Neuron
Neurotechnique

Automated Analysis of Cellular Signals from Large-Scale Calcium Imaging Data

Eran A. Mukamel,[1,3,*] Axel Nimmerjahn,[1] and Mark J. Schnitzer[1,2,*]
[1]James H. Clark Center for Biomedical Engineering and Sciences
[2]Howard Hughes Medical Institute
Stanford University, Stanford CA 94305, USA
[3]Present address: Center for Brain Science, Harvard University, Cambridge, MA 02138, USA
*Correspondence: emukamel@fas.harvard.edu (E.A.M.), mschnitz@stanford.edu (M.J.S.)
DOI 10.1016/j.neuron.2009.08.009

SUMMARY

Recent advances in fluorescence imaging permit studies of Ca^{2+} dynamics in large numbers of cells, in anesthetized and awake behaving animals. However, unlike for electrophysiological signals, 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 loading Ca^{2+} indicators into many cells have enabled recent imaging studies of the dynamics of hundreds of neurons and astrocytes (Gobel et al., 2007; Greenberg et al., 2008; Mrsic-Flogel et al., 2007; Nimmerjahn et al., 2009; Ohki et al., 2005; Orger et al., 2008; Stosiek et al., 2003). However, computational techniques for extracting cellular signals from Ca^{2+} imaging data lag behind and are mainly region of interest (ROI) analyses. These are typically manual (Dombeck et al., 2007; Gobel et al., 2007; Kerr et al., 2005; Niell and Smith, 2005) or semiautomated (Ozden et al., 2008) means of identifying cells and cannot be easily scaled to handle the largest data sets without undue human labor. Moreover, ROI analyses have largely been based on heuristic definitions of the morphology of specific cell types (Gobel et al., 2007; Ohki et al., 2005; Ozden et al., 2008) rather than general principles for decomposing a data set into constituent signal sources. Thus, current analyses are prone to crosstalk in the signals extracted from adjacent cells and surrounding neuropil. The present mismatch between the capabilities for Ca^{2+} imaging and those for analyzing the data restricts the capacity to attain biological insights.

This situation partly resembles that of the early 1990s, when multielectrode techniques were blossoming but standardized spike-sorting algorithms had yet to arise. Today, automated spike sorting is widely used to assign spikes to individual cells (Fee et al., 1996; Lewicki, 1998) and has enabled key advances in understanding neural coding (Batista et al., 2007; Csicsvari et al., 1998; Meister, 1996). An automated procedure for extracting cellular Ca^{2+} signals would be a similar enabler of scientific progress. However, the challenges in devising such a procedure are distinct from those in spike sorting.

Spike-sorting routines tend to rely on two basic ideas. First, the temporal waveforms for spikes from different cells are often sufficiently dissimilar to provide a basis for spike classification. Second, the activity of individual cells is often recorded on multiple electrodes, aiding assignment of spikes based on their relative amplitudes on different recording channels. Neither approach works well for imaging data. First, Ca^{2+} activity waveforms are strongly dictated by intracellular Ca^{2+} buffering and the dye's binding kinetics (Helmchen et al., 1996), which do not provide strong signatures of individual cells' identities. Second, single-image pixels can contain a complex mixture of signals from neuropil, neurons, astrocytes, and noise. It is nontrivial to disentangle these signals without suffering crosstalk and to find the shapes and locations of each cell. A guiding principle is needed to help extract cells' locations and activities.

We formulated such a principle by considering the manner in which intracellular [Ca^{2+}] transiently rises above background levels during cellular events such as action potentials. Brief periods of elevated [Ca^{2+}] are typically sparsely interspersed among many more background-dominated time frames. Sparseness also holds in the spatial domain if each cell occupies only a small subset of pixels. Thus, Ca^{2+} signals' sparseness should be a general attribute that is quantifiable by simple measures, such as the skewness of amplitude distributions.

This reasoning led us to an algorithm that estimates cells' locations and activities by parsing data into a combination of statistically independent signals, each with a high sparseness. The

