca^{2+} flares (bottom) as a function of locomotor activity (top), under four conditions, awake-untreated (green), PPADS- (purple), γDGG- (orange) and TTX-treated (red). Ca^{2+} activity was quantified by rate of glial fiber activation per unit area. Upper scale bars, 50 s and 150 mm/s. Lower scale bars, 1 kHz/mm^2.

(B) Traces of running speed (left) and Ca^{2+}-flares (right) triggered at movement onset (red lines) in untreated, PPADS-, γDGG-, and TTX-treated mice. For each condition, traces are arranged by the pause duration after the prior movement.

(C) Traces showing normalized time course of flare activation triggered upon movement onset (red line), for all (solid), brief (dotted), and extended movements (dashed).

(D) Probability of observing a flare at locomotor initiation versus the pause duration between movements. Flares have a refractory period (compare panel (B)).

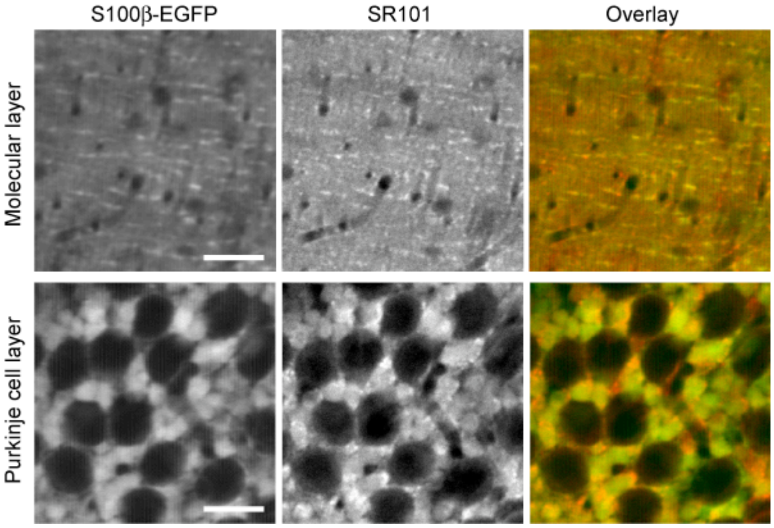

Figure 5. The Astrocyte Marker SR101 Stains Bergmann Glial Cells in the Cerebellar Cortex

Example images acquired by two-photon microscopy in anesthetized mice show co-staining (right) of Bergmann glial cell structures in the molecular (upper row) and Purkinje cell layer (lower row) after topical application of the red fluorescent dye SR101 (center) in transgenic mice expressing EGFP in astrocytes under control of the human S100β promoter (left). Scale bars: 25 μm.

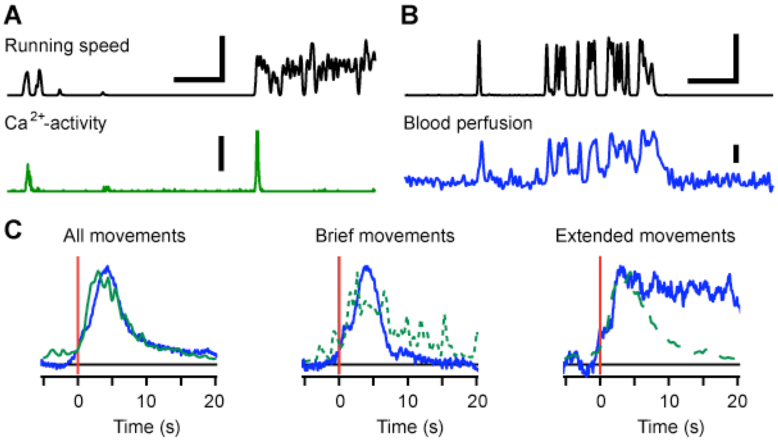

Figure 6. Motor Behavior Evoked Ca^{2+}-Excitation is Accompanied by Regional Increases in Blood Perfusion Levels

(A) Concurrent example traces showing Ca^{2+}-flare activity (bottom) and mouse running speed (top), for brief and extended running episodes. The green trace shows flare magnitudes, as quantified by the rate of Ca^{2+}-activation in Bergmann glial fiber regions per unit area. Upper scale bars, 50 s and 150 mm/s. Lower scale bar, 1 kHz/mm2.

(B) Concurrent example traces showing relative changes in regional blood perfusion levels as measured by laser Doppler flowmetry (bottom), and mouse running speed (top), for brief and extended running episodes. Upper scale bars, 50 s and 150 mm/s. Lower scale bar, 15%.

(C) Overlay of the normalized Ca^{2+}-flare activation (green) and blood perfusion increase (blue) triggered upon movement onset (vertical red lines) for all (left), brief (center), and extended movements (right), revealing similar kinetics during locomotor initiation. The solid, dotted, and dashed green traces are the same as those in Figure 4C.

EXPERIMENTAL PROCEDURES

Animals and Surgery

All procedures were approved by the Stanford APLAC. We used male C57BL/6 mice (6-10 weeks old). Surgery was done under isoflurane (1.5–2.5%, mixed with 2.0 L/min O_2). Body temperature was monitored with a thermistor underneath the torso and maintained at 36-37°C using a heating blanket. We assessed depth of anesthesia by monitoring pinch withdrawal, eyelid and corneal reflexes, breathing rate, and vibrissae movements.

1–2 weeks prior to imaging the skull was exposed and cleaned above neocortex and cerebellum. A custom metal plate allowing cranial access and repeated head-fixation was fixed to the skull with dental acrylic (Coltene/Whaledent, H00335).

On the imaging day, a craniotomy (0.5–3.5 mm diameter) was opened (-6.5 mm posterior to bregma; 0.5 mm lateral). Exposed tissue was irrigated with warm artificial cerebral spinal fluid (ACSF; 125 mM NaCl, 5 mM KCl, 10 mM D-Glucose, 10 mM HEPES, 2 mM $CaCl_2$, 2 mM $MgSO_4$; pH to 7.4 with NaOH). To dampen heartbeat- and breathing-induced brain motion, the craniotomy was filled with agarose (1.5%; Type III-A, high EEO; Sigma) in ACSF and covered by a coverslip fixed to the head plate, creating an optical window. The dura was left intact for Ca^{2+}-indicator injections and control experiments regarding mechanical stress, but generally removed just prior to agarose application to allow pharmacological interventions and electrical recordings. In some cases teflon-coated silver wires (A-M Systems, 785500 and 787000) were implanted for electro-corticogram (ECoG) recordings [262].

To note the coordinates studied, we re-anesthetized mice at the end of each experiment and marked recording sites with a water-repellent pen. We removed the brain and made digital images of the marked sites at several magnifications. We overlaid recording sites from different experiments to yield the outline of studied areas (Figure 1B).

Indicator Labeling

Targeted bulk-loading of cerebellar tissue with the Ca^{2+}-indicator Oregon green 488 1,2-bis(2-aminophenoxy)ethane-N,N,N',N'-tetraacetate-1 acetoxymethyl (OGB-1-AM; Molecular Probes) was done as described [235, 240, 263]. Concentrations of OGB-1-AM and DMSO in our pipette solution were 500 μM and 2-5%, respectively. To obtain a homogenous, widespread stain of the Purkinje cell and molecular layers we used several pressure ejections (3-3.5 psi; 1-2 min duration; Picospritzer III, General Valve, Fairfield, NJ). Typically five injections were performed in a grid pattern with ~150-250 μm spacing. To counterstain Bergmann glia, the red fluorescent astrocyte marker sulforhodamine 101 (SR101; 200-250 μM in ACSF; Invitrogen) was added to the injection solution or applied topically (1-3 min) in cases in which the dura was removed [240].

Imaging

After surgical implantation of the head-plate and a recovery of 1-7 days, mice were habituated in 3-10 sessions (usually 1 session/day) to accept increasingly longer periods (15-135 min sessions; 2.5-9.25 h total habituation time) of head restraint while on a 11.9-cm-diameter exercise ball (Habitrail, 62065) covered with copper mesh. The ball rotated on one axis, allowing the animal to stride forward or backward at liberty.

We used a custom upright two-photon microscope equipped with an ultrafast Ti:sapphire laser (Tsunami, Spectra-Physics), a nanofocusing system (P-725-PIFOC, Physik Instrumente), two detection channels, and a 20× water-immersion objective (Olympus, 0.95 NA, XLUMPlanFl). To excite OGB-1-AM, the laser was tuned to 800 nm. For simultaneous excitation of OGB-1-AM and SR101 we used 800-820 nm. To minimize aftereffects of anesthesia, mice recovered from isoflurane for >60 min prior to recordings. To minimize the possibility of photo-induced alterations in physiology, the duration and average power of continuous illumination typically were <7 min and <35 mW at a given brain location.

We used MP-Scope software for laser-scanning microscopy [264], which uses a 5 MS/s data acquisition card (PCI-6110, National Instruments). Images were typically 256 × 256 pixels (10.2–11.1 Hz frame rate) or 128 × 128 pixels (20.4 Hz). In 3D recordings, image stacks contained 5-20 optical sections acquired at 5, 10, or 20 μm axial spacing (1-4 Hz stack acquisition rate)

Behavioral Analysis

Mouse behavior was recorded in Mini-DV format using a digital Camcorder (Sony DCR-VX2000). Rotations of the exercise ball were monitored using an optical encoder (US Digital, E7PD-720-118).

Electrophysiology

For ECoG recording the stripped tip of a 140-μm-diameter Teflon-coated silver wire (A-M Systems, 785500) was placed on the pial surface in one corner of the craniotomy. A 330-μm-diameter reference electrode (A-M Systems, 787000) was placed over neocortex near bregma through a small hole in the parietal bone. ECoG signals were acquired using a differential AC amplifier (A-M Systems, 1700; gain, 1,000-10,000; Notch filtered).

Pharmacology

TTX (25-125 μM), γDGG (50 mM) and PPADS (1-2 mM) were dissolved in ACSF and applied in aqueous solution onto the optical window in the cranium. To facilitate drug entry into tissue the agarose and cover slip were cut on two sides to be flush with the craniotomy, allowing direct contact between reservoirs of solution and the agarose. Drug action was verified by observing changes in ECoG recordings and that Purkinje cells stained by OGB-1-AM showed a complete silencing of climbing fiber driven Ca^{2+}-spiking activity under TTX and γDGG. Also, TTX sometimes led to changes in mouse gait.

Data Analysis

Data analysis algorithms were written in MATLAB (Mathworks). All error values reported in the text are standard deviations, except for the errors in event percentages and rates, which are 68% confidence intervals. See Supplemental Methods for details.

SUPPLEMENTAL METHODS

Laser Doppler Studies

We tracked blood perfusion levels using a laser Doppler instrument (PF-5010-LDPM, Perimed) with a 1-mm diameter fiber-optic probe (PR 407-1) that we affixed either to the cranium or an optical window using a miniature probe holder (PH 07-6). A smoothing filter with a 0.03 s time-constant was applied to all perfusion traces. Perfusion measurements were performed either through the intact skull or using the same preparation as described in Experimental Procedures, Pharmacology. To assess the extent of motion artifacts, we performed control experiments by placing the fiber-optic probe in a colloidal suspension of latex particles (PF 1001, Perimed) that is normally used for probe calibration, within either the miniature probe holder or the head of the screw connecting the head plate and head holder.

Analysis of Laser Doppler Flowmetry Signals

We normalized laser Doppler signals by dividing by the median signal during non-movement periods, defined as when the speed of the exercise ball was <10 mm/s. We then detected transient increases in blood perfusion in a manner analogous to the detection of exercise ball rotation and Ca^{2+} flares (see below; the thresholds for blood perfusion event onset and offset were 15% and 10%, respectively).

Transgenic Animals

To confirm that the astrocyte marker SR101 [240] labels Bergmann glia in the cerebellum, we performed dual-color imaging studies of mice expressing EGFP under the control of the human S100β promoter [265] (5.5-6 months old; C57BL/6 genetic background) (Figure S1).

General Methods of Image Analysis

We first corrected any lateral motion artifacts in the imaging data using an image registration algorithm [266]. We corrected for photobleaching and, in dual color experiments for overlap in the emission spectra of the two dyes. Regions of interest (ROIs) containing glial fiber palisades were identified in images by applying a spatial high-pass filter and then finding contiguous regions above a threshold intensity value. To obtain functional Ca^{2+} signals from Bergmann glial fibers we examined the glial palisade ROIs within OGB-1-AM images and extracted traces of normalized fluorescence, $\Delta F(t)/F$, for each ROI (Figures 2 and 3).

Ca^{2+} sparkles were identified as all instances in which $\Delta F(t)/F$ traces from within an individual glial fiber ROI crossed a threshold value (10%). Ca^{2+} bursts are not confined to glial palisades and we identified these by visual screening of temporally low-pass filtered versions of the movie data. We estimated the temporal and lateral spatial profiles of Ca^{2+} bursts by fitting a convolution of two exponentials and 2D Gaussians to the time courses and maximum intensity projections, respectively (Figure 2).

Ca^{2+} flares were identified as instances in which the instantaneous rate of Ca^{2+} sparkles detected across the entire field of view crossed a threshold value (Figures 3 and 4). We used records of the exercise ball's rotational speed to perform reverse correlation analyses, in which we computed the mean rates of sparkling immediately before and after specified patterns of mouse locomotion. See below for more details.

Correction of Movement Artifacts

To correct lateral displacements of the entire field of view, we used the ImageJ plugin *TurboReg* [266] to register each movie frame with respect to a reference frame, generally the first frame of the movie, by minimizing the mean square difference between the pixel intensities. Movies were then cropped to the largest region that remained within the field of view throughout the recording duration. To estimate the severity of image displacements during periods of locomotion, we calculated the standard deviation of the displacements used for image registration. Axial displacements were estimated using movies recorded within a coronal optical plane.

Corrections for Overlap in the Emission Spectra of OGB-1-AM and SR101

Due to overlapping emission spectra, we observed modest bleed through of functional OGB-1-AM signals into the red spectral detection channel used to capture fluorescence signals from the non-functional, astrocyte marker SR101. To calibrate this overlap, we used large-amplitude bursts occurring during periods when the mouse was standing still to measure the fixed ratio of OGB-1-AM signals captured on the red and green detection channels. We calculated the relative changes in OGB-1-AM fluorescence detected in each color channel by normalizing each detector's signal by its mean over a brief epoch (<10 s) prior to each burst. We then determined the slope of the line that best fit the co-variation of the two normalized signals, which varied together in a highly linear manner ($r = 0.9 \pm 0.08$). By using data from two or more bursts we were able to fit the slope of the linear dependence between the two channels with <5% uncertainty. For example, the recording used for Figures 3B-D exhibited a bleed through ratio of 0.23 ± 0.008 (95% confidence interval). We multiplied OGB-1-AM signals by this ratio and subtracted the result from the red fluorescence signals to obtain the spectrally unmixed SR101 signals, as shown in Figure 3B.

Correction of Fluorescence Bleaching

We observed modest (<5%) declines in the mean fluorescence over the entire field of view during each recording lasting <10 min. We compensated for this decline by fitting the traces of normalized fluorescence with an exponentially decaying function. We then divided the fluorescence in each movie frame by the instantaneous value of the fitted decay function. For example, the decay time constants for Figures 3B-D were 53,000 s (OGB-1-AM) and 14,000 s (SR101), which accounted for declines in average fluorescence over the recording of 1.1% and 4.4%, respectively.

Computation of Normalized Changes in Fluorescence, ΔF(t)/F

We normalized traces of fluorescence intensity at each movie pixel by dividing by the mean fluorescence, *F,* over a reference period, generally the entire duration of the recording at that field of view. Data traces were smoothed by convolution with a Gaussian filter of unity gain and either 0.5 s (Figure 2B-D) or 0.7 s (Figures 3 and 4) standard deviation. The resulting traces were downsampled to ~2 Hz.

Identification of Glial Palisades

We first averaged the fluorescence intensity in the red detection channel, or in the green channel for single-color experiments, over all recorded time frames. In most experiments, glial palisades appeared with different relative intensity in both detection channels. In these cases we used the ratio of fluorescence in each channel to emphasize glial-specific peaks in fluorescence. We then sharpened the resulting image by convolving with the 3×3 digital high-pass filter kernel

$$K = \begin{bmatrix} -1 & -1 & -1 \\ -1 & k & -1 \\ -1 & -1 & -1 \end{bmatrix},$$

where k was either 8 or 12. We identified regions of interest (ROIs) corresponding to Bergmann glial processes by applying a threshold to the sharpened image and isolating contiguous regions within the resulting binary mask. We refined the definition of palisade ROIs by setting a maximum size criterion of ~20 μm^2. In dual-color experiments, we observed that large (>40 μm^2 area), brightly labeled structures in the OGB-1-AM channel generally corresponded to interneuron cell bodies not labeled by the astrocyte marker SR101. We therefore excluded these large ROIs from further analysis in single-color experiments, for which only OGB-1-AM signals were available.

Detection of Fluorescence Increases within Individual Glial ROIs.

We observed Ca^{2+}-transient related increases in fluorescence within individual glial ROIs during sparkles, bursts, and flares (Figure 3C). For each ROI, these transients were identified by finding peaks in the smoothed normalized fluorescence, $\Delta F(t)/F$, that exceeded a threshold value, generally a 10% increase. To obtain a conservative determination of the rate of fluorescence transients, we imposed a minimum interval, generally 10 s, between the individual fluorescence peaks detected within each ROI to prevent redundant detection of multiple peaks within individual events (Figure 2B). To avoid including noisy signals resulting in high false positive detection rates, we excluded ROIs that showed >5 times the median rate of Ca^{2+}-transients for each movie.

Calculation of Event Rates and Confidence Intervals.

Confidence intervals for all event rates that we report were estimated by making the statistical assumption that the number of event occurrences is Poisson distributed.

Statistical Tests of Significant Difference

We used Wilcoxon rank sum tests to assess the significance of differences between median Ca^{2+} transient characteristics in each of the five experimental conditions: anesthetized, awake, TTX-, γDGG-, and PPADS-treated awake. For analyzing the anisotropic spread of Ca^{2+} during bursts (see below), we hypothesized that the spread in the direction parallel to the cerebellar folium is greater than the spread perpendicular to the folium. To test this, we used a Wilcoxon signed rank test of the hypothesis that the difference between the two radii has a median of zero. Because we measured a large number of individual events (n = 95 across all conditions), we used a normal approximation for the distribution of the test statistic to compute the p-value.

To test the significance of differences in the frequencies of sparkles and bursts across the different pharmacological conditions used within individual experiments, we used a likelihood ratio test based on the assumption that the numbers of events were Poisson distributed. In each condition, i, we observed n_i events in time t_i over a field of view area a_i. The maximum likelihood estimate of the event rate was thus $r_i = n_i/(t_i\, a_i)$. To test whether the rates under the two different conditions were significantly different, we used a likelihood ratio test. This compared the likelihoods of obtaining the total $n_i + n_j$ events given two distinct rates for the two experiments versus with one estimated rate, $(n_i + n_j)/(t_ia_i + t_ja_j)$, for both conditions. The resulting likelihood ratio,

$$LR = \frac{r_i^{n_i} r_j^{n_j}}{\left[(n_i + n_j)/(t_i a_i + t_j a_j)\right]^{(n_i + n_j)}},$$

has an associated p-value, which is found by calculating $\chi^2 = -2\log(LR)$ and comparing with the cumulative chi-square distribution with 1 degree of freedom.

Analysis of Ca^{2+} Sparkles

Sparkles were detected within individual glial ROIs as described above. The duration of each event was defined as the interval from when the signal first exceeded the detection threshold (10% $\Delta F(t)/F$) to when the signal returned below this level.

Analysis of Ca^{2+} Bursts

Analysis of Two-Dimensional Spatial Profiles

We identified the approximate location and onset time of each burst by visual inspection of the normalized fluorescence data. Only events that did not extend beyond the recording field of view were used for further analysis of spatial profiles. For each event we computed the maximum of the normalized fluorescence in each pixel during the first 7 s after event onset. We rotated the resulting image to align it with a local coordinate system in which the long axes of Purkinje cell dendrites were oriented along the y-axis. We then used nonlinear least squares regression (Levenberg-Marquardt algorithm; MATLAB routine *nlinfit*) to fit the shape of the maximum intensity profile with a two-dimensional Gaussian function. The fit was constrained to have its principal axes oriented along the x and y directions. This fitting procedure produced reasonable estimates of burst spatial extent, as judged by inspection of the fit and its residual, even in cases in which animal movement or blood flow influenced the maximum intensity projection. However, for the bursts of greatest magnitude we noticed that the maximum intensity projection deviated from the fitted Gaussian, exhibiting a flatter central region and more sharply defined edges. The diameters parallel and perpendicular to the folium plotted in Figure 2H were defined as four times the standard deviation values from the Gaussian fit.

Analysis of Three-Dimensional Spatial Profiles

To characterize the volumetric structure of bursts, we analyzed as described above each of the lateral planes sampled within the three-dimensional (3D) recording. We smoothed individual image frames by convolution with a Gaussian that had a standard deviation of 1 pixel, approximately 2 μm across. Because the pixel spacing in the axial dimension was about 5 μm, or roughly twice as large as the lateral pixel spacing, we linearly interpolated the data at planes midway between each scan plane. We also smoothed the temporal dynamics, as described above.

The normalized fluorescence data, represented as a four-dimensional array of intensity values at every spatial point (*x,y,z*) and time frame (*t*), were examined in three ways. We inspected the complete set of information available about each event from the collection of movie frames acquired across a series of optical planes (Figure 2E). We viewed the maximal spatial extent of each event using projections of maximum intensity over time and one spatial dimension into the *x-y*, *x-z*, and *y-z* planes.

Finally, to compare multiple events recorded within the same imaging volume, we used a sequence of contours, each of which represented the spatial extent of the burst at a distinct depth beneath the brain surface (Figure 2G). We determined the contours using the maximum intensity projection over time for each event. We smoothed each *z*-slice of these projections by convolving with a Gaussian (5 µm s.d.), and applied a threshold to the result to produce a binary mask of the region for which $(\Delta F(t)/F)_{max} \geq 15\%$. Isolated pixels or groups of pixels that were not contiguous with the main volume of the event were excluded. We used the remaining connected volume to create the contours at each imaging plane (Figure 2G).

The grayscale background images in Figure 2G show the patterns of OGB-1-AM labeling at the boundaries of the imaging volume. The areas covered by the contours of each burst, as projected onto the *x-y* plane, are shown at the bottom of Figure 2G. The cross-sectional area at each scan plane was used to assess the depth profile of the events.

To quantify the spatial extent and total volume of bursts in 3 spatial dimensions, we computed the maximum normalized fluorescence intensity at each pixel during the first 7 s of each event recorded during 3D imaging experiments. We excluded events that occurred near any of the lateral edges of the imaging volume. Because the frame rate for volumetric imaging (~1-2.5 Hz) was lower than the frame rate for 2D-imaging, we chose to smooth the event shape in each imaging plane using a Gaussian spatial filter (5 µm s.d.), rather than using a temporal filter as in 2D analysis. The filter width was chosen to match the minimum spacing between imaging planes in the axial dimension. We fit the event shape with a three-dimensional Gaussian ellipsoid with eight free parameters (3 coordinates for the center of mass, 3 independent spatial widths, background intensity, and event amplitude). To correct for the systematic bias introduced into the fit parameters by Gaussian convolution, we subtracted the variance of the smoothing filter from the fit variances. The volume of each event was defined as the volume of the ellipsoid contained within two standard deviations:

$$V = \frac{4}{3}\pi(2\sigma_x)(2\sigma_y)(2\sigma_z).$$

Analysis of Temporal Dynamics

To characterize the temporal dynamics of bursts, we used the average fluorescence over the 7 s preceding the event onset to estimate background fluorescence. We then averaged fluorescence changes across an elliptically shaped ROI defined as the boundary at which the Gaussian spatial fit to the event (see above) falls to 1/*e* of its peak value. We found that the normalized fluorescence traces were well fit (r = 0.9 ± 0.1 s.d.) by a convolution of two exponentials, which has four free parameters, amplitude (A), time constants for the rising (τ_r) and decaying (τ_d) phases, and onset time (t_0):

$$f(t) = \begin{cases} 0 & t < t_0 \\ A\dfrac{e^{-(t-t_0)/\tau_d} - e^{-(t-t_0)/\tau_r}}{\left(\tau_r/\tau_d\right)^{\tau_r/(\tau_d-\tau_r)} - \left(\tau_r/\tau_d\right)^{\tau_d/(\tau_d-\tau_r)}} & t \ge t_0 \end{cases}.$$

We employed MATLAB's Trust-Region nonlinear least squares solver with constraints to perform the fits, which were then used to align and normalize the individual traces and to compute the average of the normalized traces in each experimental condition (Figure 2F). In many cases (81% of bursts observed in the control condition), the best fit time constants were equal, $\tau_r = \tau_d$, to within 1%. In this case the fit is equivalent to a first-order alpha function with a single time constant:

$$f(t) = \begin{cases} 0 & t < t_0 \\ Ae^{1-(t-t_0)/\tau}(t-t_0)/\tau & t \ge t_0 \end{cases}.$$

Analysis of Somatic Bergmann Glial Ca^{2+} Transients

Ca^{2+}-transient related increases in fluorescence within individual somatic ROIs were identified by finding peaks in the smoothed normalized fluorescence, $\Delta F(t)/F$, that exceeded a threshold value, generally a 10% increase.

Analysis of Ca^{2+} Flares

Analysis of Locomotor Speed

We smoothed and downsampled to ~20 Hz the data from the optical encoder that reports the rotational velocity of the exercise ball. We further smoothed the signal in time (Gaussian filter, s.d. of 0.7 s). The onset of locomotion was defined as the time at which the smoothed running speed increased over a threshold of 20 mm/s. Offset was defined as the time at which the speed subsequently fell below 10 mm/s. These thresholds were chosen to avoid the detection of small movements and to ignore small dips in speed during otherwise continuous bouts of running.