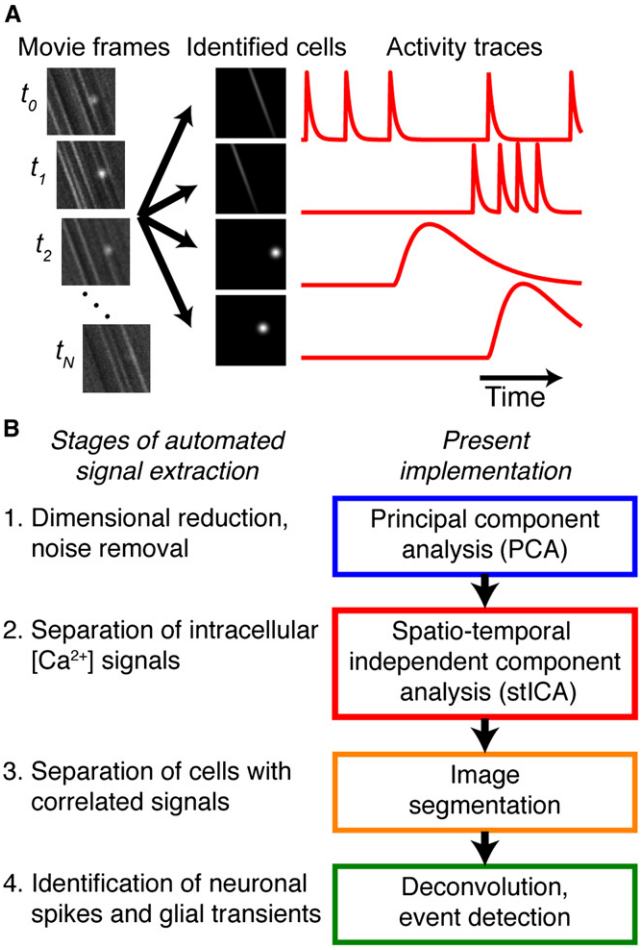

Figure 1. Analytical Stages of Automated Cell Sorting

(A) The goal of cell sorting is to extract cellular signals from imaging data (left) by estimating spatial filters (middle) and activity traces (right) for each cell. The example depicts typical fluorescence transients in the cerebellar cortex as observed in optical cross-section. Transients in Purkinje cell dendrites arise across elongated areas seen as stripes in the movie data. Transients in Bergmann glial fibers tend to be more localized, appearing ellipsoidal.

(B) Automated cell sorting has four stages that address specific analysis challenges.

algorithm requires no preconceptions of cells' appearances and little user supervision, and it relies on an independent component analysis (ICA) (Bell and Sejnowski, 1995; Brown et al., 2001; Reidl et al., 2007) (Figure 1). ICA has been used previously for analyses of electroencephalography (EEG) (Makeig et al., 1997), magnetoencephalography (MEG) (Guimaraes et al., 2007), and functional magnetic resonance imaging (fMRI) (Beckmann and Smith, 2004; McKeown et al., 1998) data, but a challenge has concerned the physiological interpretation of the identified sources, which can be mixtures of signals from different recording channels or brain areas. We reasoned that for ICA analyses of Ca^{2+} imaging data, such interpretative issues should be much reduced, since cells' properties can be corroborated by other experimental means, including in the same animals examined by imaging. In studies of human brain activity, corroborative data were much harder to obtain in living subjects.

We validated our method using simulated movies mimicking Ca^{2+} imaging data acquired in cerebellar cortex. Our sorting procedure provided superior signal estimates and lower susceptibility to crosstalk than reconstructions done by ROI analysis. We also tested our analysis on data recorded by two-photon microscopy in the cerebellar cortex of awake behaving mice, from which we extracted Ca^{2+} signals of up to >100 total Purkinje cells and Bergmann glia.

To illustrate our method's utility, we applied it to study the spatiotemporal organization of Purkinje cells' Ca^{2+} spiking activity in behaving mice. We found that synchronously active cells cluster into neighborhoods ~7–18 cells across in the mediolateral dimension. We identify these as cerebellar microzones, small patches of Purkinje cells receiving similar climbing fiber input (Andersson and Oscarsson, 1978). Our data revealed that microzones of awake animals have sharply delineated mediolateral boundaries, to a precision of about a single cell.

We addressed the longstanding question of whether microzones have stable anatomical boundaries (Andersson and Oscarsson, 1978) or are dynamic entities whose cellular constituents vary across behavioral states (Lang et al., 1999; Welsh et al., 1995). We found that during mouse locomotion microzones' spatial organization was unchanged from that in awake but resting animals, consistent with the idea that microzones are stationary anatomical units. These findings reveal basic features of cerebellar dynamics and highlight the impact of automated procedures for analyzing imaging data.

RESULTS

Principles for Extracting Cellular Signals

The purpose of cell sorting is to identify spatial filters and corresponding time traces that represent the locations and dynamics of individual cells (Figure 1A). Our procedure does this in four stages (Figure 1B and Supplemental Software Toolbox).

Physiological signals are buried in high-dimensional data that might contain, e.g., $N_x \sim 10^5$ pixels and $N_t \sim 10^4$ time frames. A first stage of analysis should reduce this dimensionality, since specifying cells' spatial filters and activities would otherwise require a large set of parameters. In principle, these parameters could be found, but tuning many parameters can hamper signal extraction (Hastie et al., 2001). For dimensional reduction, we used an established method, principal component analysis (PCA) (Table S1), to find and discard dimensions that mainly encoded noise (Mitra and Pesaran, 1999). Other means of dimensional reduction might also be suitable (Discussion).

PCA identifies a linear transformation of the data yielding basis vectors, the principal components, which are rank ordered by the variance of signals along each basis vector. Truncation of the highest-ranked components with the smallest variances often does not interfere with an approximate reconstruction of the data. On the contrary, if ongoing background noise has smaller variance than transient but strong signals, then removing higher-order components from the data can eliminate substantial noise. In our studies of cerebellum, after identifying the noise floor and truncating principal components representing noise