Detection of Ca^{2+} Flares

Within each ROI we first detected all events in which a local peak in the smoothed fluorescence signal was achieved and exceeded a fixed threshold (see above; threshold was 10% $\Delta F(t)/F$ and minimum inter-event interval was 10 s). We calculated the mean instantaneous rate of glial Ca^{2+} events within the field of view by counting all events occurring at each movie time frame, and multiplying by the frame rate and dividing by the field of view area. The rate of glial Ca^{2+} events was then smoothed using a Gaussian filter (see above; Gaussian s.d. was 2 s). The onset and offset times of Ca^{2+} flares were defined analogously to the detection of movements by the times at which the smoothed rate crossed a threshold. Because of differences in the signal/noise ratio across experiments, we adjusted these thresholds manually for each recording within the range from 10-220 Hz/mm^2. For this analysis we excluded epochs of ~10 s duration around the time of each burst, resulting in minimal loss of data since bursts occurred very sparsely.

Movement-Triggered Averaging of Glial Ca^{2+} Signals

We examined how Ca^{2+} flares depend on two aspects of locomotor behavior. We first studied the impact of the duration of running bouts on the physiological event dynamics. We assigned movements to three categories containing either all movements, brief movements (<5 s duration from onset to offset), or extended movements (>15 s duration). To prevent interference from recently completed bouts of running (see below), we excluded movements that began within a short time (10 s) after the offset of another movement or, for the first movement in each recording, after the start of the recording. For every movement in each of the three duration categories, we determined whether the corresponding signal (Ca^{2+} event rate in Bergmann glial fibers) during a period from 10 s before to 25 s after movement onset exceeded a threshold (20 Hz/mm^2). We aligned all such signals to the start of the movement and computed the average (Figure 4C).

Next we examined the impact of the pause duration, defined as the time from the offset of the previous bout of movement to the onset of the current movement, on the probability of observing a Ca^{2+} flare (Figure 4D). We classified all movements based on their inter-movement pause into logarithmically spaced bins. For this analysis we excluded the first running bout in each recording, for which we could not determine the pause time. For each movement we determined whether there was a supra-threshold increase in the corresponding glial Ca^{2+} event rate signal within 10 s of movement onset. Within each bin, the probability of a physiological response at movement onset was defined as the number of movements accompanied by a supra-threshold increase in the physiological signal divided by the total number of movements. We estimated 68% confidence intervals for our determinations of response probability by assuming the responses were governed by a binomial distribution. We performed a logistic regression to fit the binary flare occurrence data to a two-parameter sigmoid:

$$P = \frac{1}{1 + (IMI / IMI_0)^x},$$

where IMI is the inter-movement interval and IMI$_0$ is a fit parameter corresponding to the pause value for which the probability of flare occurrence was 50%. The fit parameters for the line shown in Figure 4D were $IMI_0 = 27.5$ s and $x = -1.42$.

Chapter 5: Automated analysis of cellular signals from large-scale calcium-imaging data

ABSTRACT

Recent advances in fluorescence imaging permit studies of Ca^{2+}-dynamics within large numbers of cells, in anesthetized and awake behaving animals. However, unlike for electrophysiological studies standardized algorithms for assigning optically recorded signals to individual cells have not yet emerged. Here we describe an automated sorting procedure that combines independent component analysis and image segmentation for extracting cells' locations and their dynamics with minimal human supervision. In validation studies using simulated data, automated sorting significantly improved estimation of cellular signals compared to conventional analysis based on image regions of interest. We used automated procedures to analyze data recorded by two-photon Ca^{2+}-imaging in the cerebellar vermis of awake behaving mice. Our analysis yielded simultaneous Ca^{2+}-activity traces for up to >100 Purkinje cells and Bergmann glia from single recordings. Using this approach, we found microzones of Purkinje cells that were stable across behavioral states and in which synchronous Ca^{2+}-spiking rose significantly during locomotion.

INTRODUCTION

Techniques for delivering fluorescent Ca^{2+}-indicators into large numbers of cells have enabled recent *in vivo* imaging studies of the dynamics of hundreds of neurons and astrocytes [18, 32, 48, 225, 226, 267-270]. Fast, three-dimensional laser-scanning microscopy has further increased opportunities for observing Ca^{2+} activity across many cells [219, 271, 272]. However, by comparison to these advances in staining and imaging, state-of-the-art computational techniques for analysis of Ca^{2+}-imaging data lag behind and are mainly region of interest (ROI) analyses. These ROI analyses have typically been manual [28, 267, 272, 273], or semi-automated [40] means of locating cells within movie data and cannot be easily scaled to handle the largest available data sets without undue human labor. Moreover, ROI analyses have largely been based on heuristic, pre-conceived notions of what defines the geometric and other characteristics of specific cell types [40, 269, 272], rather than general statistical principles for decomposing an imaging data set into its constituent signal sources. This mismatch between the current ability to perform imaging experiments and the capability to analyze the resulting data sets significantly restricts the capacity to attain biological insights.

In several respects, the present state of affairs for Ca^{2+}-imaging resembles that of *in vivo* electrophysiology in the early 1990s, when multi-electrode techniques were blossoming but standardized algorithms for spike sorting had yet to appear. Today, automated spike sorting is widely applied in multi-electrode studies to assign spikes to individual cells [38, 222, 274], and there are several commercial software packages for the task. Although improvements to spike sorting routines justifiably continue, the accurate and efficient assignment of spikes to single cell units has enabled key advances in our understanding of neural coding in sensory, motor, and cognitive pathways [37, 275-278]. By analogy, an automated procedure capable of extracting intracellular Ca^{2+}-signals from optical recordings of hundreds of cells would become a crucial enabler for scientific advances achieved by brain imaging. However, the computational challenges in devising such a procedure are distinct from those posed by multi-electrode recordings.

Spike sorting for multi-electrode recordings tends to rely on two basic ideas. First, the temporal waveforms for spikes generated by different individual cells are often sufficiently dissimilar so as to provide the basis for spike classification. Second, the activity of individual cells is often recorded on multiple electrodes, which generally allows a spatial location to be assigned to each somatic action potential based on the electrode positions and the relative amplitude of the spike on each recording channel.

Unfortunately, neither of these two means for sorting signals works well for imaging data. First, the temporal waveforms for Ca^{2+}-activity are strongly dictated by the dye's kinetics of Ca^{2+} binding and unbinding [279], and so do not provide strong signatures of individual cells' identities. Second, image pixels can contain a combination of signals from neuropil, multiple neurons and astrocytes, and various sources of noise and movement artifacts. It is thus nontrivial to disentangle the signals from the individual sources and to find the shapes and locations of each cell. What is needed is a guiding principle for analysis of imaging data that can be used to identify cells' locations in the field of view and to extract reconstructions of activity time courses.

We formulated such a principle by considering the regulation of intracellular [Ca^{2+}], which can transiently increase ~100-fold above background levels during cellular events such as action potentials. This dynamic implies that brief periods of elevated [Ca^{2+}] are typically sparsely interspersed in time among many more background-dominated time frames. Sparseness also holds in the spatial domain when only a small subset of pixels collects signals from each individual cell. Thus, temporal and spatial sparseness should be generic features of the Ca^{2+} signals in large-scale imaging data. Sparseness also typically distinguishes these signals from movement artifacts and other noise sources. Further, sparseness can be quantified by simple statistical measures, such as the skewness of amplitude distributions.

This reasoning led us to an algorithm that estimates sparsely distributed cellular locations and time courses by parsing our data into a combination of signals, each of which independently exhibits a high degree of skewness. To do this, our algorithm relies on an independent component analysis (ICA) [41, 42, 280], which forms the core of our 4-stage data processing protocol (Figure 1). Unlike conventional approaches using manual or semi-automated ROI analysis, ICA requires little user input or supervision and does not rely on preconceptions about the geometries or activity patterns of the cells to be found. Thus, our automated approach can be readily applied when there are many hundreds of cells, including those with unknown dynamical or morphological signatures.

We validated our protocol using simulated movies designed to mimic Ca^{2+}-imaging data. These simulations permitted detailed studies of how well Ca^{2+} signals can be extracted by cell sorting. The signals and noise in the artificial data were similar to that of real data we acquired in cerebellar cortex. The simulations revealed a graceful degradation in the performance of the automated cell sorting procedures as the signal quality of the data decreased; although fewer cells were recovered from noisy data, those cells that were identified yielded activity traces of high fidelity to the actual Ca^{2+}-dynamics. By comparison, a semi-automated ROI analysis achieved comparable performance only for movies with low noise levels and more restricted fields of view.

We further tested our analysis on data recorded by two-photon microscopy in the cerebellar vermis of awake, behaving mice. The signals included dendritic Ca^{2+} spikes and astrocytic Ca^{2+} transients from large numbers of Purkinje cells and Bergmann glia, respectively. Despite that the neuronal signals evolved one to two orders of magnitude faster than the glial signals, automated cell sorting successfully extracted Ca^{2+} signals from populations of up to >100 total Purkinje cells and Bergmann glia. The observed dendritic Ca^{2+} spikes in the Purkinje cells represent the Ca^{2+} component of these cells' complex (Na^{+} and Ca^{2+}) spikes, which we verified through simultaneous optical and electrophysiological recordings [23, 40].

To illustrate the utility of our algorithm for obtaining physiological insight, we applied our automated procedures to study the spatiotemporal organization of Purkinje cells' Ca^{2+}-spiking activity *in vivo*. We found that in cerebellar lobules V and VI synchronously active cells cluster into local neighborhoods extending ~7-18 cells across in the medio-lateral dimension. We identify these clusters as the cerebellar microzones, small patches of Purkinje cells receiving similar climbing fiber input [44]. The width of the patches has been estimated as ~200 µm [44], but these inferences from electrophysiological studies were probably overestimates and did not permit conclusions about whether adjacent microzones are precisely demarcated from one another or merge smoothly. We found repeatedly that the cerebellar microzones of awake animals have sharply delineated medio-lateral boundaries and encompass cells having significantly greater synchronous activity with each other than with cells in other microzones. Our results complement prior studies of Purkinje cells in awake animals using electrode arrays, which showed complex spiking synchrony exists in bands extending millimeter-sized distances in the rostral-caudal dimension [46, 89]. However, the 250-µm-spacing between electrodes used in those studies, which permitted the finding on the rostral-caudal extent of synchrony, also precluded fine resolution of microzone structure in the medio-lateral direction.

In distinction to prior fluorescence imaging studies of cerebellar organization in anesthetized animals [40, 281], our studies in awake animals provided the opportunity to examine the modulation of microzonal activity by motor behavior. We sought to address the longstanding debate on whether microzones have stable anatomical boundaries [44], or are dynamically changing ensembles whose cellular constituents vary across behavioral states [46, 89]. We observed that during animal locomotion on an exercise ball the basic spatial organization of correlated Purkinje cell Ca^{2+} activity was unchanged from that seen in awake but resting animals. Thus, our data provide no evidence for dynamical flexibility in the cellular composition of microzones but rather support the hypothesis that microzones are stationary anatomical entities. However, during locomotor behavior the levels of synchronous Ca^{2+} activation rose within individual microzones but decreased between adjacent microzones. Thus, microzones not only retained stable boundaries across different behavioral states but also increased in prominence during motor behavior. These findings reveal basic features of cerebellar dynamics and highlight the impact of automated procedures for analyzing imaging data.

RESULTS

Basic Principles for Extracting Cellular Signals from Imaging Data

The purpose of cell sorting is to identify a set of spatial filters and corresponding time traces that respectively represent the spatial locations and temporal dynamics of individual cells (Figure 1A). We devised a procedure to do this in four stages (Figure 1B), the first two of which respectively involve principal and independent component analyses and comprise the theoretical focus of this article (Figures 1-4).

The signals to be found are buried within high-dimensional imaging data sets that might contain $N_x \sim 10^5$ pixels and $N_t \sim 10^4$ time frames. The first stage of analysis should reduce this dimensionality, since otherwise specifying individual cells' spatial filters and time courses would require a large set of parameters for each cell. In principle these parameters could be found, but tuning many parameters can hamper signal extraction techniques [282]. Ideally, a dimensional reduction of the data should discard dimensions that contain mainly noise and not signals. We did this by using an established approach, principal component analysis (PCA), to find and then discard dimensions that mainly encoded noise [283].

PCA identifies a linear transformation of the data yielding a set of basis vectors, the principal components, that are rank ordered by the variance of the signals along each basis vector. Lower ranked components have the highest signal variances. Thus, truncation of higher ranked components with smaller variances often does not interfere with an approximate reconstruction of the data. On the contrary, if ongoing background noise has a smaller variance than transient but strong signals, then removing higher order components from the data representation can eliminate substantial noise. To do this, one calculates the principal components' variances for both the real data and a reference movie composed of uncorrelated, Gaussian noise at each pixel and time frame [284]. The two variance spectra are compared to find the data principal components whose variance is below the noise floor (Experimental Procedures). In our cerebellar Ca^{2+}-imaging studies, after this truncation of noise components we achieved data of a reduced dimensionality, K, much less than either N_x or N_t. Nonetheless, although PCA is an efficient means for separating signals from background noise, PCA alone cannot isolate Ca^{2+} signals from individual cells.

PCA by itself is ill suited for cell sorting since it relies on differences in signal variance to identify data components; in practice temporal variations in different cells' fluorescence signals tend to be of similar amplitude. Thus, each principal component generally has a mix of signals from multiple cells. Instead, one seeks a set of signal sources reflecting individual cells. The intuitive expectation is that such sources should be spatially localized, with a sparse spatial distribution of pixel weights for each cell. Thus, the distributions of pixel weights for each cell's spatial filter should be highly skewed, with most pixels having almost no weight and a few having strong weight at the cell's location. A basic measure of sparseness for such asymmetric distributions is the skewness, defined as the third moment normalized by the cube of the standard deviation. Skewness thus provides a useful means of assessing spatial sparseness and searching for signatures of individual cells' dynamics.

Similarly, one also expects many cells will have sparse signals in the temporal domain, since brief rises or bursts in fluorescence will stand out against ongoing background noise. If the times between the brief rises in signal are greater than the signals' durations, then the distribution of signals in the time traces will be sparse and skewed. (Please see Discussion concerning when this assumption is invalid). These expectations regarding the spatiotemporal distibutions of Ca^{2+} signals motivated our choice of using independent component analysis (ICA) to extract cellular signals from the K dimensions remaining after PCA.

ICA is well suited to this task, since it is designed to seek pairs of spatial filters and time traces that are sparse and statistically independent of one another [41, 42, 280, 285] (Experimental Procedures). Signals from different cells generally show individual dynamics that are sufficiently independent to make ICA a good tool for signal separation. In practice, we found that cells' activities can actually have significant correlations and ICA will still work well. This is so even for cell pairs with moderately high correlation coefficients.

We used a version of the *FastICA* algorithm, which maximizes a statistical objective function and can combine spatial and temporal information [286]. This objective function is a measure of the sparseness of the extracted signals, and we used a combination of spatial and temporal skewness for this purpose. Although ICA is conceptually more complex than PCA, the ICA stage of our analysis is the more efficient of the two computationally. PCA requires all the raw data, and with computation of K principal components the number of operations rises as $KN_x N_t^2$ (assuming $N_x > N_t$). For the ICA stage this cost rises only as $K^2 (N_x + N_t)$. For example, in the analysis below of our cerebellar recordings for which $K \sim 10^2$, $N_x \sim 10^5$, and $N_t \sim 10^4$, PCA took ~10-30 min of processing time, but ICA generally took < 1 min. A main issue in this analysis was whether ICA's assumption that Ca^{2+} signals are spatiotemporally sparse and statistically independent could be validated in the live brain.

ICA Yields Physiological Ca^{2+} Signals That Are Statistically Sparse and Skew

We examined the sparseness of Ca^{2+} signals recorded by two-photon microscopy in the cerebellar vermis of awake, head-restrained mice (Figures 2A-2E). We used multi-cell bolus loading of the Ca^{2+}-indicator Oregon Green 488 BAPTA-1-AM to label neurons and astrocytes in the cerebellar molecular layer [40, 225, 226, 287] (Supplemental Experimental Procedures). The mice stood on an exercise ball and could walk or run at liberty during imaging [28, 287]. This approach permitted optical recording of the dendritic Ca^{2+} spikes that are associated with complex (Na^+ and Ca^{2+}) action potentials in Purkinje neurons [23, 40]. We also observed Ca^{2+} activation in Bergmann glial fibers [287]. Application of PCA to the Ca^{2+}-imaging data yielded principal components that were generally not spatially localized and failed to isolate individual cells' activities (Figures 2A and 2B, *top row*). Distributions of signal values along the principal component vectors were highly symmetric and approximately Gaussian (Figure 2C, *right*) and did not separate Ca^{2+} transients from background noise. However, ICA transformed these principal components into new basis vectors with more distinctive properties.

A majority of independent components had spatial filters far more localized than those of the principal components and could reasonably represent individual cells (Figures 2A and 2B, *bottom row*). As hypothesized, spatial filters found by ICA had sparse, highly skewed distributions, with the weights of small subsets of pixels being up to 19 standard deviations greater than background pixels (Figure 2C, *left*). For example, many of the localized ICA spatial filters from our cerebellar recordings resembled a stripe in the rostral-caudal direction. This is the expected optical cross-section through the dendritic tree of a single Purkinje cell

(Figure 2A, *bottom*), since these dendritic trees are nearly planar and lie perpendicular to the field of view [23, 40]. The stripes' widths (7.0 ± 2.3 µm, mean ± s.d.; n=199 dendrites in 5 mice) and lengths (90 ± 40 µm) as estimated by twice the standard deviations of Gaussian fits were consistent with anatomically derived values [288, 289] and those reported in prior *in vivo* Ca^{2+}-imaging studies of Purkinje cells [23, 40, 226]. These elongated filters found by ICA had their long axes perpendicular to the folium of the cerebellar cortex and their short axes aligned with the folium.

In the temporal domain, Ca^{2+} activity traces of the independent components also had sparse statistics. Unlike with principal components, signal distributions were asymmetric and highly non-Gaussian (Figure 2C, *right*). The asymmetry arose because there was a large fraction of time frames during which individual cells were inactive. Transient rises in $[Ca^{2+}]$ interrupted Gaussian distributed noise fluctuations and resulted in a sparsely distributed set of time frames during which fluorescence signals increased by up to eight standard deviations (Figure 2C, *right*). For the striped independent components representing Purkinje cell dendritric trees, the time traces yielded similar Ca^{2+} spiking rates and spike waveforms as those from prior manual and semi-automated ROI analyses of cerebellar Ca^{2+}-imaging data [23, 40, 226].

After verifying the sparseness of signals identified by ICA, we also examined the independence of the signals' dynamics by examining joint distributions of signals from pairs of independent components, each representing one Purkinje cell. The example shown in Figure 2D (*bottom*) is typical of 276 such pairs we studied. During most time frames, signals in both cells clustered close to the mean background levels. Each cell was active during a subset of time frames, and the joint distribution showed these subsets for the two cells were distinct, unlike with pairs of principal components (Figure 2D, *top)*. Only a small number of time frames had high fluorescence levels in both cells simultaneously, consistent with the cells' activities being statistically independent. We did encounter some cell pairs with correlated activity, which we handled in the third stage of our analysis procedure (see below).

Interestingly, we found that although independent components are not inherently rank ordered, we could use the skewness as an effective diagnostic for distinguishing independent components representing individual cells from those representing artifacts, such as due to movement artifacts. Components we recognized by eye to be neuronal or glial signals generally had higher spatial and temporal skewness than noise or artifactual components

(Figure 2E). By sorting the independent components by their temporal skewness we could segregate and more easily reject artifactual components. The association of skewness with effective isolation of cellular signals was consistent with the increases in skewness that occurred at the ICA stage of our procedure. For example, the skewness of pixel value distributions was -0.009 ± 0.26 (mean ± s.d.; n=504 components in 5 mice) for principal components but 4.4 ± 2.7 (n=300) for independent components. When we inspected the independent components with high skewness we found several types of cellular signals in our cerebellar recordings.

Prominent signals extracted by ICA were those from one, sometimes two, and occasionally three or more of the striped regions that we interpreted to be cross-sectional profiles of the corresponding number of Purkinje cell dendritic trees (Figure 2A and Figure 5). Signals from these stripes exhibited brief increases in fluorescence (<50 ms rise), followed by a decay to baseline over 124 ± 63 ms (mean ± s.d., n=150 cells in 5 mice). In awake mice these Ca^{2+} transients occurred at a spontaneous rate of 0.76 ± 0.15 Hz (n=199 cells in 5 mice) when the mice were not moving on the exercise ball. These rates are consistent with those from electrophysiological recordings of complex spiking [89, 290] and Ca^{2+}-imaging studies in Purkinje cells [23, 40, 219]. We verified that Ca^{2+} spikes identified by ICA corresponded to complex spikes, through simultaneous optical and extracellular single-unit electrophysiological recordings (Figures 5C and 8). ICA also regularly yielded signals with slower dynamics consistent with the characteristics of Ca^{2+} activation in fibers of Bergmann glia [287], which are the sole type of astrocyte in the cerebellar molecular layer (Figure 5). Independent components with skewness values lower than those for Bergmann glia or Purkinje cells generally resembled artifacts and were excluded from all further analyses (Figure 2E).

Combining Spatial and Temporal Information Improves Cell Sorting

After obtaining these initial results, we sought to fine-tune the ability of ICA to isolate individual cells and studied in simulations how accurately the extracted signals reflected underlying physiological dynamics. We examined how this depended on whether ICA was used to optimize the skewness of the spatial filters, time courses, or some combination of the two (Figure 3A). Our initial studies (Figure 2) had followed previous ICA approaches in maximizing the sparseness of the spatial filters [280]. However, for cell sorting it is natural to include temporal statistics as well.

We added temporal statistics into our analysis by using a combination, spatio-temporal form of ICA [43, 291]. The objective function optimized by spatio-temporal ICA is a linear combination of spatial and temporal skewness values averaged over all components (Figure 3A). A single parameter, μ, set the relative contribution of spatial and temporal information (Figure 3) (Experimental Procedures). We sought μ values that improve the accuracy with which cellular signals can be reconstructed from noisy data. To explore, we created artificial movies containing known signals and noise statistics (Figures 3B and 3C). These movies mimicked our actual recordings from the mouse cerebellar molecular layer. We designed the cells' sizes, shapes, spatial distributions, and temporal dynamics, as well as the noise sources and background structures such as blood vessels, to resemble those in the experiments. We defined the fidelity of extracted signals, F, as the correlation coefficient between each cell's actual dynamics and its reconstructed signal (Experimental Procedures). F=100% implied perfect reconstruction; F=0 implied the algorithm failed to identify any signature of the cell's activity.

A purely spatial ICA ($\mu = 0$) yielded signals partly resembling the underlying dynamics ($F = 77 \pm 20\%$, mean ± s.d.; n=1100 artificial cells) (Figures 3D-3G). A purely temporal ICA ($\mu = 1$) was unsuccessful and led to low fidelities ($F = 36 \pm 13\%$) and spurious spatial filters (Figures 3D-3G). However, spatio-temporal ICA improved the fidelities of our reconstructions (86 ± 16%) ($\mu = 0.58$; Figures 3D and 3E, *middle row*). μ values of ~0.5-0.6 led to the greatest proportion (~88%) of cells reconstructed with high fidelity, which we defined as $F > 75\%$ for purposes of comparisons to conventional cell sorting (Figures 3F and 3G). In subsequent studies of real data, we verified that spatio-temporal ICA extracted the greatest number of signal traces whose dynamics resembled that expected of Purkinje cells. Thus, we next examined how spatio-temporal ICA compared to conventional ROI analysis.

Comparison of Independent Component and Region of Interest Analyses

To permit comparisons between ICA and ROI analyses we used a form of ROI analysis that represents the best a human analyst could do if she correctly identified a set of example movie frames during which each cell is active (Experimental Procedures). The most challenging situations for cell sorting that arose in our simulations and experiments involved cells with spatially overlapping pixel sets (Figures 4A and 5A). In such cases ROI analysis failed to remove cross talk between overlapping cells (Figure 4B). However, spatio-temporal ICA avoided cross talk, because the independent components had negatively weighted surround regions in the spatial filters that subtracted signals from overlapping cells (Figure 4C). Cross talk removal was particularly helpful for separating spatially overlapping neuronal and glial signals.

We studied how noise levels and the field of view area affected ICA and ROI analyses (Figure 4D). Our simulations included photon shot noise, which dominates instrumentation noise in a properly de-noised two-photon microscope. For each pixel the photon counts obeyed an approximation of Poisson counting statistics, with the standard deviation set by a signal to noise ratio (SNR) parameter (Supplemental Experimental Procedures). In simulations with 0.09 mm^2 field of view and ~100 cells, ICA used increases in SNR to improve reconstruction fidelities and reduce cross talk between neighboring cells (Figure 4E). With this cell density the performance of ICA approached an optimum with SNR > 20 (~400 photons per pixel per time frame). In this regime the median fidelity approached 95% and >80% of extracted signals had $F > 75\%$. By comparison, ROI analysis did not improve much with increases in SNR and even at high SNR failed to separate overlapping cells (Figure 4E).

Given a fixed number of pixels the difficulty of cell sorting increases with larger fields of view, because the number of overlapping cells increases as each pixel covers a larger area. Across a broad range of field sizes, ICA achieved high signal fidelity and low cross talk, whereas ROI analysis steadily degraded as increased overlap between cells led to higher levels of cross talk (Figure 4F). Overall, ICA yielded higher fidelity traces than ROI analysis and over much larger fields of view (Figure 4D). This observation is useful for experimental design, since it potentially cuts the number of photons per pixel that must be collected for high quality reconstructions (Discussion).

Image Segmentation for Separating Correlated Signals

A key advantage of imaging is its ability to sample dense networks of neurons and astroctyes. In such studies the dynamical correlations between neighboring cells are often of chief interest. Thus, a key question is whether ICA, which relies on statistical independence to identify cells, might be limited in its ability to separate cells with highly correlated signals?

In studies of artificial data with varying degrees of synchrony among Purkinje cell pairs, we found ICA could readily distinguish cells whose spike trains had correlation coefficients, r, as high as 0.8. Because the spike rates used in our simulations (≤ 1 Hz) were well below the imaging frame rate (10 Hz), this level of correlation corresponds to ~80% of each cell's spikes occurring synchronously with those of another cell. Although the correlated cells were far from independent, sorting performed well and achieved high fidelity signal estimates comparable to the performance of cell sorting in the absence of correlation (r=0). This was because the activity traces of the individual cells were still more skew than traces from cell mixtures. This result is consistent with prior reports that ICA is often robust to deviations from the underlying model assumptions [292]. Nonetheless, when $r > 0.8$, ICA often extracted signals from strongly correlated cells within a single independent component. This was because if two cells had highly correlated activity, then a time trace that averaged pixels from both cells was generally sparser, with a more skewed distribution, than the individual cells' signals.

Our sorting procedure therefore augments ICA with an image segmentation step to disentangle signals from highly correlated cells that are spatially separated (Figure 1B). Following ICA, we smoothed each component's spatial filter and applied a binary threshold to find local regions with strong signal contributions (Figure A,B). If a spatial filter contained more than one such group, we created new filters, each of which contained only one of the image segments (Figure 9C). This successfully separated distinct cells with strongly correlated activity (Figure 9D). However, if one of the image segments included only a small number of pixels, segmentation reduced the signal amplitude relative to background fluctuations since the noise was averaged over a smaller set of pixels. To avoid excessively noisy signals, we excluded any image segments covering an area smaller than 50 pixels, which was much smaller than the number of pixels occupied by a full Purkinje cell dendritic tree (478 ± 165 pixels; mean ± s.d.; n=312 dendritic trees). After these extensive tests on simulated data, we applied our procedures to the study of large-scale two-photon Ca^{2+}-imaging data acquired in the cerebellar cortex of live mice.

Sorting Cerebellar Ca^{2+} Signals From Live Mice

To illustrate 4-stage cell sorting in an experimental context, we analyzed data taken in the cerebellar molecular layer of both anesthetized and awake, head-fixed mice (Experimental Procedures). Using spatio-temporal ICA with μ = 0.1-0.2 (Supplemental Experimental

Procedures) followed by image segmentation, we extracted filters with morphologies of Purkinje cell dendrites that showed spiking dynamics (Figures 5A and 5B). There were also filters with more isotropic profiles and slower dynamics representing Ca^{2+} activation of Bergmann glial fibers (Figures 5A and 5B) [287]. Bergmann glial fibers are known to be intertwined with Purkinje cell dendrites [4, 293], so there were many pixels containing the activity of both cell types. Thus, it is very challenging if not impossible to separate Purkinje cell and Bergmann glial signals by ROI analysis. Yet, by automated cell sorting we disentangled Purkinje cell Ca^{2+} spikes (Figure 5B, *top four traces*) from stereotyped Bergmann glial Ca^{2+}-transients (Figure 5B, *bottom four traces*).

Spontaneous glial signals identified by cell sorting had mean event rates that were much less frequent than neuronal Ca^{2+} spike rates in both awake and anesthetized mice. The time course of glial Ca^{2+} activation conformed closely ($r = 0.9 \pm 0.1$, s.d.; $n = 96$ events) to a double exponential function with similar time constants for the rise (1.5 ± 0.7 s) and decay (1.8 ± 1.2 s). These kinetics resembled those observed previously in Bergmann glia in cerebellar slices and in awake mice [287, 294, 295]. Despite the spatial overlap of neuronal and glial signals and their distinct dynamical timescales, ICA separated the different signals for the two cell types while minimizing cross talk.

After isolating activity traces from individual Purkinje cells, we estimated the cells' Ca^{2+} spike trains in digital format. This involved temporal deconvolution to remove the slow decay of Ca^{2+}-dependent fluorescence following each spike and then labeling each instance of a positive-going threshold crossing as a spike [224] (Figure 5B). The deconvolution accounts for the kinetics of Ca^{2+} unbinding from the fluorescent indicator [279].

To quantify accuracies of spike detection in live subjects, we combined two-photon microscopy with simultaneous extracellular single unit electrical recordings of Purkinje cell activity in anesthetized mice (Figure 5C) (Experimental Procedures). On the working assumption the electrophysiological traces provided perfect records of complex spiking, we tested our analytical ability to extract complex spikes from the optical data. (This assumption is in general false, and imperfect electrical traces would lead to underestimation of the accuracy of optical spike detection). A natural analysis for this test is receiver operating characteristic (ROC) analysis, which compares the probabilities of correct and incorrect spike identification (Figure 8A) [296].

These probabilities each depend on the choice of threshold used for spike detection, and the ROC curve summarizes the tradeoff between sensitivity and accuracy in a two-dimensional plot (Figure 8B). The area under the ROC curve, a number between 0.5 and 1, equals the probability that an ideal classification algorithm would correctly discriminate a randomly selected time frame with a spike from one without a spike [296]. This measure is not dependent on the spike threshold. In our data, areas under the ROC curves were 0.84 ± 0.06 (mean ± s.d.; n = 7 cells) for image segments and 0.92 ± 0.05 for raw ICA signals. This demonstrates an ability to achieve ~85-90% discrimination accuracy. We also examined absolute false positive and false negative spike detection rates as a function of the spike detection threshold (Figure 8C). Both error rates in our data sets were comparable with those in recent imaging studies [40, 297] and with tetrode recordings [222] and were superior to the reported error rates of 23-30% for manual spike sorting by trained experts in simulated micro-electrode array recordings [298].

Following these validation studies in anesthetized mice, we further examined data from awake mice and extracted up to ~100 Purkinje cell dendritic segments from fields up to ~500 μm across (Figures 5D and 5E). Our automated algorithm efficiently extracted the spike trains from cells tiling the entire field of view, many of which were closely adjacent and extremely difficult to separate by manual methods.

Locomotor Behavior Increases Complex Spike Rates of Purkinje Cells

A longstanding goal of cerebellar physiology has been to determine how ongoing activity in the olivo-cerebellar circuit is modulated during motor behavior. We reasoned that automated cell sorting, when combined with two-photon microscopy in behaving mice, would permit detailed examination of the relationships between Purkinje cell complex spiking and locomotion.

Using automated cell sorting we extracted Ca^{2+} signals from Purkinje cells in cerebellar lobules V and VI in head-restrained mice allowed to run voluntarily on an exercise ball. Mean rates of complex spiking varied between periods of anesthesia, alert rest, and active locomotion for each mouse (Figure 6). For individual cells studied under all three conditions, spike rates rose from 0.48 ± 0.27 Hz (mean ± s.d., n = 199 cells in 5 mice) under isoflurane anesthesia, to 0.76 ± 0.15 in alert but resting mice, to 1.0 ± 0.18 Hz in awake mice during active locomotion (Figures 6C and 6D) ($p<0.001$ for all pairwise comparisons; Wilcoxon signed rank test). Thus, gross activity levels of complex spiking depended on behavioral state, consistent with recent optical studies [23]. Prior electrophysiological studies of locomotion in decerebrated cats have been ambiguous on this point [25].

In isoflurane-anesthetized mice, we occasionally saw adjoining groups of Purkinje cell dendritic trees with little or no complex spiking (rate < 0.01 Hz). For example, in Figure 6B dendritic trees numbered 1-8 and 35-42 by medio-lateral position were largely silent under anesthesia. This was not so for all cells, as shown by dendrites 9-34 that spiked at 0.38 ± 0.15 Hz (n=25 dendrites). After termination of anesthesia, all Purkinje cells (1-42) in the field of view resumed regular Ca^{2+} spiking (Figure 6B). This example highlights the combined utility of automated sorting and Ca^{2+}-imaging in behaving animals by revealing phenomena in dense cellular networks inaccessible to electrical recordings.

Cerebellar Microzones Remain Anatomically Stable Across Behavioral States

Our studies of the cerebellar vermis in behaving mice allowed us to reexamine the organization of parasagittal bands of correlated Purkinje cells noted previously by imaging in anesthetized animals [40, 281] and electrophysiological means [44, 89]. Using Ca^{2+}-imaging in awake behaving mice we examined whether microzones have stable anatomical boundaries [44] or represent flexible dynamic ensembles that vary across behavioral states [46].

By studying pairwise correlation coefficients for Ca^{2+}-related fluorescence signals, or for the corresponding spike trains, we repeatedly observed enhanced correlations for multiple, closely situated pairs of Purkinje cell dendrites as compared to more distally separated pairs (Figures 7A and 7B). We looked for the anatomical boundaries of such microzones of highly correlated cells by clustering Ca^{2+}-spike trains using the pairwise correlation coefficients (Figure 7C and 7D) [40] (Experimental Procedures). In alert but resting mice, cluster analysis partitioned cells into microzones with higher correlations for intra-zone pairs ($r = 0.10 \pm 0.08$ mean ± s.d.; n=1418 pairs in 5 mice) than inter-zone pairs (0.02 ± 0.03, n = 2474; $p<0.001$ Wilcoxon rank sum test). Strikingly, the boundaries between microzones were sharply delineated, with the spatial transition between microzones generally occurring within one cell width rather than in a gradual manner over multiple Purkinje cell dendritic trees.

Like the rate of complex spikes, the strength of pairwise correlations between complex spike trains varied between anesthetized, alert but resting, and actively moving behavioral states. The correlations we observed among cells within each microzone grew stronger during active movement ($p<0.001$ Wilcoxon signed rank test) both for intra-zone cell pairs (0.20 ± 0.09), and to a statistically significant but far lesser extent for inter-zone pairs (0.03 ± 0.05), as compared to alert rest. When considered together with the rises in spike rates across the field of view, this distinction in synchrony between intra- and inter-zone cell pairs shows that there are different degrees of correlated input to intra-zone versus inter-zone pairs and that the rise in synchrony is likely not simply a consequence of increased spike rate in independent cells. Likewise, it would be hard to argue that the increase in measured spike rates during locomotion was an artifact due to noise, since noise would not lead to precisely defined microzones of high pairwise synchrony.

Compared to either rest or locomotion, isoflurane anesthesia reduced correlations dramatically (0.03 ± 0.06 intra-zone; 0.001 ± 0.02 inter-zone). In each microzone we found that nearly all dendrite pairs were significantly correlated ($p<0.01$, likelihood ratio test) in resting (86% of pairs) and actively moving mice (95%), but only 26% of intrazone pairs were significantly correlated under anesthesia. By comparison, less than a third of dendrite pairs from different microzones were significantly correlated in mice at rest (29%), during movement (33%) and under isoflurane anesthesia (7%). Taken together, these results show that both complex spike rates and pairwise synchrony within microzones increased during active movement as compared to alert rest or isoflurane anesthesia.

The organization of correlated Purkinje cell complex spiking was earlier proposed to be dynamically modulated during motor behavior [46]. We tested the temporal stability of microzones in two ways. First, we divided each experiment into 60 s epochs and performed cluster analysis separately for each epoch. We found that each Purkinje cell was assigned to the same microzone during 96 ± 6% of epochs (n = 44 epochs). Second, we tested if microzones changed their organization across different behavioral states. By comparing microzones found during locomotion on an exercise ball and alert rest, we found that 98.5% ± 0.8% (mean ± 68% confidence interval; n = 199 cells) of the cells fell in the same microzone in the two conditions (Figure 7D). Our results using automated cell sorting on cerebellar imaging data from mice behaving on an exercise ball thus fail to support a dynamic modulation of microzone boundaries, and instead provide evidence for a stable anatomical organization of temporal correlations in the olivocerebellar pathway.

DISCUSSION

The approach reported here for automatically extracting cellular Ca^{2+} signals meets an important growing need for neuroscience research. We discuss below the strengths and limitations of our approach as well as the implications of our physiological findings obtained by automated sorting.

A Growing Need for Automated Cell Sorting

Emerging techniques for Ca^{2+}-imaging across large populations of individual cells [225] have created new opportunities for densely sampling the dynamics of cellular networks [18, 28, 40, 224, 226, 270, 272, 273, 299, 300]. However, the raw data from large-scale Ca^{2+}-imaging also poses new analytic challenges that largely fall into three categories: identification of cells' locations, extraction of Ca^{2+} signals, and detection of neuronal spikes. Only the third problem has received considerable attention and several algorithms for detecting spikes within fluorescence traces are now available using temporal deconvolution [224], template matching [48, 267], or supervised machine learning [297]. All of these algorithms, however, rely on the assumption that the high-dimensional, raw movie data has already been transformed into paired sets of spatial filters and time traces for each cell. This transformation necessitates solving the first two challenges.

A reason for the focus to date on spike detection may be that the problem resembles familiar challenges from spike sorting of electrophysiological data [38]. Although Ca^{2+} transients evolve more slowly and suffer from different noise sources than voltage signals, in both cases spike detection transforms a relatively low-dimensional set of traces into digital spike trains. By comparison, the problems of cellular localization and signal extraction have few precedents in electrophysiology, since the high-dimensionality of Ca^{2+}-imaging data implies thousands of parameters must be confronted for every cell. Further, the spatial intertwining of cellular processes in many tissues inherently leads to cross talk among nearby cells falling within the same pixels.

To date, localizing individual cells within Ca^{2+}-imaging data has commonly been done by manual identification of cell bodies based on their morphologies in static fluorescence images [28, 48, 272]. Some studies have used semi-automated methods that rely on morphological filters to find cellular boundaries [301]. After finding structures with the sizes and shapes expected of cell bodies, these procedures define ROIs for each cell and average the fluorescence across each ROI to extract the cellular dynamics. However, with the most commonly used AM-ester conjugated Ca^{2+} indicators this approach generally fails to identify neuronal dendrites or fine glial processes, since these do not stand out with high contrast and so cannot be readily delineated by morphological filtering. Because dendrites and glial processes often blend into the background, except during periods of Ca^{2+} transients, an approach that relies on signal dynamics is warranted.

A recent study solved the cellular localization and signal extraction problems for Ca^{2+} signals from Purkinje cell dendrites by using temporal cross-correlation analysis to find groups of pixels contributing to each cell's signal [40]. A trained user first selects a region slightly larger than each dendrite and then computes the cross-correlation among the fluorescence signals of all pixels within this region. The final ROI contains all pixels that were highly correlated with a large number of other pixels in the user-defined region. Although this procedure used signal statistics to localize neuronal dendrites, it required 2-3 hours of manual pixel selection by trained users for each data set. As the number of cells in Ca^{2+}-imaging recordings extends into the hundreds [269, 272] and beyond, this need for human selection of pixels will become increasingly prohibitive. Our automated procedure takes only a few minutes of user supervision to set the number of principal components and to screen the extracted independent components.

The use of an ICA stage within our 4-stage sorting procedure (Figure 1B) for finding cellular Ca^{2+}-signals within raw data containing $\sim 10^5$ simultaneous time traces builds upon prior applications of ICA in the analysis of physiological data of other types and of far fewer dimensions. ICA extracted unitary voltage signals from studies using a photodiode array to study the sea slug [42]. Since there were only 448 diodes in the array, the dimensionality of the voltage data was far smaller than that of the Ca^{2+}-imaging data explored here and did not require dimensional reduction. Two studies have used PCA followed by ICA to extract hemodynamic signatures and functional domains from optical imaging data [280, 302], although the signals originated from large aggregates of cells and so often involved retention of just a few (*e.g.* ~6) principal components [280].

Our approach builds upon and differs from this previous work by modifying prior ICA methodologies to incorporate both spatial and temporal statistics and by following ICA with another analytic stage for image segmentation. Due to these additions we achieved the first automated extraction of Ca^{2+} signals from individual cells, despite the high dimensionality of our raw data. Thus, our approach is uniquely suited for the analysis of data sets even larger than those examined here and could be deployed to extract signals from 3D-imaging data containing hundreds or even thousands of cells [272]. Our procedure is grounded in three suppositions: (1) cellular signals are mathematically separable into spatial and temporal components; (2) signals from different cells are statistically independent; (3) cellular spatial filters and temporal signals possess highly skewed distributions. Interestingly, our procedure proved to be effective under conditions in which there are modest deviations from these conditions. Below we discuss the strengths and limitations that arise.

Cell Sorting Separates Movie Data Into Spatial and Temporal Components

The first supposition of our approach is that cellular signals are separable into corresponding pairs of spatial filters and time courses (Figure 1A). This is based on the observation that cellular Ca^{2+} signals within large-scale imaging data generally arise from fixed locations and do not convey details at the shortest physiological time scales regarding rapid intracellular propagation of $[Ca^{2+}]$ changes. For example, for Purkinje cell complex spikes, any delay between Ca^{2+} activation in different parts of the cell is generally briefer than the acquisition time for one image (50-100 ms), thereby permitting a separation of space and time in the description of these events. On the other hand, astrocytic Ca^{2+} waves that propagate far more slowly than neuronal Ca^{2+} spikes violate the assumption of separability [303]. Nonetheless, we found that automated cell sorting still extracts a useful approximation of cellular signals even in the case of radially expanding Ca^{2+} waves in Bergmann glia that propagate over a fraction of the field of view (Figures 5A and 5B). This suggests ICA cell sorting remains a useful tool for analyzing weakly non-separable signals. However, with waves that propagate over long distances other analytic tools seem more appropriate.

Spatio-temporal separability may also be violated *in vivo* due to movement artifacts. Although lateral displacements can be readily corrected by image registration algorithms [28, 304], residual axial displacements can corrupt physiological signals. Although such artifacts violate the assumptions of our cell sorting procedures, we found that ICA can identify components that arise due to movement artifacts and thus separate them from the physiological signals. This confirms prior work showing that ICA can remove movement artifacts [280, 305].

Automated Cell Sorting is Robust to Dynamical Correlations in Cells' Activities

ICA assumes that cells' signals are statistically independent. Since no approach to cell sorting can separate completely synchronized sources without taking into account geometric information, we added an image segmentation stage following ICA. The result is a robust, hybrid procedure that can be applied even in cases of weak or strong cellular correlations. This hybrid handled strong levels of correlation $(r \sim 0.9)$ in our simulations and with cerebellar data successfully separated highly correlated dendrites that clearly belonged to different Purkinje cells based on their anatomical separation (Figure 9). High levels of complex spike synchrony are known to arise for distinct Purkinje cells with inputs from the same climbing fiber [306].

Interestingly, for cell pairs with only weak correlations ($r < 0.8$), image segmentation was unnecessary. Although correlated activity violates the assumption of independence, weak correlations between cell pairs did not prevent the ICA stage from identifying the two cells as separate entities. For example, Purkinje cells that receive different climbing fibers still often exhibit weakly correlated complex spiking [89]. This weak synchrony, with $r \sim 0.1$ [23, 40], is thought to arise due to gap-junction connections in the olivary neurons that project the climbing fiber axons [307]. The ICA stage of our analysis was not misled by this weak synchrony and correctly identified individual cells. Overall, the ability of our 4-stage procedure to find individual cells, even in cases of weak or strong synchrony, is an important aspect of the procedure's general utility that should enable applications to studies of many brain areas.

Cell Sorting Based on Statistical Skewness

A main ingredient in our approach is the assumption that the user seeks signals with skewed distributions, as mathematically embodied in the ICA objective function. Skewness is a measure of the asymmetry of a signal's distribution that can be large for sparsely active cells but that can also successfully characterize signals occurring more densely in time. An assumption of skewness also leads to some limitations.

Spatial skewness is highly valid for data in which the individual active cells occupy only a small fraction of the pixels across the field of view. Temporal skewness poses interesting tradeoffs. In our own recordings of Purkinje cell activity, the assumption of skew temporal statistics was satisfied since the mean time between Ca^{2+} spikes was greater than the duration of the individual spikes. However, such sparse activation in time is not the only means of achieving a highly skewed distribution of signal amplitudes. For example, cells with high rates of activity that are strongly modulated in time, such as during bursts, could also exhibit skewed distributions of fluorescence signal amplitudes. Thus, our approach should be well suited to neurons that exhibit irregular spiking, such as high-frequency spike bursts interrupted by periods of quiescence [28, 48, 267-270, 300, 301, 308]. However, since both PCA and ICA may ignore cells with extremely low activity levels, our 4-stage algorithm fails to identify cells that remain inactive throughout the duration of a movie. This could lead to underestimation of cells' densities and overestimation of mean activity levels in areas where many cells are largely silent [267]. At another extreme, high rates of tonic, regular Ca^{2+} activation without bursts will also lead to temporal distributions of signals that are not skew, potentially leading ICA to ignore such cells.

For example, in our own recordings we had difficulty extracting the cerebellar interneurons present in our data by spatio-temporal ICA [226], since these cells have high spontaneous spike rates compared to the image acquisition rate and do not show bursty Ca^{2+} dynamics. A purely spatial ICA combined with image segmentation might be better suited to extract these cells. Alternatively, to find cells with activity patterns that are not temporally skew, a morphological approach might be better [301].

Automated Cell Sorting Can Improve the Design of Optical Experiments

Experimenters often choose optical parameters, such as illumination power and field of view size, based on intuitive notions of how images should appear to best facilitate data analyses. Our simulations of automated cell sorting indicate that these notions may often be incorrect, and that there may sometimes be more leeway than has been appreciated to reduce the numbers of photons and pixels per cell. This stems from the observation that there are broad regimes over which increases in signal to noise and decreases in field of view do not markedly improve the fidelities of automatically extracted Ca^{2+} signals (Figures 4E and 4F). In our artificial movies the fidelity of extracted signals was not improved substantially by increasing photon counts above a threshold level. This threshold was ~400 photons per pixel per time frame at a moderate field of view size (0.09 mm^2) containing ~100 cells, and increased to 10^4 photons for a 0.25 mm^2 field of view with the same number of pixels. Increasing photon counts above the threshold for high fidelity extraction may improve the subjective appearance of raw movie data, but it may not aid the ability to test hypotheses or make conclusions about brain dynamics. Similarly, our results should encourage recordings over larger fields of view than might otherwise have been used based on cells' subjective appearances.

This extension of experimental capabilities arises from the superiority of our ICA based cell sorting procedure over conventional ROI analyses. The latter typically require sufficient morphological information to recognize cellular outlines. Sorting by ROI analysis generally will not work well if each cell occupies only a few pixels or blends into surrounding areas. These conditions need not restrict ICA, which recognizes cells based on their activity signatures, not their appearances, and can readily avoid cross talk even for cells that share overlapping pixel sets. ICA's analytical advantages thus aid experimental capabilities. Overall, our studies highlight the value of benchmarking analysis procedures against simulated data as part of the experimental design process. Simulations allow experimenters to plan studies that optimally use scarce photons and pixels to test specific hypotheses about cellular dynamics.

Cerebellar Microzones As Stable Anatomical Entities

Automated cell sorting allowed us to re-examine a longstanding debate on the spatial organization of Purkinje cells' complex spiking activity. One view holds there are flexible assemblies in the cerebellar cortex that alter their composition of Purkinje cells with synchronous complex spikes according to changing behavioral requirements [309]. In support of this view, microelectrode array recordings of complex spiking in awake rats during repetitive licking suggested that different groups of spatially separated cells were synchronously activated at distinct times corresponding to different phases of movement [46]. An opposing view holds cerebellar microzones are defined by the anatomical organization of climbing fiber projections and provide a stable architecture for controlling different parts of the body [47]. In support of this latter view, physiological mapping studies revealed that strips of neighboring Purkinje cells receive climbing fiber inputs with similar receptive fields for cutaneous stimulation [44, 45, 47]. These maps are intricate but reproducible across subjects [310]. Microzones have been recently observed with single cell resolution using Ca^{2+}-imaging of Purkinje cell dendrites in anesthetized mice [40].

By extracting complex spike trains from populations of >40 individual Purkinje cells during locomotion on an exercise ball and during alert rest we examined the stability of microzones' cellular composition. Our data revealed that both microzone boundaries and cellular compositions were stable across these physiological states and maintained sharp boundaries with single cell precision. Although the borders between microzones remained fixed, both the rates of complex spiking and pairwise correlation coefficients rose during locomotion as compared to alert rest. The substantial discrepancy in synchrony between pairs of dendrites within and across microzones, as well as the sharp microzone boundaries identified by cluster analysis, argue strongly against an interpretation of our results as being due to noise, which would not lead to precisely defined regions of high pairwise synchrony. Our results are also consistent with recent data from a study that used high-speed miniaturized microscopy to image Purkinje cell Ca^{2+} spikes in freely moving mice and reported an increase in complex spike rates and pairwise synchrony at the level of population statistics [23]. By combining automated cell sorting and two-photon microscopy we provided here a more detailed view of Purkinje cell correlations during rest and locomotion.

Our results support an anatomical view of the organization of the olivocerebellar circuit, and we found no evidence of a dynamical reorganization during motor behavior. Still, it should be kept in mind that we only explored a single motor behavior and others will need to be examined. Although we cannot eliminate the possibility that microzones dynamically reconfigure under other circumstances, our data do considerably bolster the view microzones are stable anatomical entities [47]. Nonetheless, the evidence for dynamical reconfiguration of Purkinje cell assemblies [46, 311] should not be discounted; an attractive possibility reconciling the two viewpoints is that correlated assemblies of different microzones, each stably defined, dynamically reconfigure in a behaviorally dependent manner. It will also be important to examine correlations between more than two cells to see if there are larger cellular assemblies that form [46, 312]

Our data showed that pairwise correlations and spike rates were considerably higher in awake animals than anesthetized subjects, highlighting the importance of imaging in alert subjects. These results from cerebellar cortex contrast with recent work on neocortical circuits. In the rat visual cortex, spike correlations between layer 2/3 neurons decreased during periods of activity, as assessed by whisker or forepaw movements, as compared to rest [48]. This contrasts with our data, which revealed that during animal activity complex spike correlations increased. The contrast highlights potential differences in neural coding between the neocortex and the inferior olive, which drives Purkinje cell complex spiking via the climbing fiber pathway. Further, ketamine/xylazine anesthesia reduced spike rates but increased correlations among neocortical pyramidal cells [48], which is also the opposite of our observations on

Purkinje cell complex spike correlations under isoflurane anesthesia. However, a prior study in ketatmine/xylazine anesthetized rats also found significant correlations in Purkinje cell complex spiking [40]. These observations suggest different anesthetics might have differential effects on neuronal synchrony and highlight the importance of studying concerted activity patterns in awake animals across a range of behavioral conditions.

Technological Outlook

We expect continued refinement of automated sorting techniques will further increase the utility of optical imaging for studying dense cellular networks and will allow cell sorting to assume a comparably important role as in electrophysiology. By introducing an automated means for cell sorting based on broadly applicable statistical principles, we have created a framework for moving beyond current heuristic and semi-automated approaches. Further refinements of our protocol might incorporate statistical measures of sparseness other than skewness or *a priori* knowledge of various types beyond the one-parameter tradeoff between spatial and temporal skewness explored here. Another possible generalization might include information about sensory stimuli or animal behavior to derive functional characterizations of cellular activity in an automated way.