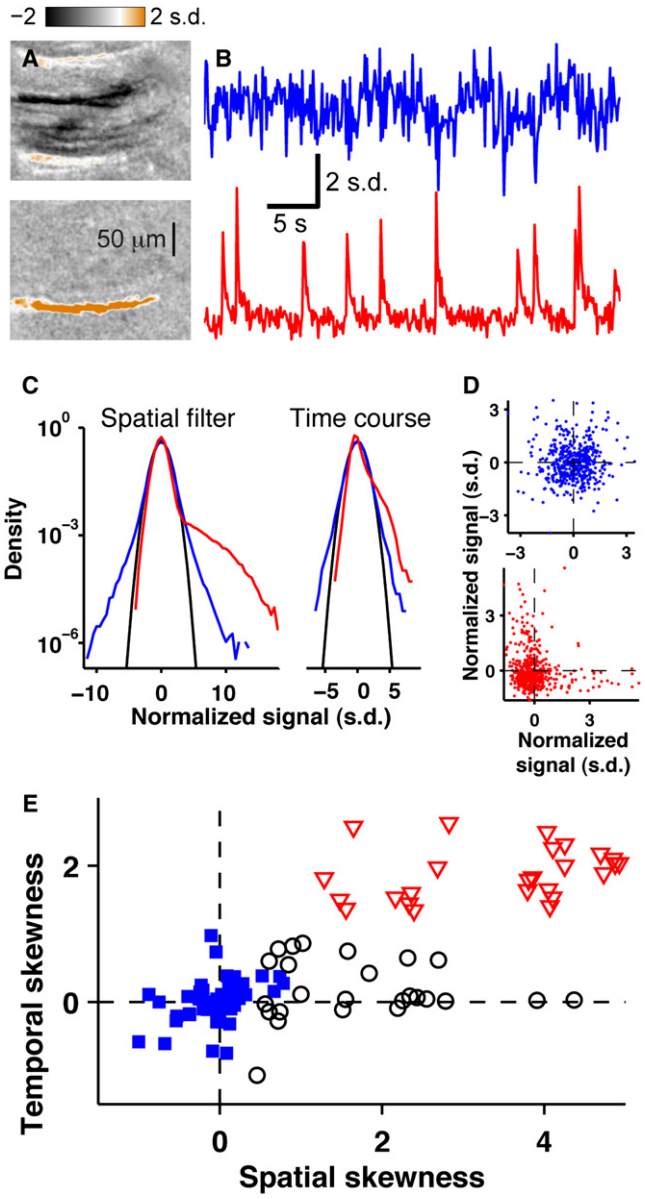

Figure 2. Independent Component Analysis Identifies Sparse, Intracellular Ca^{2+} Signals

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

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

(C) Distributions of pixel intensity values in spatial filters for the principal (blue) and independent (red) components in (A) (left). Distributions of the two components' time courses, across all movie frames (right). Gaussian distributions (black) shown for comparison, revealing the independent component's skewness.

(D) Joint distribution of time course values for two example principal (blue) and independent (red) components in data from mouse cerebellar cortex. Extended tails in the latter distribution reflect times at which Ca^{2+}-dependent signals transiently increased.

(E) Temporal versus spatial skewness for 50 principal (closed blue squares) and 50 independent (open symbols) components in data from the mouse cerebellar cortex. Using a k-means cluster analysis (k = 2) of skewness values, the independent components were divided into those representing cells (open red triangles) or other, noncellular independent components (open black circles), such as motion artifacts. These assignments made automatically match those done manually by visual inspection.

(Figure S1) (Mitra and Pesaran, 1999), we achieved a reduced dimensionality, K, much less than N_x or N_t. However, 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 variance to identify data components; in practice, time 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 expectation is that such sources should be localized, with a sparse or skewed spatial distribution of pixel weights for each cell, such that most pixels have almost no weight but a few are strongly weighted at the cell's location. Skewness, a distribution's third moment normalized by the cube of the standard deviation, provides a means of characterizing sparseness and searching for statistical signatures of individual cells. Similarly, if the intervals between cells' brief rises in $[Ca^{2+}]$ are greater than their durations, the distributions of signals' amplitudes will also be sparse and skewed (Discussion).

These expectations motivated our application of ICA following PCA. ICA is suited to extracting cellular signals since it seeks pairs of spatial filters and time traces that are sparse and statistically independent of one another. In practice, we found that cells can exhibit significant signal correlations and ICA will still work well, even for cell pairs with moderately high correlation coefficients. We used a version of the *FastICA* algorithm that maximizes the sparseness of extracted signals (Hyvarinen and Oja, 2000), which we expressed by a combination of spatial and temporal skewness (Supplemental Experimental Procedures).

ICA Yields Ca^{2+} Signals that Are Statistically Sparse

We examined the sparseness of Ca^{2+} signals recorded by two-photon microscopy in the cerebellar vermis of awake, head-restrained mice (Figure 2). 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 (Nimmerjahn et al., 2009; Ozden et al., 2008) (Supplemental Experimental Procedures). The mice stood on an exercise ball and could walk or run at liberty during imaging (Dombeck et al., 2007; Nimmerjahn et al., 2009). This permitted recording of the dendritic Ca^{2+} spikes associated with complex (Na^+ and Ca^{2+}) action potentials in Purkinje cells (Flusberg et al., 2008; Ozden et al., 2008). We also observed Ca^{2+} activation in Bergmann glial fibers (Nimmerjahn et al., 2009). Application of PCA to the data yielded principal components that were generally not spatially localized and failed to isolate 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. ICA transformed these components into new basis vectors with more distinctive properties.

A majority of independent components had spatial filters far more localized than those of principal components and could reasonably represent individual cells (Figures 2A and 2B, bottom row). Spatial filters found by ICA had skewed distributions (4.4 ± 2.7, mean skewness ± SD, n = 300 independent components in 5 mice, versus −0.009 ± 0.26 for 504 principal components), with the weights of small subsets of pixels up to 19 standard deviations greater than background pixels (Figure 2C, left). Many filters resembled a stripe in the rostral-caudal direction, with the long axis perpendicular to the folium of the cerebellar cortex (Figure 2A). This is the expected optical cross-section through the dendritic tree of a single Purkinje cell, since these cells have nearly planar dendritic trees lying perpendicular to the field of view (Flusberg et al., 2008; Ozden et al., 2008). The stripes' widths (7.0 ± 2.3 μm, mean ± SD; n = 199 dendrites in 5 mice) and lengths (90 ± 40 μm) as estimated by twice the standard deviations of Gaussian fits were consistent with values derived anatomically (Llinas et al., 2004; Palay and Chan-Palay, 1974) or in prior in vivo Ca^{2+} imaging studies of Purkinje cells (Flusberg et al., 2008; Ozden et al., 2008).