EXPERIMENTAL PROCEDURES

Automated Cell Sorting Procedures

We analyzed simulated and experimental Ca^{2+} fluorescence imaging data using *ImageJ* plug-ins and custom MATLAB routines. Movies were first corrected for lateral movement artifacts using *TurboReg* [313]. We then applied our automated cell sorting protocol (Figure 1A) comprising: (1) principal component analysis (PCA) for dimensional reduction; (2) spatio-temporal independent component analysis (ICA) for separation of intracellular cellular Ca^{2+} signals from sources of noise or artifacts; (3) image segmentation to separate activity traces from highly correlated cells; and, in the case of neuronal signals, (4) temporal deconvolution and spike detection to extract the times of action potentials. For details of each stage, see supplemental material.

Artificial Data Sets

We simulated Ca^{2+}-imaging data by combining artificial spike trains and glial transients with spatial filters designed to reflect the size, shape, and density of Bergmann glia and Purkinje cell dendritic trees as viewed in optical sections within the mouse cerebellar molecular layer. Each simulated movie had 1,000 time frames at 10 frames per second. We added a static image to the temporally modulated Ca^{2+} signals to represent background fluorescence arising from brightly labeled interneuron somata, as well as elongated dark regions representing blood vessels. These time-dependent and static background signals together defined a noiseless data set. We created noisy movie data with a defined signal to noise ratio, S, by multiplying the noiseless movie data in each pixel and each time frame by an independent random number chosen from a distribution with mean S^2 and variance S^2. This noise model approximates the expected photon distribution given an average of $\sim S^2$ collected photons per pixel per time frame using a shot-noise limited microscope under conditions of relatively uniform image contrast.

Animal Procedures

All animal procedures were approved by the Stanford Administrative Panel on Laboratory Animal Care. We used 12 male C57Bl/6 wild type mice (5.5-15 weeks old). Each experiment involved two surgeries. Several days before recordings a custom metal head plate was attached to the skull with dental acrylic. This allowed habituation of mice to accept head restraint while walking on the exercise ball and stable imaging/electrophysiological recordings in anesthetized and awake mice. On the recording day, a craniotomy was opened over the cerebellar vermis. Two-photon imaging in head-restrained mice was performed similarly as in [28] except that we used slightly different surgical procedures and the mice walked and ran on an exercise ball (Habitrail, model 62065) that rotated on one axis instead of two (Supplemental Experimental Procedures). We tracked mouse locomotion by using an optical encoder to monitor the exercise ball's rotations.

Two-photon Fluorescence Imaging

We loaded multiple neurons and glia with the fluorescent Ca^{2+}-indicator Oregon Green 488 BAPTA-1-AM (OGB-1-AM; Molecular Probes) as described [32, 225, 226]. We used a custom two-photon microscope equipped with an ultra-short pulsed Ti:sapphire laser tuned to 800 nm. See Supplemental Experimental Procedures for details.

Analysis of Cerebellar Microzones

Following extraction of complex spike trains from populations of Purkinje cells, we analyzed spiking correlations among all cells during anesthetized, awake but resting, and actively moving conditions. We used the Pearson correlation coefficient to compare binary spike trains for each pair of cells. We grouped together cells with similar spike trains using k-means clustering [314].

FIGURES

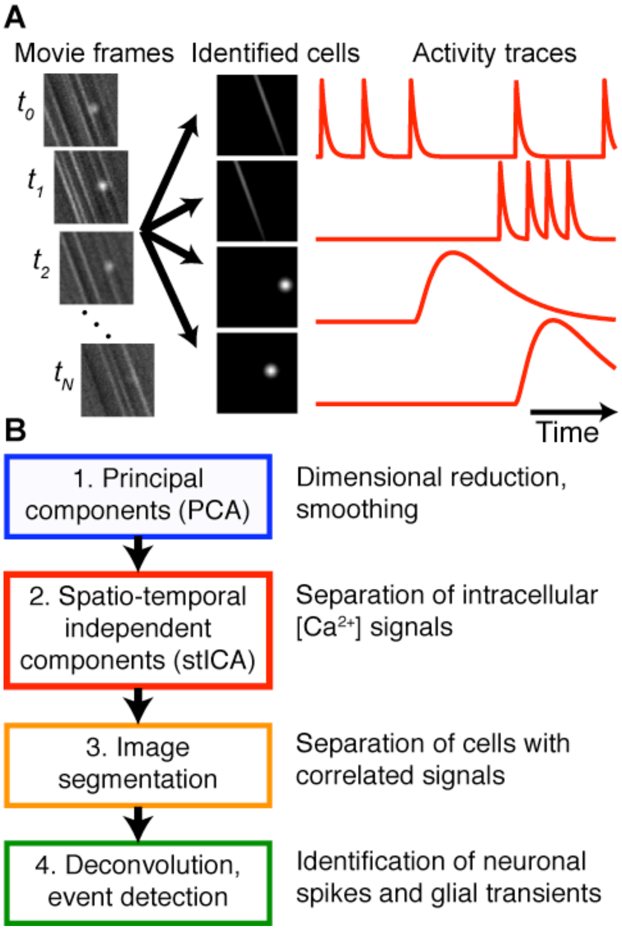

Figure 1. Automated Cell Sorting Involves a Defined Sequence of Analytical Stages.

The goal of automated cell sorting is to extract cellular signals from fluorescence imaging data sets (left) by estimating spatial filters (middle) and signal time courses (right) for each cell. The example depicts typical fluorescence transients in cellular structures of the cerebellar cortex as observed in optical cross-section. Transients in Purkinje cell dendrites appear as extended stripes while transients in Bergmann glial fibers tend to be more localized, appearing ellipsoidal in cross-section.

Automated cell sorting incorporates four stages of processing to meet specific analysis challenges.

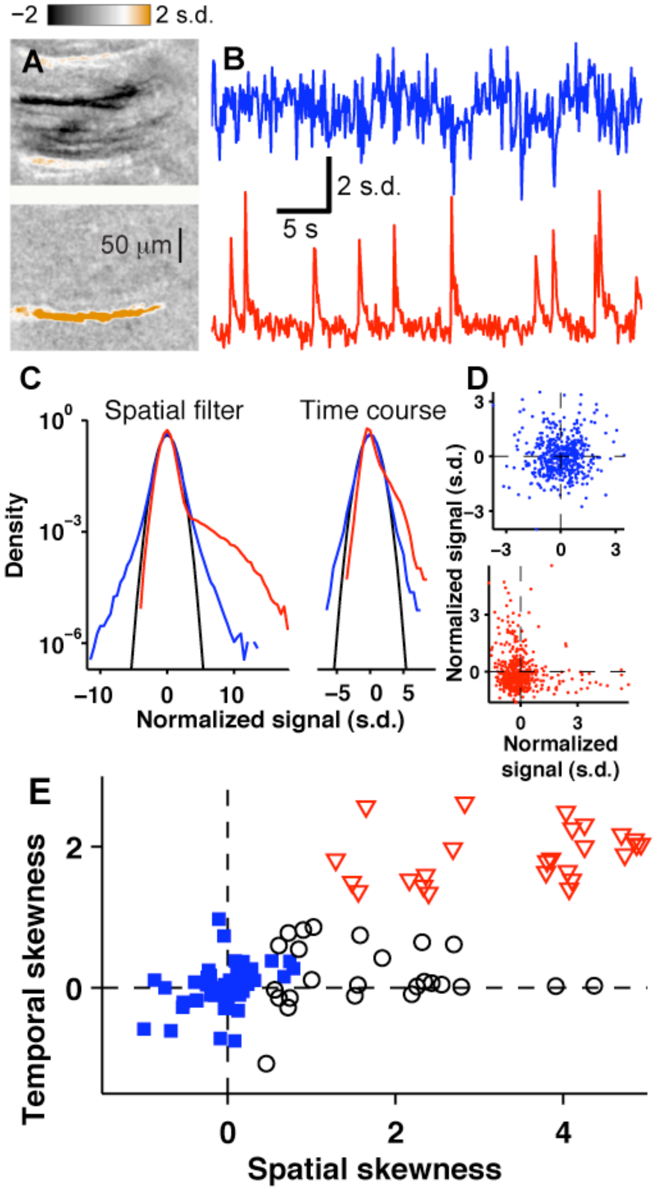

Figure 2. Intracellular Ca^{2+} Signals Are Sparse, Independent Components.

(A) Example spatial filters of individual principal (top) and neuronal independent (bottom) components identified within two-photon Ca^{2+}-imaging data from the cerebellar cortex of a live mouse. The independent component more accurately captures the form of the Purkinje cell dendritic tree (orange region).

(B) Signal time courses for example principal (blue) and independent (red) components shown in A. The latter time course provides a superior representation of the Purkinje cell's Ca^{2+} spiking activity.

(C) Distributions of intensities of pixels in spatial filters for the principal (blue) and independent (red) components shown in A (left). Distributions of the two components' time courses, taken across all movie frames (right). Gaussian distributions (black) are shown for comparison.

(D) Joint distribution of the time course values for two example principal (blue) and independent (red) components. Extended tails in the distribution of the independent components reflect time frames during which calcium dependent fluorescence increases transiently.

(E) Temporal vs. spatial skewness for 50 principal (closed blue squares) and 50 independent (open symbols) components. Cellular signals (open red triangles) have more skewed distributions than other independent components (open black circles), such as motion artifacts, or principal components.

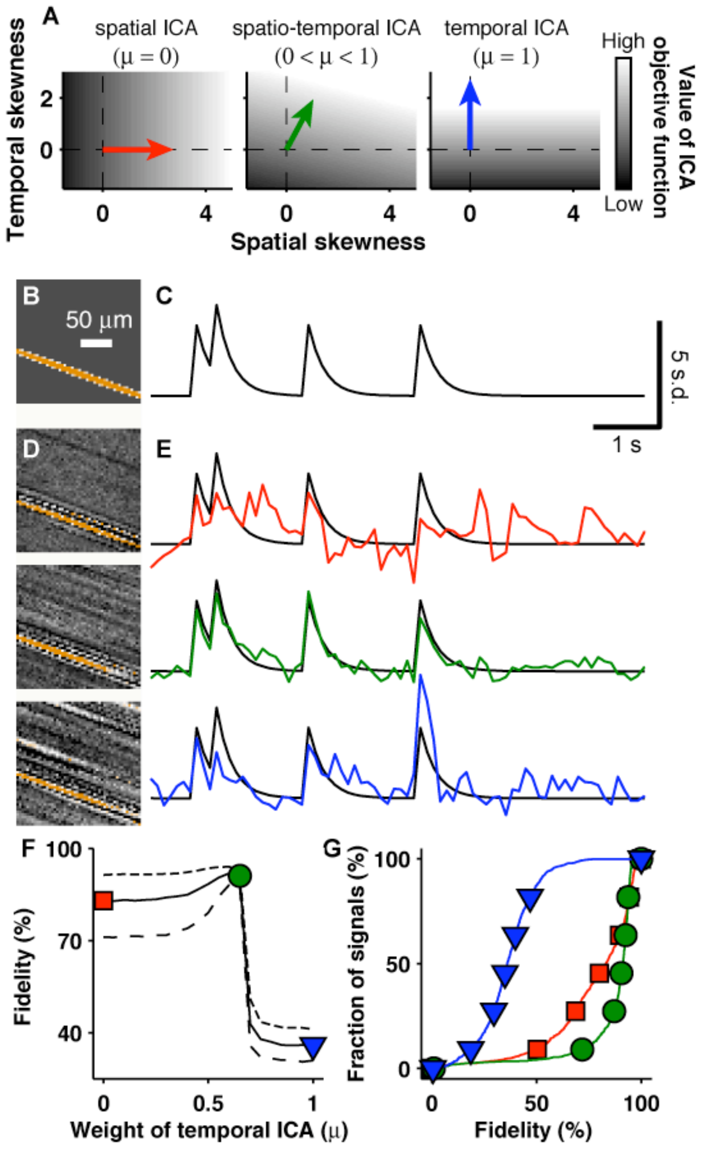

Figure 3. Combining Spatial and Temporal Information Improves Cell Sorting.

(A) Schematic illustration of the objective function (grayscale) maximized by spatial (left), temporal (right) and spatio-temporal (middle) ICA as a function of spatial and temporal signal skewness (compare with Figure 2E). Arrows indicate the combination of spatial and temporal skewness corresponding to the steepest gradient of the objective function.

(B,C) An example spatial filter (B) and signal time course (C) from an artificial data set that mimics Ca^{2+}-imaging data containing the activity of large numbers of Purkinje cell dendritic trees.

(D,E) Estimates of artificial signals extracted by spatial (top row, red trace), temporal (bottom, blue) or spatio-temporal ICA (middle, green). The spatio-temporal ICA yields a trace of higher fidelity to the underlying cellular signal than either spatial or temporal ICA.

(F) Percentile curves showing the median fidelity (solid line) of the signals recovered by ICA, as a function of the relative weight of temporal information (μ). Dashed and dotted lines indicate the range of fidelity values for signals within one standard deviation of the median.

(G) Cumulative distribution of the fidelity of recovered signals for the three different values of μ indicated by color in (F). Spatio-temporal ICA (green curve) yields the largest percentage of signal traces with high fidelity.

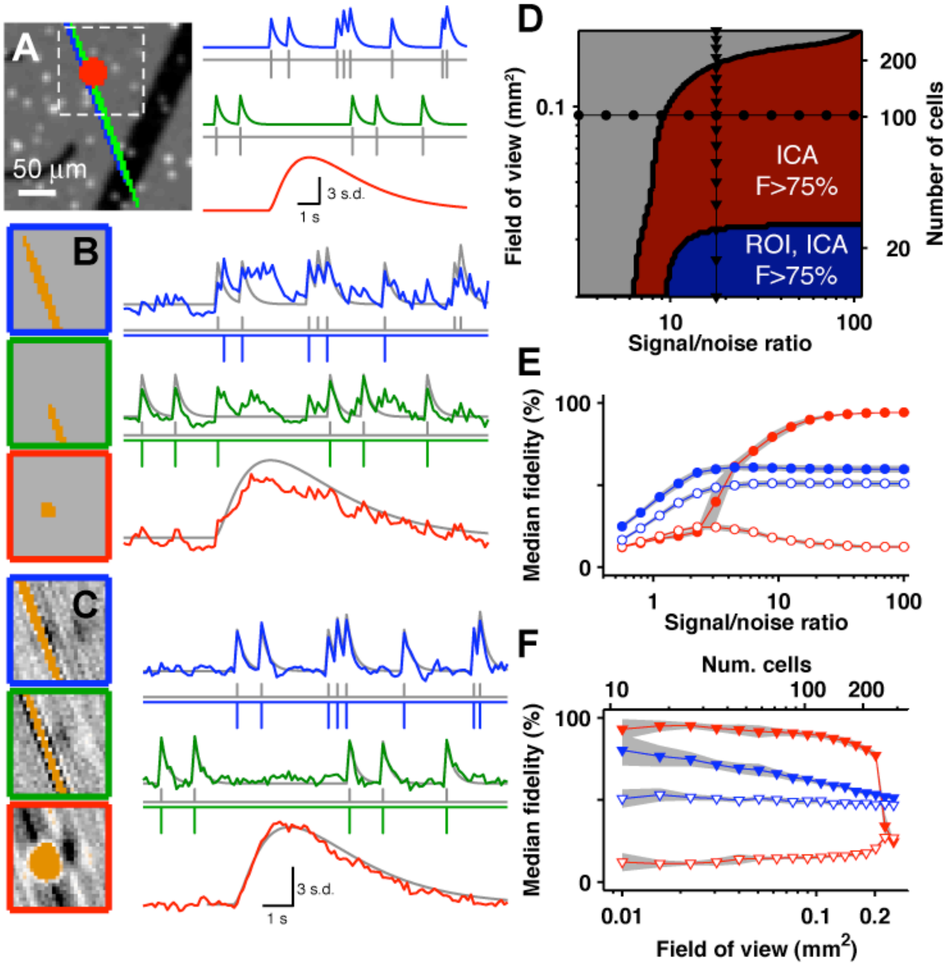

Figure 4. Automated Cell Sorting Outperforms Region of Interest (ROI) Analysis

(A) Three spatially overlapping sources of cellular signals within an artificial data set mimicking large-scale Ca^{2+}-imaging data from the cerebellar cortex. Background grayscale image shows simulated field of view containing dark blood vessels and brightly labeled interneuron somata. The signals from two Purkinje cell dendritic trees (blue, green) and a set of Bergmann glial fibers (red) show independent time courses and distinct temporal dynamics. Ca^{2+} spikes in the Purkinje cell dendrites (gray tick marks) cause the fast fluorescence transients.

(B) Spatial filters (left) and signals (right colored traces) identified by ROI analysis. The spike-triggered average fluorescence change for each cell is smoothed and thresholded to define the corresponding spatial filter. True spike times (gray ticks) match some, but not all, of the spikes estimated from the ROI signals (blue and green ticks).

(C) Spatial filters, signals and spike times estimated by spatio-temporal ICA with (μ=0.5). The higher fidelity signal estimate allows correct identification of all dendritic Ca^{2+} spikes.

(D) Diagram on logarithmic axes of how median signal fidelity depends on the signal/noise ratio and the field of view size. Colored regions show parameter ranges in which the true signals are estimated with >75% median fidelity by ROI and ICA (blue); by ICA but not ROI (red); or by neither method (gray). Black circles mark the parameter values used in (E); black triangles mark the parameter values used in (F).

(E,F) Median fidelity (solid symbols) and crosstalk (open symbols) of signals extracted by ICA (red) and ROI analysis (blue), as a function of the signal/noise ratio (E) and the field of view (F). The field of view is fixed at 0.09 mm^2 for (E); the signal/noise ratio is fixed at 18.0 for (F).

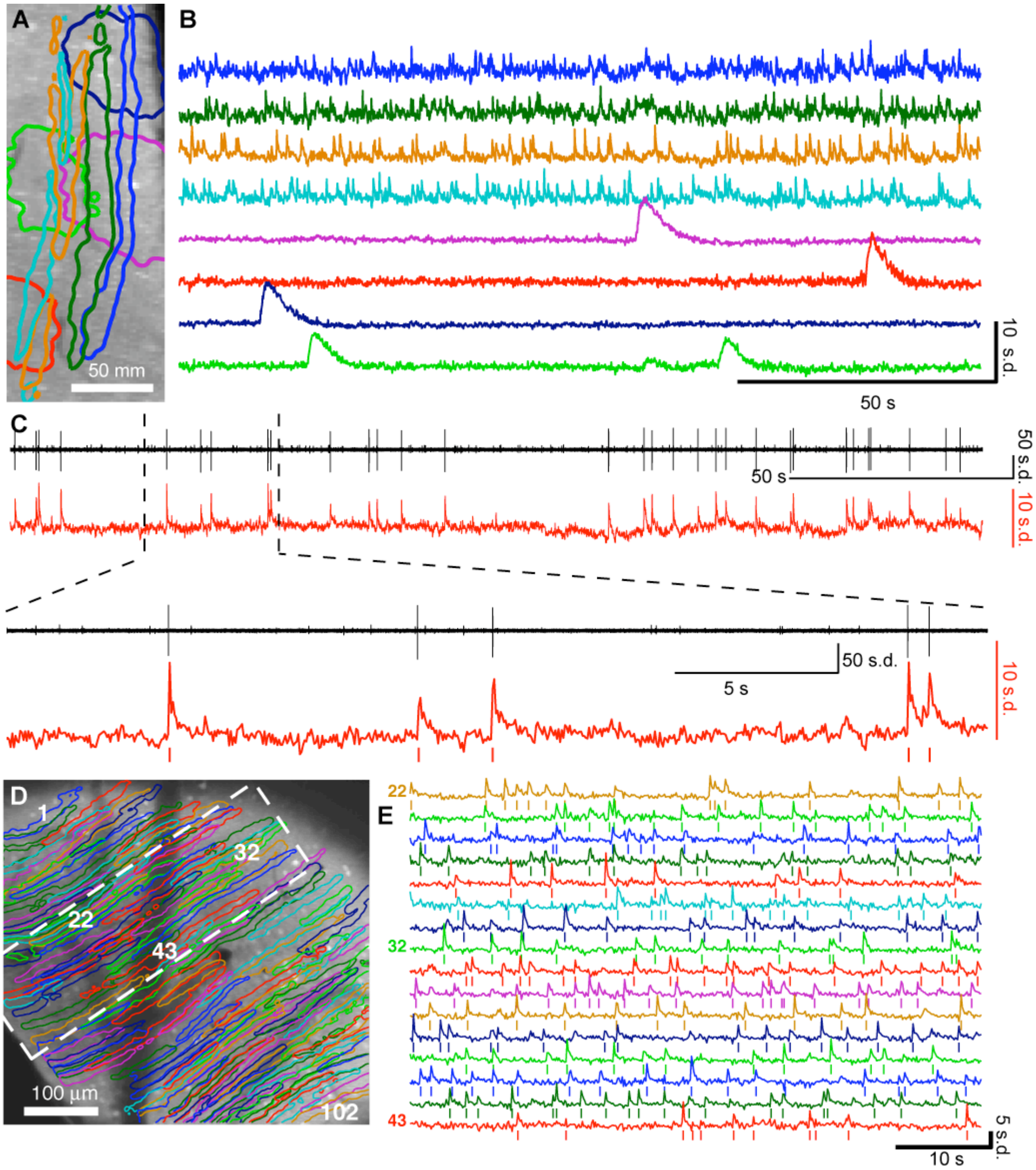

Figure 5. Automated Cell Sorting Identifies Neuronal and Glial Ca^{2+} Dynamics from Large-Scale Two-photon Imaging Data.

(A) Contours of seven spatial filters corresponding to Purkinje cell dendrites identified by automated cell sorting, as well as two independent component characteristic of the activity of Bergmann glial cells. Contours are shown superimposed on an image of the background fluorescence of the imaged region. The spatial filters partially overlap, with many pixels sharing signals from both neuronal and glial sources.

(B) Neuronal (top) and glial (bottom) temporal signals corresponding to the spatial filters of (A) demonstrate that ICA successfully suppresses crosstalk between the signals of nearby, intermingled cells.

(C) Comparison of extracellular single unit electrical recording from a Purkinje cell (black) with a signal extracted by cell sorting on imaging data recorded simultaneously (red). Asterisks indicate complex spikes identified by waveform based spike sorting.

(D) Contours of 102 spatial filters corresponding to Purkinje cell dendrites, as identified by automated cell sorting. Data were recorded in the superficial molecularl layer of an alert, restrained mouse.

(E) Example time courses of 16 independent component signals from a subset of the components within the outlined region in (A) containing cells 22-43. Tick marks represent the estimated neuronal spike trains. Each spike was identified following a temporal deconvolution to correct for the dye's Ca^{2+} binding kinetics as the occurrence of a positive-going threshold crossing within the activity traces.

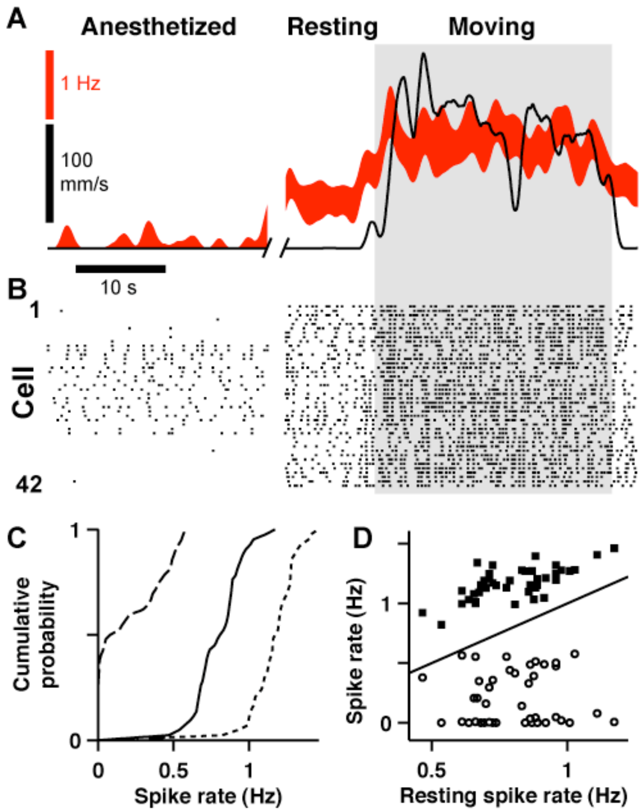

Figure 6. Ca^{2+}-Spiking in Purkinje Cells of the Cerebellar Vermis Depends on Behavioral State.

(A) Average Ca^{2+} spike rate in Purkinje cell dendrites (red; median ± s.d.) as a function of physiological condition. Left, isoflurane anesthetized. Center, alert but resting. Right shaded period, actively moving. The black trace shows the mouse's running speed on the ball.

(B) Ca^{2+} spike rasters for all 42 cells identified within the field of view. Note that under isoflurane anesthesia only a subset of Purkinje cell dendrites shows Ca^{2+} spiking.

(C) Cumulative distribution of the spike rates across isoflurane anesthetized (dashed), alert but resting (solid) and actively moving (dotted) conditions.

(D) Comparison showing each cell's spike rate during movement (solid squares) or during isoflurane anesthesia (open circles) on the ordinate, plotted against resting spike rate on the abscissa. The diagonal line delineates equal spike rates under both conditions. As also apparent in (C), spike rates are generally higher in awake than in isoflurane anesthetized animals, and increase further during active locomotion.

Figure 7. Cerebellar Microzones Exhibit Changes in Correlated Activity but Stable Anatomical Organization Across Different Behavioral States.

(A) Matrices showing the correlation coefficient of spike trains recorded from each pair of cells within a population of 41 Purkinje cell dendritic trees. Cells are arranged by location along the medio-lateral dimension. Correlation matrices are computed separately for three physiological conditions. Red and blue outlines indicate local microzones identified by a clustering analysis of the correlation coefficients obtained during locomotion.

(B) Correlation coefficients for cell pairs within the two microzones (blue and red) marked in panel A and in different (black) micro-zones as a function of their medio-lateral separation for each behavioral state.

(C) Outlines of Purkinje cell dendrites identified by automated cell sorting in five mice. Colors show that the microzones identified through the cluster analysis respect the medio-lateral ordering of cells and have sharp boundaries. The example of panels A and B is shown at far right.

(D) Schematic diagrams of microzone structure for the same five mice as in C, comparing periods of active locomotion (top) to those when the animal was awake but resting (bottom). The diagrams are based on the results of automated cell sorting and cluster analysis as shown in C, and reveal the stable anatomical organization of microzones across both behavioral states.

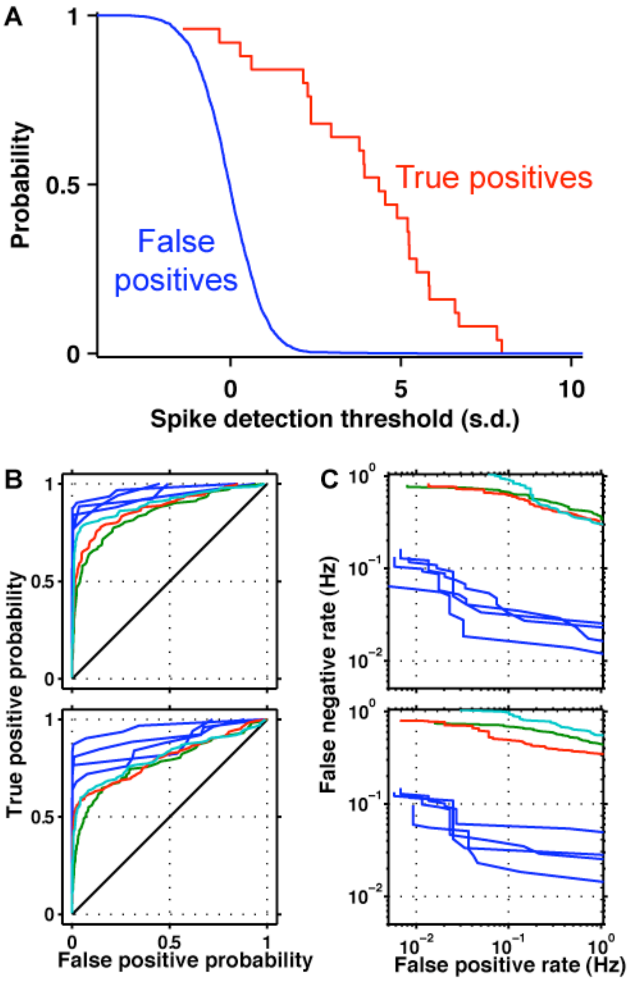

Figure 8. Simultaneous Ca^{2+} Imaging and Extracellular Recordings Quantify Discriminability of Complex Spike Trains.

(A) Cumulative distribution functions for the deconvolved fluorescence signal (normalized to unit variance) during all movie time frames coinciding with an electrically recorded complex spike (red) or during a background frame when no spike was present (blue). The detection threshold sets a particular probability of false positive errors and corresponding probability of correctly identifying true spikes.

(B) Receiver operating characteristic (ROC) analysis of the discriminability of complex spikes using independent component time courses (top) or signals resulting from image segmentation (bottom). Each curve illustrates the tradeoff between false positives and true spike detection for one of the cells studied by simultaneous electrical and optical recording. Different mice are indicated by the colors.

(C`) The same data as in (F) are displayed in terms of the rate (events per s) of false positive and false negative detection. The lower rate of complex spikes in one experiment (blue curves) improved discriminability in that subject.

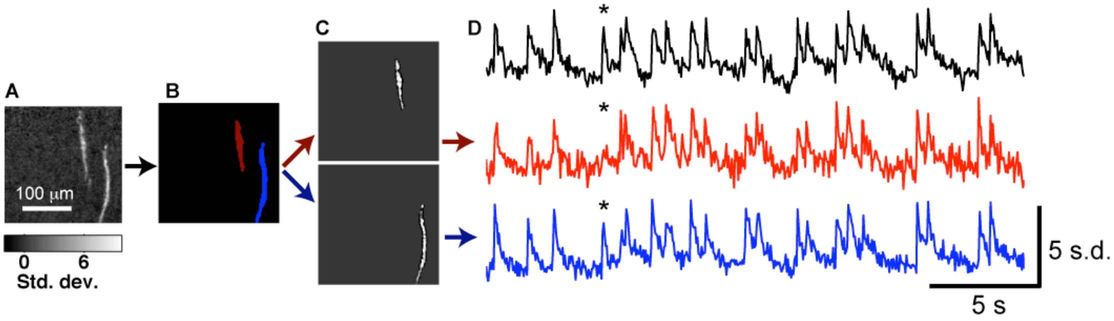

Figure 9. Image Segmentation Improves Extraction of Signals from Correlated Cells.

(A) An example of a spatial filter derived by ICA applied to an experimental data set. This component contains two well-separated groups of pixels that likely derive from dendrites of distinct cells.

(B) The first stage of segmentation smoothes the spatial filter by convolving with a Gaussian blurring function, then applies a threshold to create a binary mask showing regions contributing strongly to the signal. Each contiguous image segment is assigned to a new filter (red and blue regions).

(C) New spatial filters are created from the independent component by setting all pixel weights to zero except for those within one of the image segments. These pixels are assigned the same weight they had in the original ICA spatial filter.

(D) By applying the segmented spatial filters to the movie data, the ICA time course (black) is broken down into distinct contributions from each cell (red and blue). Segmentation reveals that the activity of the two cells is strongly correlated, as expected for cells that are grouped together by ICA. In this case the image segment spike trains are identical except for one spike (asterisk).

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Analysis Software

Data analysis was performed using custom software written in MATLAB or using *ImageJ* plugins. Some routines were modified from the publicly available *FastICA* package [286].

Data Analysis Procedures

Image Registration and Normalization

We first corrected any lateral movement artifact within each movie frame, as assessed relative to a reference frame, by using the image registration algorithm implemented in the *ImageJ* plugin *TurboReg* [313]. We cropped the movies to exclude any pixels that moved outside the field of view at any time during the recording. The resulting usable field of view was typically 1-14 pixels smaller in each spatial dimension than the original movie. We then normalized the signal in each pixel by dividing by its mean value over all movie time frames and subtracting 1. We next subtracted from each time frame the mean fluorescence averaged over all pixels. In cases when we compared data from the same animal and brain area under different physiological conditions (*e.g.* anesthetized versus awake), we spatially aligned the movies to a common reference frame, usually the first frame of the anesthetized recording.

Dimensional Reduction

We used principal component analysis (PCA) to reduce the dimensionality of the data and remove much of the noise that in each pixel causes intensity fluctuations uncorrelated to signals in other pixels [283]. Each movie was represented as a matrix, $\mathbf{M}$, of size $N_x \times N_t$, where N_x is the total number of pixels and N_t is the total number of time frames. The entries, $\mathbf{M}_{xt}$, contain the normalized fluorescence intensity in pixel *x* at time *t*. By using PCA we reduced the data to a lower-dimensional, approximate representation in terms of a set of *K* (where $K < \min(N_x, N_t)$) principal components (PCs). In practice, *K* ranged from 40 to 120 and was chosen separately for each data set (see below). Each PC comprises a spatial filter and a signal time course. The PC spatial filters, U_k, can be arranged in a $N_x \times K$ matrix with orthogonal columns, which we denote **U**. The PC time courses, V_k, similarly define a $N_t \times K$ matrix with orthogonal columns, which we denote **V.** After dimensional reduction, the new data representation approximates the original as a linear combination of the PCs,

$$\mathbf{M}_{xt} \approx \bar{\mathbf{M}}_{xt} \equiv \sum_k S_{kk} U_{kx} V_{kt} = \left[\mathbf{USV}^T\right]_{xt} ,$$

where **S** is a $K \times K$ diagonal matrix whose *k*'th diagonal entry squared, S_k^2, is the variance of the data after projecting along the *k*'th PC. PCA is defined by choosing the matrices, **U, V** and **S,** that minimize the mean square difference between the exact and approximate descriptions of the data,

$$\varepsilon\{\mathbf{U},\mathbf{S},\mathbf{V}\} = \sum_{x,t} (\mathbf{M}_{xt} - \bar{\mathbf{M}}_{xt})^2, \quad \{\mathbf{U},\mathbf{S},\mathbf{V}\} = \arg\max \varepsilon \quad .$$

We solved this optimization problem by computing the $N_t \times N_t$ temporal covariance matrix, $\mathbf{C} = \mathbf{M}^T\mathbf{M}$. (Recall that the mean time course was previously subtracted from the data). The PC time courses, $\{V_k\}$, are the *K* eigenvectors of **C** corresponding to the *K* largest eigenvalues, denoted S_k^2. In practice we sometimes excluded some of the PCs with the largest eigenvalues, sacrificing some of the accuracy of the reduced description of the data for the sake of rejecting some high-variance noise sources. We thus typically retained all components with index $k \in [K_{low}, K_{high}]$, so that the dimensionality of the reduced data was $K = K_{high} - K_{low} + 1$. Since we were only interested in the reduced data set, we needed to compute at least $\sim K_{high}$ eigenvectors and eigenvalues, not the full set of T components. We usually computed 200 PCs for all data sets in a batch pre-processing step, and then chose the two numbers $\{K_{low}, K_{high}\}$ for each data set in a subsequent semi-automated fashion (see below). Once we determined **V** and **S** in this way, we computed **U** using: $\mathbf{U} = \mathbf{MVS}^{-1}$.

This procedure for PCA, in which we computed the temporal rather than the spatial covariance matrix, allowed us to analyze efficiently data sets with a very large number of pixels. This was because **C** and **U** were computed by summing over all movie pixels, which was performed by loading the data from only a small number of movie pixels at each stage. There was therefore no need to store the entire data matrix, **M**, in memory at any time. The limiting resource for this procedure is the amount of memory required to store the matrix **C**, which scales as N_t^2.

An important step in performing PCA is the choice of the two values, $\{K_{high}, K_{low}\}$, that specify the range of PCs retained in the reduced data set. To set these parameters we calculated the set of variances, $\{S_k^2\}$, of the data principal components. We used the analytic formula for singular values of a Gaussian random matrix to find a reference set of variances for a movie of the same dimensions containing independent, Gaussian distributed random noise in each pixel and time frame [284]. By choosing the power of the noise in the reference movie so that the highest-order principal component variances match those of the data, we could compare its variance spectrum with the data to determine the data principal components whose variances were well above the noise floor. We chose the cutoff, K_{high}, to be roughly the point at which the slopes of the signal and noise spectra converged. All data dimensions whose covariance rank was greater than K_{high} were eliminated from further analysis. To reduce artifacts due to motion and galvanometer jitter, which may have large variance, we visually examined the spatial filters (the rows of **U**, reshaped and displayed as 2D images) and chose a lower cutoff, K_{low}, to eliminate modes with large variance that did not appear to contain biological signals. In practice, we usually set $K_{low} \leq 15$. With experience, setting $\{K_{high}, K_{low}\}$ can be done quickly and requires only a few minutes.

Following dimensional reduction, all subsequent analysis made use of a normalized, or whitened, representation of the data with equal covariance along every PC: $\mathbf{M}_{white} = \mathbf{U}\mathbf{V}^T$ [286]. Working with this representation simplified the process of extracting cellular signals as orthogonal transformations of the PCs.

Cell Sorting by Independent Component Analysis (ICA)

The problem of cell sorting, as schematized in Figure 1A, can be described by the generative model for signal mixtures that is at the heart of the theoretical framework of ICA. According to this model, independent biological signals mix linearly to form the data matrix:

$$[\mathbf{M}_{white}]_{x,t} = \sum_j s_j(x) a_j(t).$$

Here the fluorescence at pixel x at time t is determined by the instantaneous amplitudes of the independent signal time courses, $a_j(t)$, and by a set of corresponding spatial filters, $s_j(x)$. In temporal ICA, $s_j(x)$ is regarded as a matrix of mixing coefficients for the independent signal sources, $a_j(t)$, that combine to form the mixtures observed at each pixel. In spatial ICA, these roles are reversed, with the $a_j(t)$ serving as mixing coefficients for the spatial sources defined by $s_j(x)$.

ICA seeks to estimate $a_j(t)$ and $s_j(x)$ by finding an unmixing matrix, **C**:

$$\hat{a}_j(t) = \sum_k \mathbf{C}_{jk}\mathbf{V}_{tk},$$

$$\hat{s}_j(x) = \sum_k \mathbf{C}_{jk}\mathbf{U}_{xk},$$

where the circumflex indicates the estimate based on the movie data. We find the unmixing matrix by maximizing statistical sparseness, as quantified by the skewness of the spatial and temporal signals. The skewness of a random variable, z, is related to its third order statistical moment:

$$Skew[z] = E[(z-E[z])^3]/E[(z-E[z])^2]^{3/2} = E[\tilde{z}^3]$$

where E denotes an expectation value over the set of observations of z, and $\tilde{z} = (z-E[z])/\sigma_z$ has been centered and normalized to have zero mean and unit variance. We performed ICA using the *FastICA* algorithm [286], which uses an approximation of the Newton-Raphson method to maximize the skewness objective function [292]. This objective function for joint analysis of a set of signals, $\{z^{(j)}\}$, is the total skewness, $\sum_j Skew[z^{(j)}]$. The algorithm starts from a random initial guess for the mixing matrix and iteratively updates it to bring it closer to the maximum, by using an estimate of the local slope and curvature of the objective function at each iteration. Assuming the data have been whitened during pre-processing (see above) and the mean of each signal has been subtracted, the rule for updating $\mathbf{C}$ is

$$\mathbf{C}_{ij}^{(n+1)} = \sum_k (z_k^{(i)})^2 y_k^{(j)}$$

where $\mathbf{C}^{(n)}$ is the estimated mixing matrix after the n'th iteration, $y^{(j)}$ is the mixed signal from PCA (*e.g.* $y^{(j)} = V_j$ or $y^{(j)} = U_j$; see below), and $z^{(i)} = \sum_j \mathbf{C}_{i,j}^{(n)} y^{(j)}$ is the unmixed signal whose skewness is to be maximized. After each update, the unmixing matrix is orthogonalized via $\mathbf{C} = (\mathbf{C}\mathbf{C}^T)^{-1/2}\mathbf{C}$, which ensures that the independent components will be uncorrelated with each other [292]. To complete the definition of the ICA algorithm used for cell sorting, we must specify the identity of the unmixed signals, $\{y^{(j)}\}$, which may include the principal component spatial filters and/or time courses.

In its most basic form, ICA can be applied to cell sorting by using either spatial or temporal statistics to assess sparseness. Temporal ICA relies on the assumption that the signals of interest are temporally sparse, *i.e.* with each cell being activated in a relatively small proportion of the movie frames and with a skewed amplitude distribution. In this case, the mixed signals are the principal component time courses, $y^{(j)} = V_j$, and the objective function is $F_{time} = \sum_j Skew[\hat{a}_j]$, where the expectation value is an average over movie time frames. Imaging data from populations of individual cells are generally also sparse in the spatial domain, with each cell contributing to only a small fraction of the pixels in the field of view. *FastICA* can exploit this spatial information simply by reversing the roles of space and time in the original data matrix, leading to spatial ICA and an objective function, $F_{space} = \sum_j Skew[\hat{s}_j]$. Here, $y^{(j)} = U_j$ and the expectation values are averages over all pixels.

Cellular signals in our data sets had typically greater spatial than temporal skewness (Figure 2E). Each Purkinje cell dendritic tree occupied 1,760 ± 589 μm^2 (mean ± s.d.; n=199 dendrites), or ~3% of the area of our typical 60,000 μm^2 field of view. By comparison, the Purkinje cell Ca^{2+} spike rate was 0.76 ± 0.15 Hz and each spike lasted 182 ± 92 ms (n=155 spikes), meaning that in ~14% of time frames each cell was active. Consistent with this disparity, spatial ICA typically achieved higher fidelity signal extractions from our simulated data than temporal ICA (Figure 3F) and also extracted more components from experimental data that resembled cellular signals. Note, however, that sparse activation in time is not the only means of achieving a highly skewed distribution of signal amplitudes. For example, cells with high rates of activity that are strongly modulated in time, such as during bursts, will also exhibit skewed distributions of signal amplitudes.

Spatio-Temporal ICA

To exploit sparseness in both the spatial and temporal domains, we used spatio-temporal ICA [43]. In this case, a set of signals formed by concatenating time courses and spatial filters from each principal component plays the role of $\{y^{(j)}\}$. We used a single parameter, μ, where 1 ≥ μ ≥ 0, to control the relative contribution of spatial and temporal information:

$$y_i^{(k)} = \begin{cases} \mu U_{ki}, & i \le N_x \\ (1-\mu) V_{ki}, & N_x < i \le (N_x + N_t) \end{cases}.$$

Note that $E[y^{(k)}] = \left(\mu \sum_x U_{kx} + (1-\mu) \sum_t V_{kt}\right) \Big/ (N_x + N_t) = 0$. The objective function is therefore $F_\mu[\mathbf{C}] = \sum_j Skew[y^{(j)}] \propto \sum_j \mu^3 \; Skew[\hat{s}_j] + (1-\mu)^3 \; Skew[\hat{a}_j]$, which comprises both the spatial and temporal skewness of the extracted signals.

Spatio-temporal ICA may be viewed as a version of temporal ICA that combines a temporal skewness objective with the *a priori* assumption, weighted by a factor $\mu^3 / (1-\mu)^3$, that the spatial filters should be sparse [315]. The algorithm treats space and time symmetrically, so it can be viewed alternatively as maximizing spatial skewness under the constraint that the signal time courses should also be skewed; in this case, the prior weight is $(1-\mu)^3 / \mu^3$. Nonetheless, there remains a need for further theoretical investigation of the statistical interpretation of spatio-temporal ICA in terms of a generative model with prior assumptions on the filter weights [43, 315, 316].

Temporal Deconvolution and Spike Detection

We estimated the stereotyped waveform of a neuronal Ca^{2+} spike by averaging the time courses of spikes that were simultaneously recorded optically and by extracellular single unit electrical recording in anesthetized mice. We fit the average time course with an exponentially decaying function of time constant τ = 182 ± 97 ms (mean ± s.d., n=155 spikes in 4 experiments). Because this time course was brief, we could improve the detection of Ca^{2+} transients in neuronal signals by high-pass filtering the signals to first remove slow variations due to sources such as photobleaching. We filtered signals by first estimating the slow variations using a moving average with 2 s window size, then subtracting this smooth signal from the original. This approach was not used with glial Ca^{2+} signals, which evolve over longer time scales of several seconds.

For neuronal signals we corrected for the slow decay of fluorescence by using temporal deconvolution [224]. The deconvolved signal was defined by $\hat{s}_{deconv}(t) = \hat{s}(t) / \tau_{deconv} + [\hat{s}(t+dt) - \hat{s}(t)] / dt$, where τ_{deconv} = 150 ms and dt is the time between movie frames. The deconvolved signal for Purkinje cell Ca^{2+} traces contained one large spike corresponding to each action potential, followed by an immediate return to baseline in the next time frame. We defined spike times as local maxima in the deconvolved signal that exceeded a threshold defined as 2 standard deviations above the mean signal. Applying this spike detection procedure in our awake data sets, we found that the average time constant of decay following each detected spike was ~124 ± 63 ms (mean ± s.d., n=150 cells in 5 mice), within the range observed in our anesthetized recordings with paired electrophysiology.

Image Segmentation

Spatio-temporal ICA sometimes identified individual spatial filters that appeared to comprise multiple, spatially separated Purkinje cells. We therefore segmented these filters into distinct sub-components, each representing an individual Purkinje cell. To do this, the original filter was smoothed by a convolution with a Gaussian kernel of 1.5 pixels (2.9 µm) standard deviation. We next transformed the filter into a binary mask by applying a threshold 1.5 standard deviations above the mean intensity of all pixels. We used the MATLAB function `bwlabel`, which gives distinct labels to all spatially connected components within the binary mask [317]. In some cases we identified small image segments that were cut off by the edges of the field of view, or small pieces of dendrites separated from the main processes by a blood vessel. In such cases the signal amplitude relative to background fluctuations was often smaller than for dendritic segments with larger area, due to the smaller number of photons contributing. To avoid excessively noisy signals, we excluded any image segments covering an area smaller than 50 pixels. (For comparison, Purkinje cell dendrites typically occupy 481 ± 161 pixels (mean ± s.d.) in our recordings). Finally, for each image segment we created a spatial filter filled by setting to zero the weights of all pixels outside of the connected region. The filter weights inside the connected region were equal to those of the corresponding pixels in the original filter. The result was a set of spatial filters that had non-overlapping support and retained the fine spatial structure of the original filter.

Artificial Movie Data Sets for Validation of Cell Sorting Procedures

We created artificial movies designed to mimic Ca^{2+}-imaging data acquired in the cerebellar molecular layer. The artificial data were constructed by simulating cellular signals with temporal dynamics and spatial characteristics, such as cell sizes, shape, and density, conforming to our expectations of intracellular Ca^{2+} signals from Purkinje cell dendrites and Bergmann glial processes. Each movie covered a square field of view either 50 × 50 or 64 × 64 pixels in size. The number of Purkinje cells and total rate of Bergmann glial Ca^{2+} transients increased in proportion to the field of view area, so as to maintain a constant density of ~1025 Purkinje cells per mm^2 and 13 glial Ca^{2+} transients per mm^2 per second. This density of Purkinje cells was somewhat higher than the density of cellular signals we extracted from experimental data sets (663 ± 45 per mm^2), reflecting the presence of blood vessels occluding parts of the field of view in the experimental data as well as the possibility that our analysis missed some of the cells that were present.

We distributed the centroids of the Purkinje cell dendritic trees, $\{x_k, y_k\}$, randomly within the middle 64% of the field of view; we chose not to include cells in the periphery of the field of view so that all of the simulated cells would cover roughly the same number of pixels. We modeled Purkinje cell signals by assuming that the pixels near the centroid of the dendrite contribute the most weight to the signal. The relative weight of each pixel in a cell's spatial filter was given by an elongated Gaussian function:

$$U_k(x,y) \propto \exp[-(x-x_k)^2 / 2\sigma_x^2 - (y-y_k)^2 / 2\sigma_y^2]$$

where σ_x = 75 μm, σ_y = 2.5 μm defined the major and minor axis standard deviations. We imposed a sharp spatial cutoff, setting U_k to zero for all pixels whose weight would have been $< 10^{-4}$ times that of the center pixel. These simulated Purkinje cell dendrite shapes were roughly consistent with the total length (90 ± 40 μm) and breadth (7.0 ± 2.3 μm) of Purkinje cell dendrite image segments observed in our experimental data. Bergmann glial spatial filters were defined identically but had isotropic width parameters, $\sigma_x = \sigma_y = 40$ μm. We rotated all spatial filters relative to the square borders of the field of view by 20°. Finally, we normalized each spatial filter so that $\sum_{x,y} U_k(x,y) = 1$.

The time courses of Ca^{2+}-dependent fluorescent signals were modeled by first generating underlying spike trains by using independent but biased random coin flips for each Purkinje cell and for each time frame. The probability of a spike in each time frame was chosen to give a mean spike rate of 0.6 Hz, within the range reported for non-anesthetized mice [23, 221, 290]. We convolved each cell's spike train, represented as a vector of 1's and 0's, with an exponentially decaying function with time constant 240 ms to account for the slow decline of Ca^{2+} fluorescence following a spike.

For simplicity and to avoid multiple simultaneous glial Ca^{2+} transients, which we rarely observed in the experimental data, we chose onset times for glial events, t_i, that were distributed at regularly spaced intervals over the duration of the simulated movies. The time course of each glial Ca^{2+} transient was modeled by the function $(t-t_i)e^{-(t-t_i)/\tau}$, where τ=1.6 s.

Each artificial movie was 100 s in duration and the frame rate was 10 Hz. Cellular signals were added to a static fluorescence background that included bright circular structures resembling the somata of molecular layer interneurons and elongated dark regions resembling blood vessels. These time-dependent and static background signals defined a noiseless data set.

We created noisy movie data with a defined signal to noise ratio, S, by multiplying the noiseless movie data in each pixel, x, and each time frame, t, by an independent random number, α_{xt} chosen from a distribution with mean S^2 and variance S^2. This noise model approximates the statistical distribution of signals we would expect to observe given an average of $\sim S^2$ photons collected per pixel per time frame using a shot-noise limited microscope under conditions in which the contrast is relatively uniform, such as with OGB-1-AM staining. For $S<5$, we used Poisson statistics, so that $\alpha_{x,t}$ is a non-negative integer chosen according to the distribution

$$\Pr[\alpha_{xt}] = e^{-S^2}\frac{(S^2)^{\alpha_{xt}}}{\alpha_{xt}!}.$$

For $S \geq 5$, to improve computational efficiency we approximated Poisson statistics by a Gaussian distribution, $\Pr[\alpha_{xt}] = e^{-(\alpha_{xt}-S^2)^2/(2S^2)} / \sqrt{2\pi S^2}$. Both noise distributions are chosen such that the mean-to-standard deviation ratio is S. To avoid unphysical fluorescence values when using the Gaussian approximation, we set any negative values of α_{xt} to 0.

When analyzing artificial data sets we always chose the number of retained PCs, K, to be equal to the total number of neuronal and glial signals. We set $K_{low}=1, K_{high}=K$. We used μ=0.5 when performing spatio-temporal ICA since this value was in the range for which cell sorting performed optimally for artificial data (Fig. 3). We noticed that the optimal value of μ for each data set depended on factors such as the average spike rate of cells, the imaging frame rate, and the numbers of pixels and time frames. Our experimental data had a larger ratio of pixels to time frames than our artificial data sets, hence our choice of μ=0.1-0.2 for analyzing the experimental data. For computational efficiency we limited the number of iterations of the gradient descent algorithm to 100.