In the temporal domain, Ca^{2+} activity traces of the independent components also had sparse statistics. Signal distributions were asymmetric and non-Gaussian, with sparse sets of time frames during which fluorescence signals rose by up to eight standard deviations (Figure 2C, right). For the striped independent components representing Purkinje cell dendritic trees, the time traces yielded Ca^{2+} spike rates and waveforms similar to those from prior manual or semiautomated analyses of cerebellar Ca^{2+} imaging data (Flusberg et al., 2008; Ozden et al., 2008).

After verifying the sparseness of signals found by ICA, we examined their independence by plotting joint distributions of signals from pairs of components, each representing one Purkinje cell. The data of Figure 2D (bottom) is typical of 276 pairs we studied. In most time frames, signals were close to background levels. Each cell was active during a subset of time frames, but the 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 signal amplitudes in both cells, consistent with the cells' activities being statistically independent. We did encounter some cell pairs whose activity was sufficiently correlated that the pair was extracted as a single independent component. We handled these instances in the third stage of our analysis (see below).

Interestingly, the skewness often provided a means for distinguishing components representing individual cells from those representing artifacts, such as due to brain motion. Nearly all independent components in experimental data had higher spatial skewness than principal components (Figure 2E). Among the independent components, those we recognized by eye to be Purkinje cell or Bergmann glial signals, based on our prior studies of these cells (Flusberg et al., 2008; Nimmerjahn et al., 2009), generally had higher temporal skewness than those representing noise or artifacts (Figure 2E). After sorting independent components by their temporal skewness, it was usually quick and straightforward to segregate and reject artifactual components by inspection. Alternatively, signal components could often be isolated via automated clustering of signals according to their skewness (Figures 2E and S1). To be conservative, we regularly examined spatial filters and time traces visually before accepting them as cellular signals. When we inspected components with high skewness, we found several types of 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-sections of the corresponding number of Purkinje cell dendritic trees (Figures 2A, 4, and 5). Signals from these stripes exhibited brief increases in fluorescence (<50 ms rise), followed by an approximately exponential decay with time constant 124 ± 63 ms (mean ± SD, n = 150 cells in 5 mice). In awake mice, these events 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 were determined by performing a temporal deconvolution on the time traces to account for the Ca^{2+} indicator's kinetics, followed by a threshold operation for spike detection (Figure 1B) (Yaksi and Friedrich, 2006). The rates are consistent with those from physiological recordings of complex spiking (Lang et al., 1999; Servais et al., 2004) and Ca^{2+} imaging studies of Purkinje cells (Flusberg et al., 2008; Gobel and Helmchen, 2007b; Ozden et al., 2008). Through simultaneous optical and single-unit electrophysiological recordings, we verified that Ca^{2+} spikes found by our procedure corresponded to complex spikes (Figures 5C and S2). ICA also regularly yielded signals with slower dynamics consistent with Ca^{2+} activation in fibers of Bergmann glia (Nimmerjahn et al., 2009), 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 in having diffuse spatial weighting, strong correlations with brain movement, and kinetics too fast to be consistent with those of the Ca^{2+} indicator (Figure S6). We excluded such components from further analyses (Figure 2E).

Combining Spatial and Temporal Information Can Improve Cell Sorting

After obtaining these results, we sought to fine tune ICA's ability to isolate cells. Most prior usages of ICA in brain imaging have maximized spatial sparseness (Reidl et al., 2007), but it is reasonable to consider also temporal sparseness. We did this using spatiotemporal ICA, which optimizes a linear combination of spatial and temporal skewness (Stone et al., 2002). One parameter, μ, set the relative weight of the two. We sought μ values that improved estimation accuracies (Figure S3). To explore, we created artificial movies that mimicked our recordings from the cerebellum and contained signals from Purkinje cells and Bergmann glia. We included photon shot noise, the main noise source in two-photon microscopy, 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, with *F* = 100% implying a perfect reconstruction.

Distributions of *F* as a function of μ depended on multiple parameters, including the field of view size and thus the number of cells, the numbers of pixels and time frames, cells' mean spike