Definitions of Signal Fidelity and Cross Talk

To quantify the performance of automated cell sorting, we compared the signals extracted from our artificial data sets to the underlying cellular signals used to generate the artificial movie. We computed the correlation coefficient between each true signal, $a_k(t)$, and each extracted signal trace, $\hat{a}_j(t)$:

$$cc_{j,k} = \frac{E[\hat{a}_j a_k] - E[\hat{a}_j]E[a_k]}{\left(E[\hat{a}_j^2] - E[\hat{a}_j]^2\right)^{1/2}\left(E[a_k^2] - E[a_k]^2\right)^{1/2}}$$

where E denotes averaging over all time frames. We then paired each extracted signal with one of the true signals in an iterative fashion. First, we identified the pair of signals with the largest correlation coefficient and assigned these two signals to each other. We then removed these two signals from the set of available pairs and repeated the procedure with the remaining true and extracted signals. We proceeded in this way until each extracted signal was assigned to one true signal. We defined the fidelity for each extracted signal as the cross correlation with the true signal with which it was paired in this process. The correlation coefficients between all the other true signals and the extracted signal yielded the cross talk. Our overall measure of cross talk is the median of the K largest such correlation coefficients, where K is the total number of extracted signals.

Comparison with Standard ROI Analysis

We implemented a form of ROI analysis that represents a best-case scenario for the results a human data analyst could achieve with an ROI approach. The underlying assumption is that for each cell the analyst possesses several examples of movie time frames that she knows to contain examples of activations of that cell, caused by spikes in neurons or Ca^{2+} transients in glial processes. In practice, uncertainty and error in the estimation of these exemplary spike times will degrade the performance of an ROI cell sorting procedure compared with the idealized version we studied. Thus, our results likely understate the superiority of ICA over ROI analysis.

We defined ROIs for each Purkinje dendrite by computing the average of the simulated fluorescence in each pixel over all movie time frames coinciding with the occurrences of an action potential in that cell. This spike-triggered average generally yielded a filter with high amplitudes in the pixels contained within the cell, and low amplitudes in pixels reflecting only background fluctuations in fluorescence or signals from other, uncorrelated cells. We used a similar approach to define an ROI for each glial Ca^{2+} transient. Each glial signal contained only one fluorescence transient during the simulated movie, but the transient persisted ~10 times longer than neuronal action potential transients. We therefore defined glial ROIs by averaging fluorescence during a period of 500 ms duration (5 movie time frames) around the peak of the transient.

For both neuronal and glial ROIs, we smoothed the triggered average image by convolving with a Gaussian blurring function of 2 pixels standard deviation. We then applied a threshold, defined as 0.8 times the filter's maximum amplitude, to generate a binary mask. This mask defined the final ROI (Figure 4B). To compute the signal corresponding to each ROI, we averaged the fluorescence in all pixels within the corresponding binary mask for each time frame.

Animals and Animal Surgery

Mouse surgeries were performed under isoflurane anesthesia (1.5–2.5%, mixed with 2.0 L/min O_2). Body temperature was monitored with a rectal thermistor placed underneath the animal's belly and maintained at 36-37°C using a heating blanket. We assessed the depth of anesthesia by monitoring pinch withdrawal, eyelid and corneal reflexes, breathing rate, and vibrissae movements.

Our experiments involved two surgical procedures. Several days prior to recordings the animal's skull was exposed and cleaned above neocortex and cerebellum. A custom metal plate allowing cranial access and repeated head-fixation was attached to the skull with dental acrylic cement (Coltene/Whaledent, H00335). On the recording day, a craniotomy (1-4 mm in diameter) was opened (-6.5 mm posterior to bregma; 0.5 mm lateral). Exposed tissue was irrigated with warm artificial cerebral spinal fluid (ACSF; 125 mM NaCl, 5 mM KCl, 10 mM D-Glucose, 10 mM HEPES, 2 mM $CaCl_2$, 2 mM $MgSO_4$; pH to 7.4 with NaOH). To dampen heartbeat- and breathing-induced brain motion, the craniotomy was filled with agarose (1.5%; Type III-A, high EEO; Sigma) in ACSF and covered with a coverslip immobilized onto the head plate, creating an optical window for imaging. The dura was left intact for Ca^{2+}-indicator injections, but removed immediately prior to agarose application.

Fluorescence Labeling

Targeted bulk-loading of cerebellar tissue with the fluorescent Ca^{2+}-indicator Oregon green 488 1,2-bis(2-aminophenoxy)ethane-N,N,N',N'-tetraacetate-1 acetoxymethyl (OGB-1-AM; Molecular Probes) was performed as described [226] Nimmerjahn et al., 2004; [225]. The final concentration of OGB-1-AM and DMSO in our pipette solution was 500 μM and 2-5%, respectively. To obtain a homogenous, widespread stain of the Purkinje cell and molecular layers we performed several pressure ejections (3-3.5 psi pressure) of 1-2 min duration using a picrospritzer (Picospritzer III, General Valve, Fairfield, NJ). Typically five injections were performed in a grid-like pattern with ~150-250 μm spacing. To counterstain Bergmann glia, the red fluorescent astrocyte marker sulforhodamine 101 (SR101; 200-250 μM in ACSF; Invitrogen) was generally added to the injection solution or applied topically (1-3 min) [240].

Two-Photon Imaging in Behaving Mice

After surgical implantation of a head-plate into the cranium and a recovery period of 1-4 days, mice were habituated in 3-9 sessions (typically 1 session/day) to accept increasingly longer periods (15-90 min sessions; 1.5-7.5 h total habituation time) of head restraint while walking on a 11.9-cm-diameter exercise ball (Habitrail, model 62065) covered with copper mesh. The ball rotated on one axis, allowing the animal to stride forward or backward at liberty.

We used a custom upright two-photon microscope equipped with an ultrashort-pulsed Ti:sapphire laser (Tsunami, Spectra-Physics), two fluorescence detection channels, and a 20× water-immersion objective (Olympus, 0.95 NA, XLUMPlanFl). To excite OGB-1-AM, the laser was tuned to 800 nm. For simultaneous excitation of OGB-1-AM and SR101 we used 800-820 nm. Prior to Ca^{2+}-imaging mice generally recovered from isoflurane for >60 min to minimize aftereffects of anesthesia. To minimize the possibility of photo-induced alterations in tissue physiology, the duration and average power of continuous illumination were <12 min and <35 mW at a given brain location.

We used MP-Scope data-acquisition software for laser-scanning microscopy [318], which uses a 5 MS/s data acquisition card (PCI-6110, National Instruments) for simultaneous acquisition and display of up to two imaging and two analog input channels. Images contained 256×256 pixels (10.2–11.1 Hz frame rate) or 128×128 pixels (20.4 Hz).

Behavioral Assessment

Mouse behavior was recorded in Mini-DV format using a digital Camcorder (Sony DCR-VX2000). Rotational displacements of the exercise ball were monitored using a rotary encoder (US Digital, E7PD-720-118).

Electrophysiology

To record the electrophysiological complex spikes corresponding to optically detected Ca^{2+} spikes in Purkinje cell dendrites [23, 40, 46], we performed targeted extracellular single-unit recordings in the cerebellar cortex of isoflurane anesthetized mice studied simultaneously by two-photon imaging as described above. Ca^{2+}-transients in distal Purkinje cell dendrites were visualized using a 40× 0.8 NA water-immersion objective (Zeiss, Achroplan). Extracellular Purkinje cell activity was recorded using beveled borosilicate glass electrodes (A-M Systems, Inc.; OD, 1.0 mm; ID, 0.58 mm; internal solution, 2 M NaCl and in some cases 200-500 μM sulforhodamine 101; pipette resistance, 5-10 MΩ). Electrodes were positioned under visual control near OGB-1-AM labeled Purkinje cell somata. Purkinje cells were identified by the brief pause in simple spike activity following each complex spike. The raw electrode signal was amplified, filtered (A-M Systems, Model 1800 microelectrode AC amplifier; low cut-off, 300 Hz; high cut-off, 5 kHz), digitized (20 kHz) and stored on disk for off-line analysis. Two-photon imaging data of dendritic Ca^{2+} activity was recorded simultaneously (frame rate, 20.35 Hz) in superficial regions of the molecular layer (depth, 20-50 μm) located directly above the microelectrode tip. The correspondence between dendritic Ca^{2+} spikes and electrophysiological complex spike activity was verified through a spike-triggered average analysis written in Matlab, in agreement with prior combined optical and electrophysiological studies of Purkinje cells [23, 40].

ROC Analysis

Receiver operating characteristic (ROC) analysis characterizes the tradeoff between two competing objectives in the detection of spikes. On the one hand, the threshold for a spiking event should be set low enough so that most of the true spikes are correctly detected (a high hit rate). On the other hand, the threshold should be kept high enough so that very few spikes are reported at times when no spike actually occurred (low rate of false positives). The ROC curve plots the false positive rate versus the hit rate for a range of values of the spike threshold. The area under this curve is a measure of the spike discriminability and can be interpreted as the probability of correctly distinguishing a spike from a non-spike [296].

We wrote custom MATLAB routines to generate ROC curves for paired electrophysiological and two-photon imaging recordings, and to compute the ROC area. After extracting an estimate of each cellular signal by ICA cell sorting, we applied temporal deconvolution and used the resulting deconvolved signal for ROC analysis.

Identification of Periods of Active Locomotion

We used an optical encoder to record the mouse's running speed on the exercise ball. The encoder was sufficiently sensitive to detect even 0.5 deg rotations of the exercise ball during periods when the mouse did not actively walk or run. We first smoothed the encoder signal by convolving with a Gaussian kernel of standard deviation 0.5 s. We defined the onset times of active movement as times at which the smoothed velocity signal surpassed a threshold of 2 mm/s. The offset times of the periods of movement were defined as when the speed fell below 1 mm/s. As a conservative measure, to avoid including any data from actively moving periods in the analysis of resting data, we padded the period of active movement with 2 s before the onset and after the offset of each movement. This procedure allowed us to detect periods of continuous movement, which may include brief pauses lasting less than a few seconds.

Detection of Cerebellar Microzones

We partitioned all identified cells into disjoint microzones using k-means clustering [40, 314], implemented using the MATLAB function *kmeans*. We represented each spike train as a vector, with value 1 for time frames at which a spike was detected and 0 otherwise. The cross-correlation between spike trains provided the distance function that we used in the clustering procedure [46]. We repeated k-means clustering for each data set 20 times with random initial cluster centroids, and chose the solution with the largest average cross-correlation between each spike train and the centroid of its cluster. To assess the stability of microzones over time, we divided each recording lasting 7-10 minutes into 1-minute duration data subsets and repeated the clustering procedure for each subset. We also performed clustering separately on data recorded during rest and active locomotion.

References

1. Abbott, L.F., *Theoretical Neuroscience Rising.* 2008. **60**(3): p. 489-495.
2. Dickinson, E., *The complete poems of Emily Dickinson*. 1924, Boston,: Little, Brown, and Company.
3. Ramón y Cajal, S., *Histologie du système nerveux de l'homme & des vertébrés*. Éd. française rev. & mise à jour ed. 1911, Paris: Maloine.
4. Livet, J., T.A. Weissman, H. Kang, R.W. Draft, J. Lu, R.A. Bennis, J.R. Sanes, and J.W. Lichtman, *Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system.* Nature, 2007. **450**(7166): p. 56-62.
5. Eccles, J.C., M. Ito, and J. Szentágothai, *The cerebellum as a neuronal machine*. 1967, Berlin, New York [etc.]: Springer-Verlag.
6. Braitenberg, V., *Is the cerebellar cortex a biological clock in the millisecond range?* Prog Brain Res, 1967. **25**: p. 334-46.
7. Braitenberg, V., *In defense of the cerebellum.* Ann N Y Acad Sci, 2002. **978**: p. 175-83.
8. Bell, C.C., V. Han, and N.B. Sawtell, *Cerebellum-like structures and their implications for cerebellar function.* Annu Rev Neurosci, 2008. **31**: p. 1-24.
9. McCormick, D.A. and R.F. Thompson, *Cerebellum: essential involvement in the classically conditioned eyelid response.* Science, 1984. **223**(4633): p. 296-9.
10. Ito, M., *The cerebellum and neural control*. 1984.
11. Marr, D., *A theory of cerebellar cortex.* J Physiol, 1969. **202**(2): p. 437-70.
12. Albus, J., *A theory of cerebellar function.* Math. Biosci., 1971. **10**: p. 25-61.
13. Kehoe, E.J. and R.M. Napier, *Temporal specificity in cross-modal transfer of the rabbit nictitating membrane response.* J Exp Psychol Anim Behav Process, 1991. **17**(1): p. 26-35.
14. Maass, W., T. Natschlager, and H. Markram, *Real-time computing without stable states: a new framework for neural computation based on perturbations.* Neural Comput, 2002. **14**(11): p. 2531-60.
15. White, O.L., D.D. Lee, and H. Sompolinsky, *Short-term memory in orthogonal neural networks.* Phys Rev Lett, 2004. **92**(14): p. 148102.
16. Buonomano, D.D., *Neural network model of the cerebellum: temporal discrimination and the timing of motor responses.* Neural computation, 1994. **6**: p. 38.
17. Medina, J.F., K.S. Garcia, W.L. Nores, N.M. Taylor, and M.D. Mauk, *Timing mechanisms in the cerebellum: testing predictions of a large-scale computer simulation.* J Neurosci, 2000. **20**(14): p. 5516-25.
18. Sumbre, G., A. Muto, H. Baier, and M.M. Poo, *Entrained rhythmic activities of neuronal ensembles as perceptual memory of time interval.* Nature, 2008.

19. Karmarkar, U.R. and D.V. Buonomano, *Timing in the absence of clocks: encoding time in neural network states.* Neuron, 2007. **53**(3): p. 427-38.

20. Gormezano, I., N. Schneiderman, E. Deaux, and I. Fuentes, *Nictitating membrane: classical conditioning and extinction in the albino rabbit.* Science, 1962. **138**: p. 33-4.

21. Ohyama, T., W.L. Nores, M. Murphy, and M.D. Mauk, *What the cerebellum computes.* Trends Neurosci, 2003. **26**(4): p. 222-7.

22. Wetmore, D.Z., E.A. Mukamel, and M.J. Schnitzer, *Lock-and-Key Mechanisms of Cerebellar Memory Recall Based on Rebound Currents.* J Neurophysiol, 2008. **100**(4): p. 2328-47.

23. Flusberg, B.A., A. Nimmerjahn, E.D. Cocker, E.A. Mukamel, R.P. Barretto, T.H. Ko, L.D. Burns, J.C. Jung, and M.J. Schnitzer, *High-speed, miniaturized fluorescence microscopy in freely moving mice.* Nat Methods, 2008.

24. Pitkow, X., H. Sompolinsky, and M. Meister, *A neural computation for visual acuity in the presence of eye movements.* PLoS Biology, 2007. **5**(12): p. e331.

25. Andersson, G. and D.M. Armstrong, *Complex spikes in Purkinje cells in the lateral vermis (b zone) of the cat cerebellum during locomotion.* J Physiol, 1987. **385**: p. 107-34.

26. Dahmen, H.J., *A simple apparatus to investigate the orientation of walking insects.* Experientia, 1980. **36**: p. 685.

27. Borst, A. and M. Heisenberg, *Osmotropotaxis in Drosophila melanogaster.* Journal of Comparative Physiology A, 2004. **147**: p. 479-484.

28. Dombeck, D.A., A.N. Khabbaz, F. Collman, T.L. Adelman, and D.W. Tank, *Imaging large-scale neural activity with cellular resolution in awake, mobile mice.* Neuron, 2007. **56**(1): p. 43-57.

29. Nimmerjahn, A., E. Mukamel, and M. Schnitzer, *Two-photon imaging of neuronal and glial network activity in the cerebellar cortex of awake mice.* Soc. for Neurosci., 2007. **252.5/K21**.

30. Cornell-Bell, A.H., S.M. Finkbeiner, M.S. Cooper, and S.J. Smith, *Glutamate induces calcium waves in cultured astrocytes: long-range glial signaling.* Science, 1990. **247**(4941): p. 470-3.

31. Hirase, H., L.F. Qian, P. Bartho, and G. Buzsaki, *Calcium dynamics of cortical astrocytic networks in vivo.* Plos Biology, 2004. **2**(4): p. 494-499.

32. Nimmerjahn, A., F. Kirchhoff, J.N. Kerr, and F. Helmchen, *Sulforhodamine 101 as a specific marker of astroglia in the neocortex in vivo.* Nat Methods, 2004. **1**(1): p. 31-7.

33. Wang, X.H., N.H. Lou, Q.W. Xu, G.F. Tian, W.G. Peng, X.N. Han, J. Kang, T. Takano, and M. Nedergaard, *Astrocytic Ca2+ signaling evoked by sensory stimulation in vivo.* Nature Neuroscience, 2006. **9**(6): p. 816-823.

34. Schummers, J., H.B. Yu, and M. Sur, *Tuned responses of astrocytes and their influence on hemodynamic signals in the visual cortex.* Science, 2008. **320**(5883): p. 1638-1643.

35. Press, O.U., *The Oxford English dictionary.* 2nd ed. 1989, Oxford

Oxford ; New York: Clarendon Press ;

Oxford University Press.

36. Iadecola, C. and M. Nedergaard, *Glial regulation of the cerebral microvasculature.* Nature Neuroscience, 2007. **10**: p. 1369-1376.

37. Wilson, M.A. and B.L. McNaughton, *Dynamics of the hippocampal ensemble code for space.* Science, 1993. **261**(5124): p. 1055-8.

38. Lewicki, M.S., *A review of methods for spike sorting: the detection and classification of neural action potentials.* Network, 1998. **9**(4): p. R53-78.

39. Greenberg, D.S., A.R. Houweling, and J.N.D. Kerr, *Population imaging of ongoing neuronal activity in the visual cortex of awake rats.* Nature Neuroscience, 2008. **11**(7): p. 749-751.

40. Ozden, I., H.M. Lee, M.R. Sullivan, and S.S. Wang, *Identification and clustering of event patterns from in vivo multiphoton optical recordings of neuronal ensembles.* J Neurophysiol, 2008. **100**(1): p. 495-503.

41. Bell, A.J. and T.J. Sejnowski, *An information-maximization approach to blind separation and blind deconvolution.* Neural Comput, 1995. **7**(6): p. 1129-59.

42. Brown, G.D., S. Yamada, and T.J. Sejnowski, *Independent component analysis at the neural cocktail party.* Trends Neurosci, 2001. **24**(1): p. 54-63.

43. Stone, J.V., J. Porrill, N.R. Porter, and I.D. Wilkinson, *Spatiotemporal independent component analysis of event-related fMRI data using skewed probability density functions.* Neuroimage, 2002. **15**(2): p. 407-21.

44. Andersson, G. and O. Oscarsson, *Climbing fiber microzones in cerebellar vermis and their projection to different groups of cells in the lateral vestibular nucleus.* Exp Brain Res, 1978. **32**(4): p. 565-79.

45. Jorntell, H., C. Ekerot, M. Garwicz, and X.L. Luo, *Functional organization of climbing fibre projection to the cerebellar anterior lobe of the rat.* The Journal of Physiology, 2000. **522 Pt 2**: p. 297-309.

46. Welsh, J.P., E.J. Lang, I. Suglhara, and R. Llinas, *Dynamic organization of motor control within the olivocerebellar system.* Nature, 1995. **374**(6521): p. 453-7.

47. Apps, R. and M. Garwicz, *Anatomical and physiological foundations of cerebellar information processing.* Nat Rev Neurosci, 2005. **6**(4): p. 297-311.

48. Greenberg, D.S., A.R. Houweling, and J.N. Kerr, *Population imaging of ongoing neuronal activity in the visual cortex of awake rats.* Nat Neurosci, 2008. **11**(7): p. 749-751.

49. Eagleman, D.M., P.U. Tse, D. Buonomano, P. Janssen, A.C. Nobre, and A.O. Holcombe, *Time and the brain: how subjective time relates to neural time.* J Neurosci, 2005. **25**(45): p. 10369-71.

50. Ivry, R.B. and R.M. Spencer, *The neural representation of time.* Curr Opin Neurobiol, 2004. **14**(2): p. 225-32.

51. Mauk, M.D. and D.V. Buonomano, *The neural basis of temporal processing.* Annu Rev Neurosci, 2004. **27**: p. 307-40.

52. Bullock, D., J.C. Fiala, and S. Grossberg, *A neural model of timed response learning in the cerebellum.* Neural Networks, 1994.
53. Buonomano, D.V., *The biology of time across different scales.* Nature chemical biology, 2007. **3**(10): p. 594-7.
54. Desmond, J.E. and J.W. Moore, *Adaptive timing in neural networks: the conditioned response.* Biological cybernetics, 1988. **58**(6): p. 405-15.
55. Hong, S., L. Optican, and K. Friston, *Interaction between Purkinje Cells and Inhibitory Interneurons May Create Adjustable Output Waveforms to Generate Timed Cerebellar Output.* PLoS ONE, 2008. **3**(7): p. e2770.
56. Yamazaki, T. and S. Tanaka, *Neural modeling of an internal clock.* Neural computation, 2005. **17**(5): p. 1032-58.
57. Mauk, M.D. and B.P. Ruiz, *Learning-dependent timing of Pavlovian eyelid responses: differential conditioning using multiple interstimulus intervals.* Behavioral neuroscience, 1992. **106**(4): p. 666-81.
58. Schneiderman, N., I. Fuentes, and I. Gormezano, *Acquisition and extinction of the classically conditioned eyelid response in the albino rabbit.* Science, 1962. **136**: p. 650-2.
59. Thompson, R.F., *Neural mechanisms of classical conditioning in mammals.* Philosophical transactions of the Royal Society of London Series B, Biological sciences, 1990. **329**(1253): p. 161-70.
60. Hesslow, G., P. Svensson, and M. Ivarsson, *Learned movements elicited by direct stimulation of cerebellar mossy fiber afferents.* Neuron, 1999. **24**(1): p. 179-85.
61. Aitkin, L.M. and J. Boyd, *Acoustic input to the lateral pontine nuclei.* Hearing research, 1978. **1**(1): p. 67-77.
62. Berthier, N.E. and J.W. Moore, *Cerebellar Purkinje cell activity related to the classically conditioned nictitating membrane response.* Experimental brain research, 1986. **63**(2): p. 341-50.
63. Walter, J. and K. Khodakhah, *The linear computational algorithm of cerebellar Purkinje cells.* J Neurosci, 2006. **26**(50): p. 12861-72.
64. Brunel, N., V. Hakim, P. Isope, J.P. Nadal, and B. Barbour, *Optimal information storage and the distribution of synaptic weights: perceptron versus Purkinje cell.* Neuron, 2004. **43**(5): p. 745-57.
65. Minsky, M.L. and S. Papert, *Perceptrons : an introduction to computational geometry.* 1988: p. xv, 292 p.
66. Rosenblatt, F., *Principles of neurodynamics; perceptrons and the theory of brain mechanisms.* 1962: p. 616 p.
67. Fiete, I.R., R.H. Hahnloser, M.S. Fee, and H.S. Seung, *Temporal sparseness of the premotor drive is important for rapid learning in a neural network model of birdsong.* J Neurophysiol, 2004. **92**(4): p. 2274-82.

68. Rumelhart, D., G.E. Hinton, and R. Williams, *Learning representations by back-propagating errors.* Nature, 2004. **323**: p. 533.

69. Lev-Ram, V., S.B. Mehta, D. Kleinfeld, and R.Y. Tsien, *Reversing cerebellar long-term depression.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(26): p. 15989-93.

70. Chadderton, P., T.W. Margrie, and M. Hausser, *Integration of quanta in cerebellar granule cells during sensory processing.* Nature, 2004. **428**(6985): p. 856-60.

71. Hastie, T., R. Tibshirani, and J.H. Friedman, *The elements of statistical learning : data mining, inference, and prediction*. 2001, New York: Springer.

72. Seung, H.S., *How the brain keeps the eyes still.* Proceedings of the National Academy of Sciences of the United States of America, 1996. **93**(23): p. 13339-44.

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

74. Hastie, T., R. Tibshirani, and J. H. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* 2001: p. 533.

75. Buonomano, D.V. and M.D. Mauk, *Neural network model of the cerebellum: temporal discrimination and the timing of motor responses.* Neural Comput, 1994. **6**: p. 38-55.

76. Ermentrout, B., *Neural networks as spatio-temporal pattern-forming systems.* Reports on Progress in Physics, 1998. **61**(4): p. 353-430.

77. Rajan, K. and L.F. Abbott, *Eigenvalue spectra of random matrices for neural networks.* Physical Review Letters, 2006. **97**(18): p. 188104.

78. Sommers, H.J., A. Crisanti, H. Sompolinsky, and Y. Stein, *Spectrum of large random asymmetric matrices.* Physical Review Letters, 1988. **60**(19): p. 1895-1898.

79. Sompolinsky, H., A. Crisanti, and H.J. Sommers, *Chaos in random neural networks.* Physical Review Letters, 1988. **61**(3): p. 259-262.

80. Debiec, J., V. Doyere, K. Nader, and J.E. Ledoux, *Directly reactivated, but not indirectly reactivated, memories undergo reconsolidation in the amygdala.* Proc Natl Acad Sci U S A, 2006. **103**(9): p. 3428-33.

81. Doyere, V., J. Debiec, M.H. Monfils, G.E. Schafe, and J.E. Ledoux, *Synapse-specific reconsolidation of distinct fear memories in the lateral amygdala.* Nat Neurosci, 2007.

82. Wills, T.J., C. Lever, F. Cacucci, N. Burgess, and J. O'Keefe, *Attractor dynamics in the hippocampal representation of the local environment.* Science, 2005. **308**(5723): p. 873-6.

83. Hopfield, J.J., *Neural networks and physical systems with emergent collective computational abilities.* Proc Natl Acad Sci U S A, 1982. **79**(8): p. 2554-8.

84. Medina, J.F. and M.D. Mauk, *Computer simulation of cerebellar information processing.* Nat Neurosci, 2000. **3 Suppl**: p. 1205-11.

85. Mauk, M.D. and N.H. Donegan, *A model of Pavlovian eyelid conditioning based on the synaptic organization of the cerebellum.* Learn Mem, 1997. **4**(1): p. 130-58.

86. du Lac, S., J.L. Raymond, T.J. Sejnowski, and S.G. Lisberger, *Learning and memory in the vestibulo-ocular reflex.* Annu Rev Neurosci, 1995. **18**: p. 409-41.

87. Nakazawa, K., M.C. Quirk, R.A. Chitwood, M. Watanabe, M.F. Yeckel, L.D. Sun, A. Kato, C.A. Carr, D. Johnston, M.A. Wilson, and S. Tonegawa, *Requirement for hippocampal CA3 NMDA receptors in associative memory recall.* Science, 2002. **297**(5579): p. 211-8.

88. De Zeeuw, C.I. and C.H. Yeo, *Time and tide in cerebellar memory formation.* Curr Opin Neurobiol, 2005. **15**(6): p. 667-74.

89. Lang, E.J., I. Sugihara, J.P. Welsh, and R. Llinas, *Patterns of spontaneous purkinje cell complex spike activity in the awake rat.* J Neurosci, 1999. **19**(7): p. 2728-39.

90. Boyden, E.S., A. Katoh, and J.L. Raymond, *Cerebellum-dependent learning: the role of multiple plasticity mechanisms.* Annu Rev Neurosci, 2004. **27**: p. 581-609.

91. Hansel, C., D.J. Linden, and E. D'Angelo, *Beyond parallel fiber LTD: the diversity of synaptic and non-synaptic plasticity in the cerebellum.* Nat Neurosci, 2001. **4**(5): p. 467-75.

92. Ohyama, T. and M. Mauk, *Latent acquisition of timed responses in cerebellar cortex.* J Neurosci, 2001. **21**(2): p. 682-90.

93. Ohyama, T., W.L. Nores, and M.D. Mauk, *Stimulus generalization of conditioned eyelid responses produced without cerebellar cortex: implications for plasticity in the cerebellar nuclei.* Learn Mem, 2003. **10**(5): p. 346-54.

94. Mauk, M.D., *Roles of cerebellar cortex and nuclei in motor learning: contradictions or clues?* Neuron, 1997. **18**(3): p. 343-6.

95. Miles, F.A. and S.G. Lisberger, *Plasticity in the vestibulo-ocular reflex: a new hypothesis.* Annu Rev Neurosci, 1981. **4**: p. 273-99.

96. Napper, R.M. and R.J. Harvey, *Number of parallel fiber synapses on an individual Purkinje cell in the cerebellum of the rat.* J Comp Neurol, 1988. **274**(2): p. 168-77.

97. Mauk, M.D. and T. Ohyama, *Extinction as new learning versus unlearning: considerations from a computer simulation of the cerebellum.* Learn Mem, 2004. **11**(5): p. 566-71.

98. Kimpo, R.R., E.S. Boyden, A. Katoh, M.C. Ke, and J.L. Raymond, *Distinct patterns of stimulus generalization of increases and decreases in VOR gain.* J Neurophysiol, 2005. **94**(5): p. 3092-100.

99. Barnet, R.C., R.P. Cole, and R.R. Miller, *Temporal integration in second-order conditioning and sensory preconditioning.* Animal Learning and Behavior, 1997. **25**(2): p. 221-233.

100. Christian, K.M. and R.F. Thompson, *Neural substrates of eyeblink conditioning: acquisition and retention.* Learn Mem, 2003. **10**(6): p. 427-55.

101. Ito, M., *Cerebellar control of the vestibulo-ocular reflex--around the flocculus hypothesis.* Annu Rev Neurosci, 1982. **5**: p. 275-96.

102. Raymond, J.L. and S.G. Lisberger, *Behavioral analysis of signals that guide learned changes in the amplitude and dynamics of the vestibulo-ocular reflex.* J Neurosci, 1996. **16**(23): p. 7791-802.

103. Jahnsen, H., *Electrophysiological characteristics of neurones in the guinea-pig deep cerebellar nuclei in vitro.* J Physiol, 1986. **372**: p. 129-47.

104. Aizenman, C.D., P.B. Manis, and D.J. Linden, *Polarity of long-term synaptic gain change is related to postsynaptic spike firing at a cerebellar inhibitory synapse.* Neuron, 1998. **21**(4): p. 827-35.

105. Llinas, R. and M. Muhlethaler, *Electrophysiology of guinea-pig cerebellar nuclear cells in the in vitro brain stem-cerebellar preparation.* J Physiol, 1988. **404**: p. 241-58.

106. Sekirnjak, C. and S. du Lac, *Intrinsic firing dynamics of vestibular nucleus neurons.* J Neurosci, 2002. **22**(6): p. 2083-95.

107. Aizenman, C.D. and D.J. Linden, *Regulation of the rebound depolarization and spontaneous firing patterns of deep nuclear neurons in slices of rat cerebellum.* J Neurophysiol, 1999. **82**(4): p. 1697-709.

108. Smith, M.C., S.R. Coleman, and I. Gormezano, *Classical conditioning of the rabbit's nictitating membrane response at backward, simultaneous, and forward CS-US intervals.* J Comp Physiol Psychol, 1969. **69**(2): p. 226-31.

109. Salafia, W.R., R.W. Lambert, K.C. Host, N.L. Chiala, and J.J. Ramirez, *Rabbit nictitating membrane conditioning: lower limit of the effective interstimulus interval.* Animal Learning & Behavior, 1980. **8**(1): p. 85-91.

110. Smith, M.C., *CS-US interval and US intensity in classical conditioning of the rabbit's nictitating membrane response.* J Comp Physiol Psychol, 1968. **66**(3): p. 679-87.

111. Schneiderman, N. and I. Gormezano, *Conditioning of the Nictitating Membrane of the Rabbit as a Function of Cs-Us Interval.* J Comp Physiol Psychol, 1964. **57**: p. 188-95.

112. Ito, M. and M. Kano, *Long-lasting depression of parallel fiber-Purkinje cell transmission induced by conjunctive stimulation of parallel fibers and climbing fibers in the cerebellar cortex.* Neurosci Lett, 1982. **33**(3): p. 253-8.

113. Ito, M., *Long-term depression.* Annu Rev Neurosci, 1989. **12**: p. 85-102.

114. Mauk, M.D., J.E. Steinmetz, and R.F. Thompson, *Classical conditioning using stimulation of the inferior olive as the unconditioned stimulus.* Proc Natl Acad Sci U S A, 1986. **83**(14): p. 5349-53.

115. McCormick, D.A., J.E. Steinmetz, and R.F. Thompson, *Lesions of the inferior olivary complex cause extinction of the classically conditioned eyeblink response.* Brain Res, 1985. **359**(1-2): p. 120-30.

116. Steinmetz, J.E., D.G. Lavond, and R.F. Thompson, *Classical conditioning in rabbits using pontine nucleus stimulation as a conditioned stimulus and inferior olive stimulation as an unconditioned stimulus.* Synapse, 1989. **3**(3): p. 225-33.

117. Steinmetz, J.E., D.J. Rosen, P.F. Chapman, D.G. Lavond, and R.F. Thompson, *Classical conditioning of the rabbit eyelid response with a mossy-fiber stimulation CS: I.*

Pontine nuclei and middle cerebellar peduncle stimulation. Behav Neurosci, 1986. **100**(6): p. 878-87.

118. Feil, R., J. Hartmann, C. Luo, W. Wolfsgruber, K. Schilling, S. Feil, J.J. Barski, M. Meyer, A. Konnerth, C.I. De Zeeuw, and F. Hofmann, *Impairment of LTD and cerebellar learning by Purkinje cell-specific ablation of cGMP-dependent protein kinase I.* J Cell Biol, 2003. **163**(2): p. 295-302.

119. Kishimoto, Y., S. Kawahara, M. Suzuki, H. Mori, M. Mishina, and Y. Kirino, *Classical eyeblink conditioning in glutamate receptor subunit delta 2 mutant mice is impaired in the delay paradigm but not in the trace paradigm.* Eur J Neurosci, 2001. **13**(6): p. 1249-53.

120. Koekkoek, S.K., H.C. Hulscher, B.R. Dortland, R.A. Hensbroek, Y. Elgersma, T.J. Ruigrok, and C.I. De Zeeuw, *Cerebellar LTD and learning-dependent timing of conditioned eyelid responses.* Science, 2003. **301**(5640): p. 1736-9.

121. Miyata, M., H.T. Kim, K. Hashimoto, T.K. Lee, S.Y. Cho, H. Jiang, Y. Wu, K. Jun, D. Wu, M. Kano, and H.S. Shin, *Deficient long-term synaptic depression in the rostral cerebellum correlated with impaired motor learning in phospholipase C beta4 mutant mice.* Eur J Neurosci, 2001. **13**(10): p. 1945-54.

122. Shibuki, K., H. Gomi, L. Chen, S. Bao, J.J. Kim, H. Wakatsuki, T. Fujisaki, K. Fujimoto, A. Katoh, T. Ikeda, C. Chen, R.F. Thompson, and S. Itohara, *Deficient cerebellar long-term depression, impaired eyeblink conditioning, and normal motor coordination in GFAP mutant mice.* Neuron, 1996. **16**(3): p. 587-99.

123. Koekkoek, S.K., K. Yamaguchi, B.A. Milojkovic, B.R. Dortland, T.J. Ruigrok, R. Maex, W. De Graaf, A.E. Smit, F. VanderWerf, C.E. Bakker, R. Willemsen, T. Ikeda, S. Kakizawa, K. Onodera, D.L. Nelson, E. Mientjes, M. Joosten, E. De Schutter, B.A. Oostra, M. Ito, and C.I. De Zeeuw, *Deletion of FMR1 in Purkinje cells enhances parallel fiber LTD, enlarges spines, and attenuates cerebellar eyelid conditioning in Fragile X syndrome.* Neuron, 2005. **47**(3): p. 339-52.

124. Medina, J.F. and M.D. Mauk, *Simulations of cerebellar motor learning: computational analysis of plasticity at the mossy fiber to deep nucleus synapse.* J Neurosci, 1999. **19**(16): p. 7140-51.

125. Boyden, E.S., A. Katoh, J.L. Pyle, T.A. Chatila, R.W. Tsien, and J.L. Raymond, *Selective engagement of plasticity mechanisms for motor memory storage.* Neuron, 2006. **51**(6): p. 823-34.

126. Boyden, E.S. and J.L. Raymond, *Active reversal of motor memories reveals rules governing memory encoding.* Neuron, 2003. **39**(6): p. 1031-42.

127. Jahnsen, H., *Extracellular activation and membrane conductances of neurones in the guinea-pig deep cerebellar nuclei in vitro.* J Physiol, 1986. **372**: p. 149-68.

128. Abbott, L.F. and S.B. Nelson, *Synaptic plasticity: taming the beast.* Nat Neurosci, 2000. **3 Suppl**: p. 1178-83.

129. Coesmans, M., J.T. Weber, C.I. De Zeeuw, and C. Hansel, *Bidirectional parallel fiber plasticity in the cerebellum under climbing fiber control.* Neuron, 2004. **44**(4): p. 691-700.

130. Wang, S.S., W. Denk, and M. Hausser, *Coincidence detection in single dendritic spines mediated by calcium release.* Nat Neurosci, 2000. **3**(12): p. 1266-73.

131. Doi, T., S. Kuroda, T. Michikawa, and M. Kawato, *Inositol 1,4,5-trisphosphate-dependent Ca2+ threshold dynamics detect spike timing in cerebellar Purkinje cells.* J Neurosci, 2005. **25**(4): p. 950-61.

132. Lev-Ram, V., S.B. Mehta, D. Kleinfeld, and R.Y. Tsien, *Reversing cerebellar long-term depression.* Proc Natl Acad Sci U S A, 2003. **100**(26): p. 15989-93.

133. Hines, M.L. and N.T. Carnevale, *The NEURON simulation environment.* Neural Comput, 1997. **9**(6): p. 1179-209.

134. Berthier, N.E. and J.W. Moore, *Cerebellar Purkinje cell activity related to the classically conditioned nictitating membrane response.* Exp Brain Res, 1986. **63**(2): p. 341-50.

135. King, D.A., D.J. Krupa, M.R. Foy, and R.F. Thompson, *Mechanisms of neuronal conditioning.* Int Rev Neurobiol, 2001. **45**: p. 313-37.

136. Freeman, J.H., Jr. and D.A. Nicholson, *Neuronal activity in the cerebellar interpositus and lateral pontine nuclei during inhibitory classical conditioning of the eyeblink response.* Brain Res, 1999. **833**(2): p. 225-33.

137. Nicholson, D.A. and J.H. Freeman, Jr., *Neuronal correlates of conditioned inhibition of the eyeblink response in the anterior interpositus nucleus.* Behav Neurosci, 2002. **116**(1): p. 22-36.

138. Raman, I.M., A.E. Gustafson, and D. Padgett, *Ionic currents and spontaneous firing in neurons isolated from the cerebellar nuclei.* J Neurosci, 2000. **20**(24): p. 9004-16.

139. Anchisi, D., B. Scelfo, and F. Tempia, *Postsynaptic currents in deep cerebellar nuclei.* J Neurophysiol, 2001. **85**(1): p. 323-31.

140. Hille, B., *Ion channels of excitable membranes*. 3rd ed. 2001, Sunderland, Mass.: Sinauer.

141. McRory, J.E., C.M. Santi, K.S. Hamming, J. Mezeyova, K.G. Sutton, D.L. Baillie, A. Stea, and T.P. Snutch, *Molecular and functional characterization of a family of rat brain T-type calcium channels.* J Biol Chem, 2001. **276**(6): p. 3999-4011.

142. Talley, E.M., L.L. Cribbs, J.H. Lee, A. Daud, E. Perez-Reyes, and D.A. Bayliss, *Differential distribution of three members of a gene family encoding low voltage-activated (T-type) calcium channels.* J Neurosci, 1999. **19**(6): p. 1895-911.

143. Gauck, V., M. Thomann, D. Jaeger, and A. Borst, *Spatial distribution of low- and high-voltage-activated calcium currents in neurons of the deep cerebellar nuclei.* J Neurosci, 2001. **21**(15): p. RC158.

144. Mainen, Z.F. and T.J. Sejnowski, *Influence of dendritic structure on firing pattern in model neocortical neurons.* Nature, 1996. **382**(6589): p. 363-6.

145. Chan-Palay, V., *Axon terminals of the intrinsic neurons in the nucleus lateralis of the cerebellum. An electron microscope study.* Z Anat Entwicklungsgesch, 1973. **142**(2): p. 187-206.

146. Pinsky, P.F. and J. Rinzel, *Intrinsic and network rhythmogenesis in a reduced Traub model for CA3 neurons.* J Comput Neurosci, 1994. **1**(1-2): p. 39-60.

147. Aksenov, D.P., N.A. Serdyukova, J.R. Bloedel, and V. Bracha, *Glutamate neurotransmission in the cerebellar interposed nuclei: involvement in classically conditioned eyeblinks and neuronal activity.* J Neurophysiol, 2005. **93**(1): p. 44-52.

148. Pape, H.C., *Queer current and pacemaker: the hyperpolarization-activated cation current in neurons.* Annu Rev Physiol, 1996. **58**: p. 299-327.

149. Santoro, B., S. Chen, A. Luthi, P. Pavlidis, G.P. Shumyatsky, G.R. Tibbs, and S.A. Siegelbaum, *Molecular and functional heterogeneity of hyperpolarization-activated pacemaker channels in the mouse CNS.* J Neurosci, 2000. **20**(14): p. 5264-75.

150. Spain, W.J., P.C. Schwindt, and W.E. Crill, *Anomalous rectification in neurons from cat sensorimotor cortex in vitro.* J Neurophysiol, 1987. **57**(5): p. 1555-76.

151. Straka, H., N. Vibert, P.P. Vidal, L.E. Moore, and M.B. Dutia, *Intrinsic membrane properties of vertebrate vestibular neurons: function, development and plasticity.* Prog Neurobiol, 2005. **76**(6): p. 349-92.

152. Chun, S.W., J.H. Choi, M.S. Kim, and B.R. Park, *Characterization of spontaneous synaptic transmission in rat medial vestibular nucleus.* Neuroreport, 2003. **14**(11): p. 1485-8.

153. du Lac, S. and S.G. Lisberger, *Cellular processing of temporal information in medial vestibular nucleus neurons.* J Neurosci, 1995. **15**(12): p. 8000-10.

154. du Lac, S. and S.G. Lisberger, *Membrane and firing properties of avian medial vestibular nucleus neurons in vitro.* J Comp Physiol [A], 1995. **176**(5): p. 641-51.

155. Perrett, S.P., B.P. Ruiz, and M.D. Mauk, *Cerebellar cortex lesions disrupt learning-dependent timing of conditioned eyelid responses.* J Neurosci, 1993. **13**(4): p. 1708-18.

156. Fitzhugh, R., *Thresholds and plateaus in the Hodgkin-Huxley nerve equations.* J Gen Physiol, 1960. **43**: p. 867-96.

157. Loewenstein, Y., S. Mahon, P. Chadderton, K. Kitamura, H. Sompolinsky, Y. Yarom, and M. Hausser, *Bistability of cerebellar Purkinje cells modulated by sensory stimulation.* Nat Neurosci, 2005. **8**(2): p. 202-11.

158. Rinzel, J. and B. Ermentrout, *Analysis of neural excitability and oscillations*, in *Methods in Neuronal Modeling*, C. Koch and I. Segev, Editors. 2001, MIT Press: Cambridge. p. 671.

159. Rush, M.E. and J. Rinzel, *Analysis of bursting in a thalamic neuron model.* Biol Cybern, 1994. **71**(4): p. 281-91.

160. Sekirnjak, C., B. Vissel, J. Bollinger, M. Faulstich, and S. du Lac, *Purkinje cell synapses target physiologically unique brainstem neurons.* J Neurosci, 2003. **23**(15): p. 6392-8.

161. Serafin, M., C. de Waele, A. Khateb, P.P. Vidal, and M. Muhlethaler, *Medial vestibular nucleus in the guinea-pig. II. Ionic basis of the intrinsic membrane properties in brainstem slices.* Exp Brain Res, 1991. **84**(2): p. 426-33.

162. Smith, M.R., A.B. Nelson, and S. Du Lac, *Regulation of firing response gain by calcium-dependent mechanisms in vestibular nucleus neurons.* J Neurophysiol, 2002. **87**(4): p. 2031-42.

163. Gould, T.J. and J.E. Steinmetz, *Changes in rabbit cerebellar cortical and interpositus nucleus activity during acquisition, extinction, and backward classical eyelid conditioning.* Neurobiol Learn Mem, 1996. **65**(1): p. 17-34.

164. Sekirnjak, C. and S. du Lac, *Physiological and anatomical properties of mouse medial vestibular nucleus neurons projecting to the oculomotor nucleus.* J Neurophysiol, 2006.

165. Hesslow, G., *Correspondence between climbing fibre input and motor output in eyeblink-related areas in cat cerebellar cortex.* J Physiol, 1994. **476**(2): p. 229-44.

166. Chen, F.P. and C. Evinger, *Cerebellar modulation of trigeminal reflex blinks: interpositus neurons.* J Neurosci, 2006. **26**(41): p. 10569-76.

167. Pedroarena, C.M. and C. Schwarz, *Efficacy and short-term plasticity at GABAergic synapses between Purkinje and cerebellar nuclei neurons.* J Neurophysiol, 2003. **89**(2): p. 704-15.

168. Heck, D., F. Kummell, W.T. Thach, and A. Aertsena, *Dynamic correlation of neuronal activity in rat cerebellar cortex modulated by behavior.* Ann N Y Acad Sci, 2002. **978**: p. 156-63.

169. Thier, P., P.W. Dicke, R. Haas, and S. Barash, *Encoding of movement time by populations of cerebellar Purkinje cells.* Nature, 2000. **405**(6782): p. 72-6.

170. Berthier, N.E. and J.W. Moore, *Activity of deep cerebellar nuclear cells during classical conditioning of nictitating membrane extension in rabbits.* Exp Brain Res, 1990. **83**(1): p. 44-54.

171. Choi, J.S. and J.W. Moore, *Cerebellar neuronal activity expresses the complex topography of conditioned eyeblink responses.* Behav Neurosci, 2003. **117**(6): p. 1211-9.

172. McCormick, D.A., G.A. Clark, D.G. Lavond, and R.F. Thompson, *Initial localization of the memory trace for a basic form of learning.* Proc Natl Acad Sci U S A, 1982. **79**(8): p. 2731-5.

173. McCormick, D.A. and R.F. Thompson, *Neuronal responses of the rabbit cerebellum during acquisition and performance of a classically conditioned nictitating membrane-eyelid response.* J Neurosci, 1984. **4**(11): p. 2811-22.

174. Jirenhed, D.A., F. Bengtsson, and G. Hesslow, *Acquisition, extinction, and reacquisition of a cerebellar cortical memory trace.* J Neurosci, 2007. **27**(10): p. 2493-502.

175. Kotani, S., S. Kawahara, and Y. Kirino, *Purkinje cell activity during learning a new timing in classical eyeblink conditioning.* Brain Res, 2003. **994**(2): p. 193-202.

176. McDonough, S.I. and B.P. Bean, *Mibefradil inhibition of T-type calcium channels in cerebellar purkinje neurons.* Mol Pharmacol, 1998. **54**(6): p. 1080-7.

177. Porcello, D.M., S.D. Smith, and J.R. Huguenard, *Actions of U-92032, a T-type Ca2+ channel antagonist, support a functional linkage between I(T) and slow intrathalamic rhythms.* J Neurophysiol, 2003. **89**(1): p. 177-85.

178. De Zeeuw, C.I. and A.S. Berrebi, *Postsynaptic targets of Purkinje cell terminals in the cerebellar and vestibular nuclei of the rat.* Eur J Neurosci, 1995. **7**(11): p. 2322-33.

179. Teune, T.M., J. van der Burg, C.I. de Zeeuw, J. Voogd, and T.J. Ruigrok, *Single Purkinje cell can innervate multiple classes of projection neurons in the cerebellar nuclei of the rat: a light microscopic and ultrastructural triple-tracer study in the rat.* J Comp Neurol, 1998. **392**(2): p. 164-78.

180. Medina, J.F., W.L. Nores, and M.D. Mauk, *Inhibition of climbing fibres is a signal for the extinction of conditioned eyelid responses.* Nature, 2002. **416**(6878): p. 330-3.

181. Bengtsson, F., P. Svensson, and G. Hesslow, *Feedback control of Purkinje cell activity by the cerebello-olivary pathway.* Eur J Neurosci, 2004. **20**(11): p. 2999-3005.

182. Hesslow, G. and M. Ivarsson, *Inhibition of the inferior olive during conditioned responses in the decerebrate ferret.* Exp Brain Res, 1996. **110**(1): p. 36-46.

183. Zhang, F., L.P. Wang, M. Brauner, J.F. Liewald, K. Kay, N. Watzke, P.G. Wood, E. Bamberg, G. Nagel, A. Gottschalk, and K. Deisseroth, *Multimodal fast optical interrogation of neural circuitry.* Nature, 2007. **446**(7136): p. 633-9.

184. Green, J.T. and J.E. Steinmetz, *Purkinje cell activity in the cerebellar anterior lobe after rabbit eyeblink conditioning.* Learn Mem, 2005. **12**(3): p. 260-9.

185. Kotani, S., S. Kawahara, and Y. Kirino, *Purkinje cell activity during classical eyeblink conditioning in decerebrate guinea pigs.* Brain Res, 2006. **1068**(1): p. 70-81.

186. Tracy, J.A. and J.E. Steinmetz, *Purkinje cell responses to pontine stimulation CS during rabbit eyeblink conditioning.* Physiol Behav, 1998. **65**(2): p. 381-6.

187. Christian, K.M., A.M. Poulos, D.G. Lavond, and R.F. Thompson, *Comment on "Cerebellar LTD and learning-dependent timing of conditioned eyelid responses".* Science, 2004. **304**(5668): p. 211; author reply 211.

188. Kistler, W.M. and C.I. De Zeeuw, *Time windows and reverberating loops: a reverse-engineering approach to cerebellar function.* Cerebellum, 2003. **2**(1): p. 44-54.

189. Hesslow, G., *Inhibition of classically conditioned eyeblink responses by stimulation of the cerebellar cortex in the decerebrate cat.* J Physiol, 1994. **476**(2): p. 245-56.

190. Pugh, J.R. and I.M. Raman, *Potentiation of Mossy Fiber EPSCs in the Cerebellar Nuclei by NMDA Receptor Activation followed by Postinhibitory Rebound Current.* Neuron, 2006. **51**(1): p. 113-23.

191. Aizenman, C.D. and D.J. Linden, *Rapid, synaptically driven increases in the intrinsic excitability of cerebellar deep nuclear neurons.* Nat Neurosci, 2000. **3**(2): p. 109-11.

192. Zhang, W., J.H. Shin, and D.J. Linden, *Persistent changes in the intrinsic excitability of rat deep cerebellar nuclear neurones induced by EPSP or IPSP bursts.* J Physiol, 2004. **561**(Pt 3): p. 703-19.

193. D'Angelo, E., P. Rossi, S. Armano, and V. Taglietti, *Evidence for NMDA and mGlu receptor-dependent long-term potentiation of mossy fiber-granule cell transmission in rat cerebellum.* J Neurophysiol, 1999. **81**(1): p. 277-87.

194. Soler-Llavina, G.J. and B.L. Sabatini, *Synapse-specific plasticity and compartmentalized signaling in cerebellar stellate cells.* Nat Neurosci, 2006. **9**(6): p. 798-806.

195. Kassardjian, C.D., Y.F. Tan, J.Y. Chung, R. Heskin, M.J. Peterson, and D.M. Broussard, *The site of a motor memory shifts with consolidation.* J Neurosci, 2005. **25**(35): p. 7979-85.

196. Shutoh, F., M. Ohki, H. Kitazawa, S. Itohara, and S. Nagao, *Memory trace of motor learning shifts transsynaptically from cerebellar cortex to nuclei for consolidation.* Neuroscience, 2006. **139**(2): p. 767-77.

197. Molineux, M.L., J.E. McRory, B.E. McKay, J. Hamid, W.H. Mehaffey, R. Rehak, T.P. Snutch, G.W. Zamponi, and R.W. Turner, *Specific T-type calcium channel isoforms are associated with distinct burst phenotypes in deep cerebellar nuclear neurons.* Proc Natl Acad Sci U S A, 2006. **103**(14): p. 5555-60.

198. Medina, J.F., M.R. Carey, and S.G. Lisberger, *The representation of time for motor learning.* Neuron, 2005. **45**(1): p. 157-67.

199. Chen, C. and R.F. Thompson, *Temporal specificity of long-term depression in parallel fiber--Purkinje synapses in rat cerebellar slice.* Learn Mem, 1995. **2**(3-4): p. 185-98.