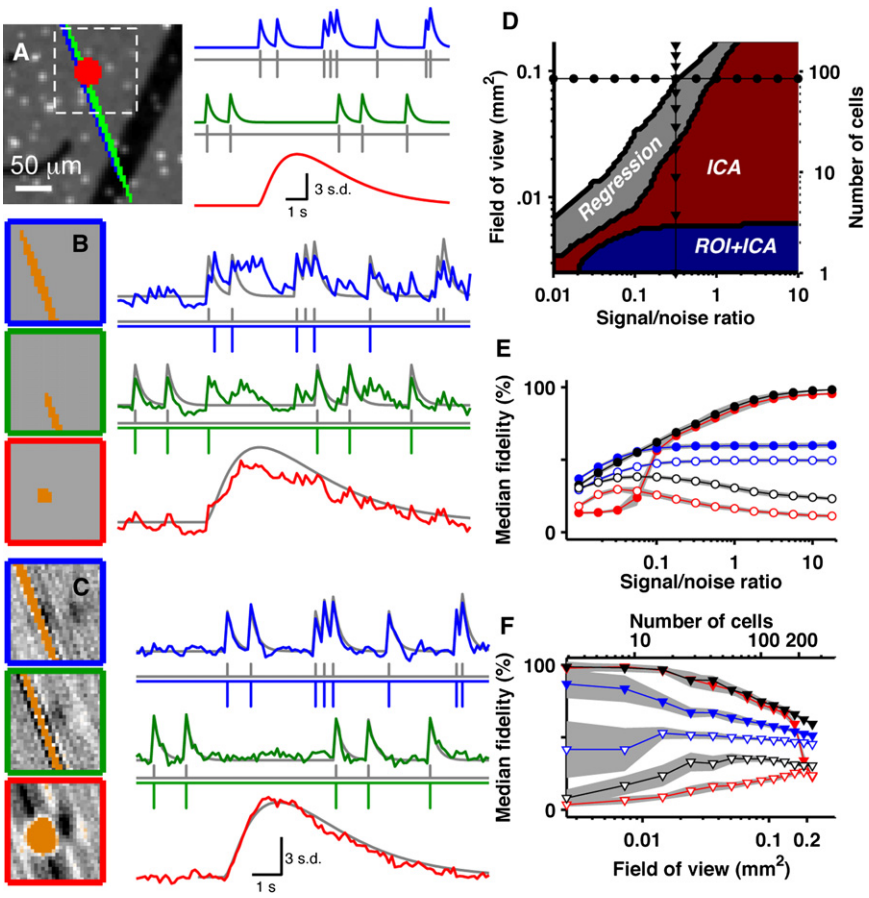

Figure 3. Automated Cell Sorting Outperforms Region of Interest Analysis

(A) Three spatially overlapping sources of cellular signals in an artificial data set mimicking Ca^{2+} imaging data from cerebellar cortex. Background grayscale image shows simulated field of view containing dark blood vessels and brightly labeled interneuron somata. 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 spatiotemporal ICA ($\mu = 0.5$). The high-fidelity signal estimate allows correct identification of all dendritic Ca^{2+} spikes.

(D) Diagram on logarithmic axes of how median signal fidelity depends on signal/noise ratio and field of view size. Shaded 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 (white). An estimator derived by linear regression achieved >75% fidelity in the gray region in cross-validation with distinct testing and training data sets. Black circles mark parameter values used in (E); black triangles mark values used in (F).

(E and F) Median fidelity (solid symbols) and crosstalk (open symbols) of signals extracted by ICA (red), ROI (blue), and linear regression analysis (black), as a function of the signal/noise ratio (E) and the field of view (F). Shaded areas indicate ±1 SD across ten simulation batches. Field of view is fixed at 0.09 mm^2 for (E); signal/noise ratio is 0.31 for (F).

rate relative to the frame rate, and the amplitudes of cells' signals above background. In nearly all situations representative of our experiments, spatial ICA ($\mu = 0$) was superior to temporal ICA ($\mu = 1$), leading to higher median fidelities and components easily recognizable as Purkinje cells. Temporal ICA often failed to yield components recognizable as cells. Spatiotemporal ICA with $\mu < 0.5$ nearly always led to the highest fidelities, although the gains over a purely spatial ICA were only weakly sensitive to μ and modest, with ~1%–10% improvements in median fidelity and ~0%–15% improvements in fidelities for individual cells (Figure S3). An exception was that, if the frame rate was much higher than the spike rates, the time traces became very sparse, and a temporal weighted ICA ($\mu \approx 1$) yielded the highest fidelities. Tuning μ yielded the most benefit in challenging cases with substantial mixing of cells' signals in individual pixels. With real data, we found by exploration that spatiotemporal ICA ($\mu \approx$ 0.1–0.2) extracted the most components resembling Purkinje cells, so we habitually used this approach.

Comparison of Independent Component and Region of Interest Analyses

To compare ICA and ROI analyses, we used a form of ROI analysis that is the best a human analyst could do if she correctly identified sets of example movie frames during which each cell is active. The most challenging situations in our simulations and experiments involved cells with overlapping pixel sets (Figures 3A and 5A) or pixels with signals from cells and neuropil (Figure S4). In such cases, ROI analysis had difficulty removing crosstalk (Figures 3B and S4). With spatiotemporal ICA, crosstalk was much reduced, since independent components had negatively weighted surround regions in the spatial filters that subtracted signals from overlapping cells (Figures 3C and S4).

We studied how signal amplitudes and the field of view area affected signal extraction (Figure 3D). For each pixel, photon counts obeyed Poisson statistics, with the mean intensity set by a signal-to-noise ratio parameter expressing signals' dynamic range compared to background fluorescence levels (Supplemental Data). In simulations with a fixed field of view (0.09 mm^2; ~100 cells), ICA used increases in signal-to-noise ratio to improve reconstruction fidelities and reduce crosstalk between cells (Figure 3E). ICA's performance approached an optimum for signal-to-noise ratios >0.3, at which the median fidelity approached 95% and >80% of extracted signals had $F > 75\%$. This fidelity was close to the theoretical limit set by a linear regression analysis, which defines the best achievable by any linear combination of principal components (Figure 3D and

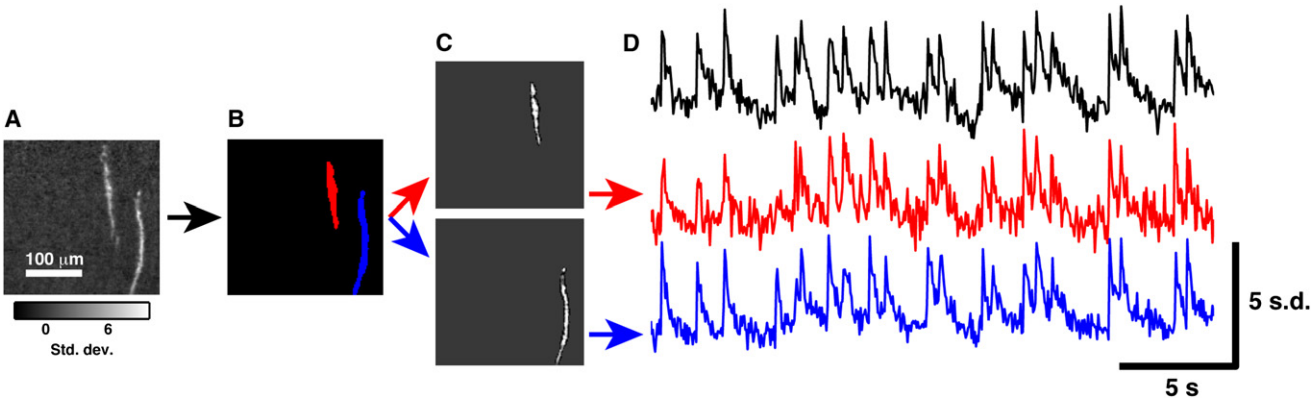

Figure 4. Image Segmentation Improves Extraction of Signals from Correlated Cells

(A) Example of a spatial filter found by ICA in an experimental data set recorded in an awake, head-restrained mouse. This component has two well-separated groups of pixels. These groups likely represent the dendritic trees of two distinct Purkinje cells that receive input from the same olivo-cerebellar climbing fiber.

(B) The first stage of segmentation smooths the spatial filter by convolving with a Gaussian blurring function and applying a threshold to the result to create a binary mask that marks the regions contributing strongly to the signal. Each contiguous image segment is assigned to its own 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 in 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 into distinct contributions from each cell (red and blue). Segmentation reveals that the activities of the two cells are strongly correlated, as expected for cells found together by ICA. In this case, the spike trains from the two image segments are identical except for one spike.

Supplemental Data). By comparison, ROI analysis did not improve much with increases in signal-to-noise ratio and failed to separate overlapping cells (Figure 3E). Alternatively, when we distributed a fixed number of pixels across variable field of view areas, ICA achieved high fidelities and low crosstalk across a broad range of field sizes, whereas ROI analysis steadily degraded as increased overlap between cells led to greater crosstalk (Figure 3F). With the largest fields of view, both methods failed, yielding fidelities scarcely better than the crosstalk values. Overall, ICA was superior to ROI analysis and much closer to performance limits set by linear regression (Figure 3D).

Image Segmentation for Separating Correlated Signals

An advantage of imaging is its ability to sample dense networks of cells, since the dynamical correlations between neighboring cells are often of prime interest. A key question is whether ICA, which relies on statistical independence to identify cells, is limited in its ability to separate cells with correlated signals?

In simulations with varying degrees of pairwise Purkinje cell synchrony, we found that ICA could readily distinguish cells whose spike trains had correlation coefficients, r, as high as 0.8 at the spike ($\leq$1 Hz) and frame (10 Hz) rates simulated. The correlated cells were far from independent, but sorting performed well, yielding high-fidelity estimates comparable to those attained without correlations (r = 0). This was because activity traces of individual cells were still more skew than those from cell mixtures, consistent with prior reports that ICA is often robust to deviations from the underlying model assumptions (Hyvarinen, 1999). Still, when $r > 0.8$, ICA often extracted signals from strongly correlated cells in a single independent component.

Our sorting procedure thus augments ICA with an image segmentation step to disentangle signals from highly correlated cells that are spatially separated (Figures 1B and 4). Following ICA, we smoothed each component's spatial filter and applied a binary threshold to find local regions with strong signal contributions. If a filter contained more than one such region, we created new filters, each of which contained only one of the image segments. These steps separated distinct cells with strongly correlated activity (Figure 4). After testing our procedures on simulated data, we studied data from the cerebellar cortex of live mice.

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

To illustrate cell sorting in an experimental context, we studied data from the molecular layer of cerebellar lobules V and VI of anesthetized and awake head-fixed mice. Using spatiotemporal ICA with μ = 0.1–0.2 followed by image segmentation, we extracted filters with shapes of Purkinje cell dendrites that showed spontaneous (Figures 5A and 5B) and evoked Ca^{2+} spiking activity (Figures 6 and S5).