200. Gallistel, C.R. and J. Gibbon, *Time, rate, and conditioning.* Psychol Rev, 2000. **107**(2): p. 289-344.

201. Baccus, S.A. and M. Meister, *Fast and slow contrast adaptation in retinal circuitry.* Neuron, 2002. **36**(5): p. 909-19.

202. Korenberg, M.J. and I.W. Hunter, *The identification of nonlinear biological systems: LNL cascade models.* Biol Cybern, 1986. **55**(2-3): p. 125-34.

203. Fino, E., J. Glowinski, and L. Venance, *Bidirectional activity-dependent plasticity at corticostriatal synapses.* J Neurosci, 2005. **25**(49): p. 11279-87.

204. Nambu, A. and R. Llinas, *Electrophysiology of globus pallidus neurons in vitro.* J Neurophysiol, 1994. **72**(3): p. 1127-39.

205. Lashley, K.S., *In search of the engram.* Symp. Soc. Exp. Biol., 1950. **4**: p. 454-482.

206. Bliss, T.V. and T. Lomo, *Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path.* J Physiol, 1973. **232**(2): p. 331-56.

207. Gallistel, C.R., *The organization of learning*. 1993, Cambridge: MIT Press.

208. Martin, S.J., P.D. Grimwood, and R.G. Morris, *Synaptic plasticity and memory: an evaluation of the hypothesis.* Annu Rev Neurosci, 2000. **23**: p. 649-711.

209. Hodgkin, A.L. and A.F. Huxley, *The dual effect of membrane potential on sodium conductance in the giant axon of Loligo.* J Physiol, 1952. **116**(4): p. 497-506.

210. Ferezou, I., S. Bolea, and C.C. Petersen, *Visualizing the cortical representation of whisker touch: voltage-sensitive dye imaging in freely moving mice.* Neuron, 2006. **50**(4): p. 617-29.

211. Yamaguchi, S., M. Kobayashi, S. Mitsui, Y. Ishida, G.T. van der Horst, M. Suzuki, S. Shibata, and H. Okamura, *View of a mouse clock gene ticking.* Nature, 2001. **409**(6821): p. 684.

212. Adelsberger, H., O. Garaschuk, and A. Konnerth, *Cortical calcium waves in resting newborn mice.* Nat Neurosci, 2005. **8**(8): p. 988-90.

213. Murayama, M., E. Perez-Garci, H.R. Luscher, and M.E. Larkum, *Fiberoptic system for recording dendritic calcium signals in layer 5 neocortical pyramidal cells in freely moving rats.* J Neurophysiol, 2007. **98**(3): p. 1791-805.

214. Poe, G.R., M.P. Kristensen, D.M. Rector, and R.M. Harper, *Hippocampal activity during transient respiratory events in the freely behaving cat.* Neuroscience, 1996. **72**(1): p. 39-48.

215. Helmchen, F., M.S. Fee, D.W. Tank, and W. Denk, *A miniature head-mounted two-photon microscope. high-resolution brain imaging in freely moving animals.* Neuron, 2001. **31**(6): p. 903-12.

216. Flusberg, B.A., E.D. Cocker, W. Piyawattanametha, J.C. Jung, E.L. Cheung, and M.J. Schnitzer, *Fiber-optic fluorescence imaging.* Nat Methods, 2005. **2**(12): p. 941-50.

217. Jung, J.C., A.D. Mehta, E. Aksay, R. Stepnoski, and M.J. Schnitzer, *In vivo mammalian brain imaging using one- and two-photon fluorescence microendoscopy.* J Neurophysiol, 2004. **92**(5): p. 3121-33.

218. Schaffer, C.B., B. Friedman, N. Nishimura, L.F. Schroeder, P.S. Tsai, F.F. Ebner, P.D. Lyden, and D. Kleinfeld, *Two-photon imaging of cortical surface microvessels reveals a robust redistribution in blood flow after vascular occlusion.* PLoS Biol, 2006. **4**(2): p. e22.

219. Gobel, W. and F. Helmchen, *New angles on neuronal dendrites in vivo.* J Neurophysiol, 2007. **98**(6): p. 3770-9.

220. Llinas, R.R., K.D. Walton, and E.J. Lang, *Cerebellum*, in *The Synaptic Organization of the Brain* G. Shepherd, Editor. 2004, Oxford University Press: Oxford.

221. Schiffmann, S.N., G. Cheron, A. Lohof, P. d'Alcantara, M. Meyer, M. Parmentier, and S. Schurmans, *Impaired motor coordination and Purkinje cell excitability in mice lacking calretinin.* Proc Natl Acad Sci U S A, 1999. **96**(9): p. 5257-62.

222. Harris, K.D., D.A. Henze, J. Csicsvari, H. Hirase, and G. Buzsaki, *Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements.* J Neurophysiol, 2000. **84**(1): p. 401-14.

223. Kleinfeld, D., P.P. Mitra, F. Helmchen, and W. Denk, *Fluctuations and stimulus-induced changes in blood flow observed in individual capillaries in layers 2 through 4 of rat neocortex.* Proc Natl Acad Sci U S A, 1998. **95**(26): p. 15741-6.

224. Yaksi, E. and R.W. Friedrich, *Reconstruction of firing rate changes across neuronal populations by temporally deconvolved Ca2+ imaging.* Nat Methods, 2006. **3**(5): p. 377-83.

225. Stosiek, C., O. Garaschuk, K. Holthoff, and A. Konnerth, *In vivo two-photon calcium imaging of neuronal networks.* Proc Natl Acad Sci U S A, 2003. **100**(12): p. 7319-24.

226. Sullivan, M.R., A. Nimmerjahn, D.V. Sarkisov, F. Helmchen, and S.S. Wang, *In vivo calcium imaging of circuit activity in cerebellar cortex.* J Neurophysiol, 2005. **94**(2): p. 1636-44.

227. Fiacco, T.A. and K.D. McCarthy, *Astrocyte calcium elevations: Properties, propagation, and effects on brain signaling.* Glia, 2006. **54**(7): p. 676-690.

228. Petzold, G.C., D.F. Albeanu, T.F. Sato, and V.N. Murthy, *Coupling of neural activity to blood flow in olfactory glomeruli is mediated by astrocytic pathways.* Neuron, 2008. **58**(6): p. 897-910.

229. Takano, T., G.F. Tian, W.G. Peng, N.H. Lou, W. Libionka, X.N. Han, and M. Nedergaard, *Astrocyte-mediated control of cerebral blood flow.* Nature Neuroscience, 2006. **9**(2): p. 260-267.

230. Rinberg, D., A. Koulakov, and A. Gelperin, *Sparse odor coding in awake behaving mice.* Journal of Neuroscience, 2006. **26**(34): p. 8857-8865.

231. Movshon, J.A., T.D. Albright, G.R. Stoner, N.J. Majaj, and M.A. Smith, *Cortical responses to visual motion in alert and anesthetized monkeys.* Nature Neuroscience, 2003. **6**(1): p. 3-3.

232. Takata, N. and H. Hirase, *Cortical layer 1 and layer 2/3 astrocytes exhibit distinct calcium dynamics in vivo.* PLoS ONE, 2008. **3**(6): p. e2525.

233. Metea, M.R. and E.A. Newman, *Calcium signaling in specialized glial cells.* Glia, 2006. **54**(7): p. 650-655.

234. Shibuki, K., H. Gomi, L. Chen, S.W. Bao, J.S.K. Kim, H. Wakatsuki, T. Fujisaki, J. Fujimoto, A. Katoh, T. Ikeda, C. Chen, R.F. Thompson, and S. Itohara, *Deficient cerebellar long-term depression, impaired eyeblink conditioning, and normal motor coordination in GFAP mutant mice.* Neuron, 1996. **16**(3): p. 587-599.

235. Stosiek, C., O. Garaschuk, K. Holthoff, and A. Konnerth, *In vivo two-photon calcium imaging of neuronal networks.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(12): p. 7319-7324.

236. Andersson, G. and D.M. Armstrong, *Complex Spikes in Purkinje-Cells in the Lateral Vermis (B-Zone) of the Cat Cerebellum during Locomotion.* Journal of Physiology-London, 1987. **385**: p. 107-134.

237. Leicht, R. and R.F. Schmidt, *Somatotopic Studies on Vermal Cortex of Cerebellar Anterior Lobe or Unanesthetized Cats.* Experimental Brain Research, 1977. **27**(5): p. 479-490.

238. Reichenbach, A., A. Siegel, M. Rickmann, J.R. Wolff, D. Noone, and S.R. Robinson, *Distribution of Bergmann Glial Somata and Processes - Implications for Function.* Journal Fur Hirnforschung, 1995. **36**(4): p. 509-517.

239. Deblas, A.L., *Monoclonal-Antibodies to Specific Astroglial and Neuronal Antigens Reveal the Cytoarchitecture of the Bergmann Glia Fibers in the Cerebellum.* Journal of Neuroscience, 1984. **4**(1): p. 265-273.

240. Nimmerjahn, A., F. Kirchhoff, J.N.D. Kerr, and F. Helmchen, *Sulforhodamine 101 as a specific marker of astroglia in the neocortex in vivo.* Nature Methods, 2004. **1**(1): p. 31-37.

241. Grosche, J., V. Matyash, T. Moller, A. Verkhratsky, A. Reichenbach, and H. Kettenmann, *Microdomains for neuron-glia interaction: parallel fiber signaling to Bergmann glial cells.* Nature Neuroscience, 1999. **2**(2): p. 139-143.

242. Gobel, W., B.M. Kampa, and F. Helmchen, *Imaging cellular network dynamics in three dimensions using fast 3D laser scanning.* Nature Methods, 2007. **4**(1): p. 73-79.

243. Beierlein, M. and W.G. Regehr, *Brief bursts of parallel fiber activity trigger calcium signals in Bergmann glia.* Journal of Neuroscience, 2006. **26**(26): p. 6958-6967.

244. Piet, R. and C.E. Jahr, *Glutamatergic and purinergic receptor-mediated calcium transients in Bergmann glial cells.* Journal of Neuroscience, 2007. **27**(15): p. 4027-4035.

245. Charles, A.C., J.E. Merrill, E.R. Dirksen, and M.J. Sanderson, *Intercellular Signaling in Glial-Cells - Calcium Waves and Oscillations in Response to Mechanical Stimulation and Glutamate.* Neuron, 1991. **6**(6): p. 983-992.

246. Ozden, I., H.M. Lee, M.R. Sullivan, and S.S.H. Wang, *Identification and clustering of event patterns from in vivo multiphoton optical recordings of neuronal ensembles.* J Neurophysiol, 2008. **100**(1): p. 495-503.

247. Huettel, S.A. and G. McCarthy, *Evidence for a refractory period in the hemodynamic response to visual stimuli as measured by MRI.* Neuroimage, 2000. **11**(5): p. 547-553.

248. Martindale, J., J. Berwick, C. Martin, Y.Z. Kong, Y. Zheng, and J.E.W. Mayhew, *Long duration stimuli and nonlinearities in the neural-haemodynamic coupling.* Journal of Cerebral Blood Flow and Metabolism, 2005. **25**(5): p. 651-661.

249. Mantz, J., J. Cordier, and C. Giaume, *Effects of General-Anesthetics on Intercellular Communications Mediated by Gap-Junctions between Astrocytes in Primary Culture.* Anesthesiology, 1993. **78**(5): p. 892-901.

250. Miyazaki, H., Y. Nakamura, T. Arai, and K. Kataoka, *Increase of glutamate uptake in astrocytes - A possible mechanism of action of volatile anesthetics.* Anesthesiology, 1997. **86**(6): p. 1359-1366.

251. Dombeck, D.A., A.N. Khabbaz, F. Collman, T.L. Adelman, and D.W. Tank, *Imaging large-scale neural activity with cellular resolution in awake, mobile mice.* Neuron, 2007. **56**: p. 43-57.

252. Iino, M., K. Goto, W. Kakegawa, H. Okado, M. Sudo, S. Ishiuchi, A. Miwa, Y. Takayasu, I. Saito, K. Tsuzuki, and S. Ozawa, *Glia-synapse interaction through Ca2+-permeable AMPA receptors in Bergmann glia.* Science, 2001. **292**(5518): p. 926-929.

253. Nett, W.J., S.H. Oloff, and K.D. McCarthy, *Hippocampal astrocytes in situ exhibit calcium oscillations that occur independent of neuronal activity.* Journal of Neurophysiology, 2002. **87**(1): p. 528-537.

254. Napper, R.M.A. and R.J. Harvey, *Number of Parallel Fiber Synapses on an Individual Purkinje-Cell in the Cerebellum of the Rat.* Journal of Comparative Neurology, 1988. **274**(2): p. 168-177.

255. Heinsen, H. and Y.L. Heinsen, *Quantitative Studies on Regional Differences in Purkinje-Cell Dendritic Spines and Parallel Fiber Synaptic Density.* Anatomy and Embryology, 1983. **168**(3): p. 361-370.

256. Zonta, M., M.C. Angulo, S. Gobbo, B. Rosengarten, K.A. Hossmann, T. Pozzan, and G. Carmignoto, *Neuron-to-astrocyte signaling is central to the dynamic control of brain microcirculation.* Nature Neuroscience, 2003. **6**(1): p. 43-50.

257. Metea, M.R. and E.A. Newman, *Glial cells dilate and constrict blood vessels: A mechanism of neurovascular coupling.* Journal of Neuroscience, 2006. **26**(11): p. 2862-2870.

258. Mulligan, S.J. and B.A. MacVicar, *Calcium transients in astrocyte endfeet cause cerebrovascular constrictions.* Nature, 2004. **431**(7005): p. 195-199.

259. Martin, C., J. Martindale, J. Berwick, and J. Mayhew, *Investigating neural-hemodynamic coupling and the hemodynamic response function in the awake rat.* Neuroimage, 2006. **32**(1): p. 33-48.

260. Logothetis, N.K. and B.A. Wandell, *Interpreting the BOLD signal.* Annual Review of Physiology, 2004. **66**: p. 735-769.

261. Moore, C.I. and R. Cao, *The hemo-neural hypothesis: On the role of blood flow in information processing.* Journal of Neurophysiology, 2008. **99**(5): p. 2035-2047.

262. Nimmerjahn, A., F. Kirchhoff, and F. Helmchen, *Resting microglial cells are highly dynamic surveillants of brain parenchyma in vivo.* Science, 2005. **308**(5726): p. 1314-1318.

263. Sullivan, M.R., A. Nimmerjahn, D.V. Sarkisov, F. Helmchen, and S.S.H. Wang, *In vivo calcium imaging of circuit activity in cerebellar cortex.* Journal of Neurophysiology, 2005. **94**(2): p. 1636-1644.

264. Nguyen, Q.T., P.S. Tsai, and D. Kleinfeld, *MPScope: A versatile software suite for multiphoton microscopy.* Journal of Neuroscience Methods, 2006. **156**(1-2): p. 351-359.

265. Zuo, Y., J.L. Lubischer, H. Kang, L. Tian, M. Mikesh, A. Marks, V.L. Scofield, S. Maika, C. Newman, P. Krieg, and W.J. Thompson, *Fluorescent proteins expressed in mouse transgenic lines mark subsets of glia, neurons, macrophages, and dendritic cells for vital examination.* Journal of Neuroscience, 2004. **24**(49): p. 10999-11009.

266. Thevenaz, P., U.E. Ruttimann, and M. Unser, *A pyramid approach to subpixel registration based on intensity.* Ieee Transactions on Image Processing, 1998. **7**(1): p. 27-41.

267. Kerr, J.N., D. Greenberg, and F. Helmchen, *Imaging input and output of neocortical networks in vivo.* Proc Natl Acad Sci U S A, 2005. **102**(39): p. 14063-8.

268. Mrsic-Flogel, T.D., S.B. Hofer, K. Ohki, R.C. Reid, T. Bonhoeffer, and M. Hubener, *Homeostatic regulation of eye-specific responses in visual cortex during ocular dominance plasticity.* Neuron, 2007. **54**(6): p. 961-72.

269. Ohki, K., S. Chung, P. Kara, M. Hubener, T. Bonhoeffer, and R.C. Reid, *Highly ordered arrangement of single neurons in orientation pinwheels.* Nature, 2006. **442**(7105): p. 925-8.

270. Orger, M.B., A.R. Kampff, K.E. Severi, J.H. Bollmann, and F. Engert, *Control of visually guided behavior by distinct populations of spinal projection neurons.* Nat Neurosci, 2008. **11**(3): p. 327-33.

271. Duemani Reddy, G., K. Kelleher, R. Fink, and P. Saggau, *Three-dimensional random access multiphoton microscopy for functional imaging of neuronal activity.* Nature Neuroscience, 2008. **11**(6): p. 713-20.

272. Gobel, W., B.M. Kampa, and F. Helmchen, *Imaging cellular network dynamics in three dimensions using fast 3D laser scanning.* Nat Methods, 2007. **4**(1): p. 73-9.

273. Niell, C.M. and S.J. Smith, *Functional imaging reveals rapid development of visual response properties in the zebrafish tectum.* Neuron, 2005. **45**(6): p. 941-51.

274. Fee, M.S., P.P. Mitra, and D. Kleinfeld, *Automatic sorting of multiple unit neuronal signals in the presence of anisotropic and non-Gaussian variability.* J Neurosci Methods, 1996. **69**(2): p. 175-88.

275. Batista, A.P., G. Santhanam, B.M. Yu, S.I. Ryu, A. Afshar, and K.V. Shenoy, *Reference frames for reach planning in macaque dorsal premotor cortex.* J Neurophysiol, 2007. **98**(2): p. 966-83.

276. Csicsvari, J., H. Hirase, A. Czurko, and G. Buzsaki, *Reliability and state dependence of pyramidal cell-interneuron synapses in the hippocampus: an ensemble approach in the behaving rat.* Neuron, 1998. **21**(1): p. 179-89.

277. Meister, M., *Multineuronal codes in retinal signaling.* Proc Natl Acad Sci U S A, 1996. **93**(2): p. 609-14.

278. Leonardo, A. and M.S. Fee, *Ensemble coding of vocal control in birdsong.* J Neurosci, 2005. **25**(3): p. 652-61.

279. Helmchen, F., K. Imoto, and B. Sakmann, *Ca2+ buffering and action potential-evoked Ca2+ signaling in dendrites of pyramidal neurons.* Biophys J, 1996. **70**(2): p. 1069-81.

280. Reidl, J., J. Starke, D.B. Omer, A. Grinvald, and H. Spors, *Independent component analysis of high-resolution imaging data identifies distinct functional domains.* Neuroimage, 2007. **34**(1): p. 94-108.

281. Reinert, K.C., W. Gao, G. Chen, and T.J. Ebner, *Flavoprotein autofluorescence imaging in the cerebellar cortex in vivo.* J Neurosci Res, 2007. **85**(15): p. 3221-32.

282. Hastie, T., R. Tibshirani, and J.H. Friedman, *The elements of statistical learning : data mining, inference, and prediction : with 200 full-color illustrations*. 2001, New York: Springer.

283. Mitra, P.P. and B. Pesaran, *Analysis of dynamic brain imaging data.* Biophys J, 1999. **76**(2): p. 691-708.

284. Sengupta, A.M. and P.P. Mitra, *Distributions of singular values for some random matrices.* Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics, 1999. **60**(3): p. 3389-92.

285. Bell, A.J. and T.J. Sejnowski, *The "independent components" of natural scenes are edge filters.* Vision Res, 1997. **37**(23): p. 3327-38.

286. Hyvarinen, A. and E. Oja, *Independent component analysis: algorithms and applications.* Neural Netw, 2000. **13**(4-5): p. 411-30.

287. Nimmerjahn, A., E. Mukamel, and M. Schnitzer, *Locomotion triggers concerted calcium activation in cerebellar Bergmann glial networks that conincides with hemodynamic modulation*, in *Soc. for Neurosci.* Society for Neuroscience, 2008.

288. Palay, S.L. and V. Chan-Palay, *Cerebellar cortex: cytology and organization*. 1974, Berlin, Heidelberg, New York,: Springer.

289. Llinas, R.R., K.D. Walton, and E.J. Lang, *Cerebellum*, in *The synaptic organization of the brain*, G.M. Shepherd, Editor. 2004, Oxford University Press: Oxford ; New York. p. 271-310.

290. Servais, L., B. Bearzatto, R. Hourez, B. Dan, S.N. Schiffmann, and G. Cheron, *Effect of simple spike firing mode on complex spike firing rate and waveform in cerebellar Purkinje cells in non-anesthetized mice.* Neuroscience letters, 2004. **367**(2): p. 171-6.

291. Karthikesh, A.H.a.R. *Sparse priors on the mixing matrix in independent component analysis*. in *Proc. Int. Workshop on Independent Component Analysis and Blind Signal Separation (ICA2000)*. 2000.

292. Hyvarinen, A., *Fast and robust fixed-point algorithms for independent component analysis.* IEEE Trans Neural Netw, 1999. **10**(3): p. 626-34.

293. Grosche, J., V. Matyash, T. Möller, A. Verkhratsky, A. Reichenbach, and H. Kettenmann, *Microdomains for neuron-glia interaction: parallel fiber signaling to Bergmann glial cells.* Nature Neuroscience, 1999. **2**(2): p. 139-43.

294. Beierlein, M. and W.G. Regehr, *Brief bursts of parallel fiber activity trigger calcium signals in bergmann glia.* J Neurosci, 2006. **26**(26): p. 6958-67.

295. Piet, R. and C.E. Jahr, *Glutamatergic and purinergic receptor-mediated calcium transients in Bergmann glial cells.* J Neurosci, 2007. **27**(15): p. 4027-35.

296. Fawcett, T., *An introduction to ROC analysis.* Pattern Recognition Letters, 2006. **27**(8): p. 861-874.

297. Sasaki, T., N. Takahashi, N. Matsuki, and Y. Ikegaya, *Fast and accurate detection of action potentials from somatic calcium fluctuations.* J Neurophysiol, 2008.

298. Wood, F., M.J. Black, C. Vargas-Irwin, M. Fellows, and J.P. Donoghue, *On the variability of manual spike sorting.* IEEE Trans Biomed Eng, 2004. **51**(6): p. 912-8.

299. Bhatt, D.H., D.L. McLean, M.E. Hale, and J.R. Fetcho, *Grading movement strength by changes in firing intensity versus recruitment of spinal interneurons.* Neuron, 2007. **53**(1): p. 91-102.

300. Busche, M.A., G. Eichhoff, H. Adelsberger, D. Abramowski, K.H. Wiederhold, C. Haass, M. Staufenbiel, A. Konnerth, and O. Garaschuk, *Clusters of hyperactive neurons near amyloid plaques in a mouse model of Alzheimer's disease.* Science, 2008. **321**(5896): p. 1686-9.

301. Ohki, K., S. Chung, Y.H. Ch'ng, P. Kara, and R.C. Reid, *Functional imaging with cellular resolution reveals precise micro-architecture in visual cortex.* Nature, 2005. **433**(7026): p. 597-603.

302. Siegel, R.M., J.R. Duann, T.P. Jung, and T. Sejnowski, *Spatiotemporal dynamics of the functional architecture for gain fields in inferior parietal lobule of behaving monkey.* Cereb Cortex, 2007. **17**(2): p. 378-90.

303. Fiacco, T. and K. McCarthy, *Astrocyte calcium elevations: Properties, propagation, and effects on brain signaling.* Glia, 2006. **54**(7): p. 676-690.

304. Greenberg, D.S. and J.N. Kerr, *Automated correction of fast motion artifacts for two-photon imaging of awake animals.* Journal of Neuroscience Methods, 2008. **(in press)**.

305. McKeown, M.J., S. Makeig, G.G. Brown, T.P. Jung, S.S. Kindermann, A.J. Bell, and T.J. Sejnowski, *Analysis of fMRI data by blind separation into independent spatial components.* Hum Brain Mapp, 1998. **6**(3): p. 160-88.

306. Sugihara, I., *Microzonal projection and climbing fiber remodeling in single olivocerebellar axons of newborn rats at postnatal days 4-7.* J Comp Neurol, 2005. **487**(1): p. 93-106.

307. Marshall, S.P., R.S. van der Giessen, C.I. de Zeeuw, and E.J. Lang, *Altered olivocerebellar activity patterns in the connexin36 knockout mouse.* Cerebellum, 2007: p. 1-13.

308. Margrie, T.W., M. Brecht, and B. Sakmann, *In vivo, low-resistance, whole-cell recordings from neurons in the anaesthetized and awake mammalian brain.* Pfl√°gers Archiv : European journal of physiology, 2002. **444**(4): p. 491-8.

309. Llinas, R., in *Motor Control: Concepts and Issues*, D.R. Humphrey and H.-J. Freund, Editors. 1991, Wiley: New York. p. 223-242.

310. Ekerot, C.F., M. Garwicz, and J. Schouenborg, *Topography and nociceptive receptive fields of climbing fibres projecting to the cerebellar anterior lobe in the cat.* The Journal of Physiology, 1991. **441**: p. 257-74.

311. Fukuda, M., T. Yamamoto, and R. Llinas, *The isochronic band hypothesis and climbing fibre regulation of motricity: an experimental study.* Eur J Neurosci, 2001. **13**(2): p. 315-26.

312. Schnitzer, M.J. and M. Meister, *Multineuronal firing patterns in the signal from eye to brain.* Neuron, 2003. **37**(3): p. 499-511.

313. Thevenaz, P., U.E. Ruttimann, and M. Unser, *A pyramid approach to subpixel registration based on intensity.* IEEE Trans Image Process, 1998. **7**(1): p. 27-41.

314. MacKay, D.J.C., *Information theory, inference, and learning algorithms*. 2003, Cambridge, UK ; New York: Cambridge University Press.

315. Karthikesh, R. and A. Hyvarinen. *Sparse priors on the mixing matrix in independent component analysis*. in *Proc. Int. Workshop on Independent Component Analysis and Blind Signal Separation (ICA2000)*. 2000.

316. Zhang, K. and L. Chan. *ICA with Sparse Connections*. in *IDEAL*. 2006. Burgos, Spain: Springer.

317. Haralick, R.M. and L.G. Shapiro, *Computer and Robot Vision*. Vol. 1. 1992: Addison-Wesley.

318. Nguyen, Q.T., P.S. Tsai, and D. Kleinfeld, *MPScope: a versatile software suite for multiphoton microscopy.* J Neurosci Methods, 2006. **156**(1-2): p. 351-9.