There were also filters with more isotropic profiles and slower dynamics representing Ca^{2+} activation of Bergmann glial fibers (Figures 5A and 5B) (Nimmerjahn et al., 2009). Since Bergmann glial fibers entwine Purkinje cell dendrites (Grosche et al., 1999), many pixels contained the activity of both cell types. Thus, it is challenging to separate Purkinje cell and Bergmann glial signals cleanly by ROI analysis. Yet, by automated sorting we disentangled Purkinje cell Ca^{2+} spikes (Figure 5B, top four traces) from Bergmann glial Ca^{2+} transients (Figure 5B, bottom four traces) without noticeable crosstalk. Spontaneous glial signals found by cell sorting covered ellipsoidal areas and had event rates much lower 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 ± 0.1, mean ± SD; n = 96 events in 11 mice) to a double-exponential function with similar

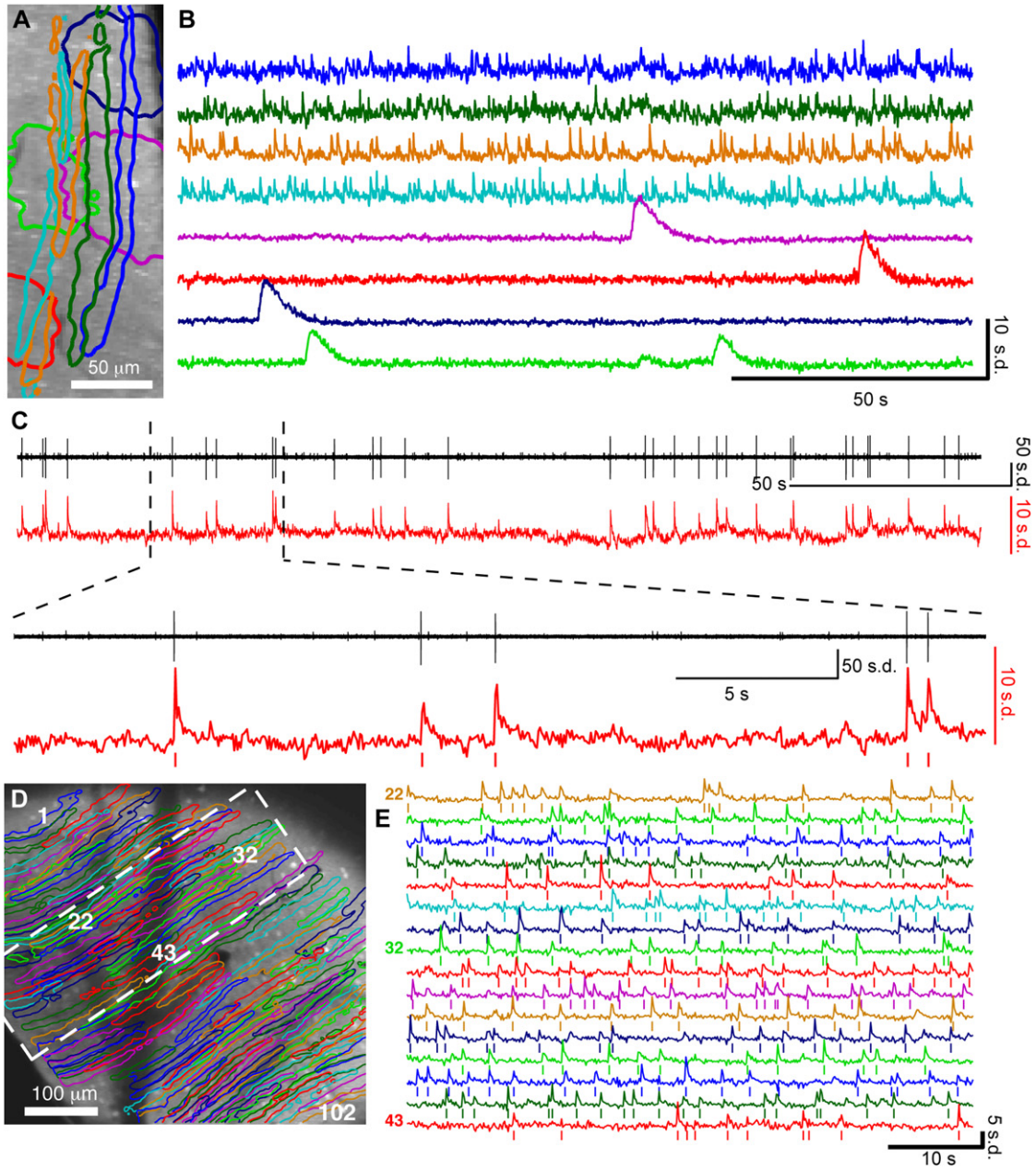

Figure 5. Automated Cell Sorting Identifies Neuronal and Glial Ca^{2+} Dynamics from Large-Scale Two-Photon Imaging Data

(A) Contours of four spatial filters corresponding to Purkinje cell dendrites identified by automated cell sorting, as well as four independent components characteristic of Bergmann glial cells' activity. Contours are superimposed on an image of the background fluorescence. The spatial filters partially overlap, with many pixels sharing signals from both neurons and glia.

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

(C) Single-unit electrical recording from a Purkinje cell (black traces) and the corresponding signal extracted by cell sorting (red traces) from simultaneously recorded Ca^{2+} imaging data. Dashed lines mark the interval in the top two traces over which the data is replotted in the bottom two traces. Red tick marks indicate the times of estimated Ca^{2+} spikes, each identified as the occurrence of a positive-going threshold crossing in the activity traces, following a temporal deconvolution to correct for the dye's Ca^{2+} binding kinetics.

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

(E) Example time courses of 16 of the independent-component signals identified in the outlined region in (D) containing cells 22–43. Tick marks represent the times of Ca^{2+} spikes, estimated as in (C).

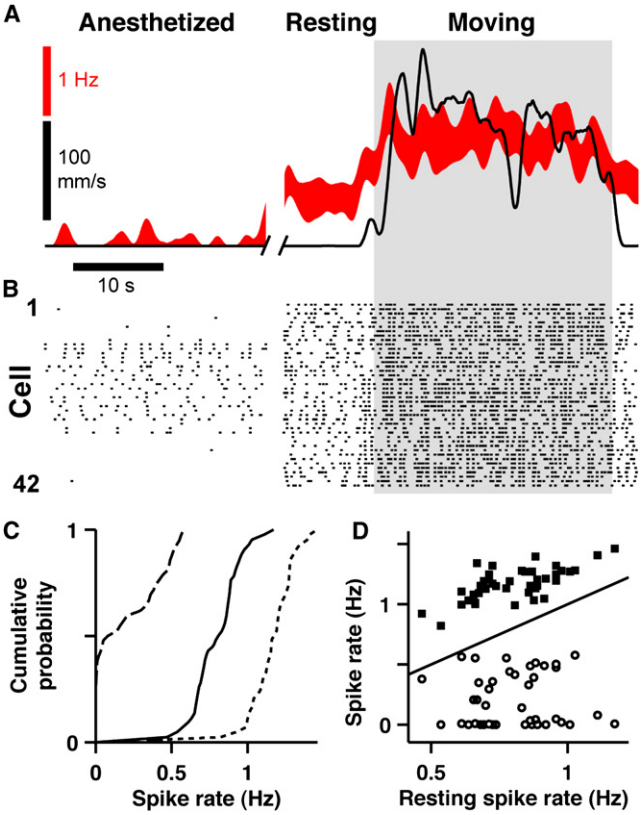

Figure 6. Ca^{2+} Spiking in Purkinje Cells of the Cerebellar Vermis Depends on Behavioral State

(A) Average Ca^{2+} spike rate (red; median ± SD) in 42 Purkinje cell dendrites identified by cell sorting from one mouse under different physiological conditions. Left, isoflurane-anesthetized. Center, alert but resting. Right shaded period, actively moving. Black trace shows the mouse's running speed on the ball.

(B) Ca^{2+} spike rasters for each of the cells in the data set of (A). Under anesthesia only a subset of Purkinje cells found in the analysis of the complete data set 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 versus resting spike rate on the abscissa. The diagonal line delineates equal spike rates under both conditions. As in (C), spike rates are generally higher in awake than in anesthetized animals and rise further during locomotion.

time constants for the rise (1.5 ± 0.7 s) and decay (1.8 ± 1.2 s). These kinetics resembled those of Ca^{2+} "bursts" seen previously in Bergmann glia in cerebellar slices (Beierlein and Regehr, 2006; Piet and Jahr, 2007), anesthetized (Hoogland et al., 2009) and awake mice (Nimmerjahn et al., 2009).

After extracting Purkinje cells' activities, we estimated the cells' Ca^{2+} spike trains in binary format using a deconvolution and threshold approach (Yaksi and Friedrich, 2006) (Figure 1B). To quantify spike-detection accuracies, we combined two-photon microscopy with simultaneous extracellular single-unit electrical recordings of Purkinje cell activity in anesthetized mice (Figure 5C). On the assumption that the electrophysiological traces yielded perfect records of complex spiking, we tested our ability to extract complex spikes from the optical data. (This assumption is in general false, as imperfect electrical traces could lead to misestimates of the accuracy of optical spike detection.) We used receiver operating characteristic (ROC) analysis, which compares probabilities of correct and incorrect spike identification (Figure S2) (Fawcett, 2006).

These probabilities depend on the spike-detection threshold, and the ROC curve summarizes the tradeoff between sensitivity and accuracy in a two-dimensional plot (Figure S2). The area under this curve equals the probability that an ideal classification algorithm would correctly discriminate a randomly selected time frame with a spike from one without a spike (Fawcett, 2006), a measure not dependent on the spike threshold. In our data, areas under the ROC curves were 0.84 ± 0.06 (mean ± SD; n = 7 cells) for image segments and 0.92 ± 0.05 for raw ICA signals. This implies ~85%–90% discrimination accuracy. False positive and false negative spike-detection rates (Figure S2) were comparable to those in recent imaging studies (Ozden et al., 2008; Sasaki et al., 2008) and tetrode recordings (Harris et al., 2000).

Following these validation studies in anesthetized mice, we studied awake mice and extracted up to ~100 Purkinje cells from fields up to ~500 μm wide (Figures 5D and 5E). Our algorithm extracted spike trains from cells tiling the entire field of view, many of which were closely adjacent and extremely difficult to separate by manual methods. Automated cell sorting also separated data components that seemed to represent movement artifacts (Figure S6), similarly to ICA's isolation of movement artifacts in fMRI studies (Beckmann and Smith, 2004; McKeown et al., 1998). This meant that our estimates of cellular activity were relatively uncorrupted by animal motion, showing that ICA can complement image-registration techniques for removal of motion artifacts. As a test of ICA's robustness, we compared cells identified across an entire data set to those identified when only periods of the mouse running were used for analysis. The latter analysis identified ~50% (range, 20%–74%; 5 mice) of the cells from the full data set, but these cells appeared to be correctly identified since they were found in both cases. This test is extreme, for it assumes that an entire experiment involves continuous running, and illustrates ICA's ability to find cells despite the presence of motion artifacts.

Locomotor Behavior Increases Purkinje Cells' Complex Spike Rates

A longstanding goal has been to determine how activity in the olivo-cerebellar circuit is modulated during motor behavior. By combining automated cell sorting and two-photon imaging in behaving mice, we examined relationships between Purkinje cell complex spiking and locomotion.

Using automated 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 in all three conditions, spike rates rose from 0.48 ± 0.27 Hz (mean ± SD, 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

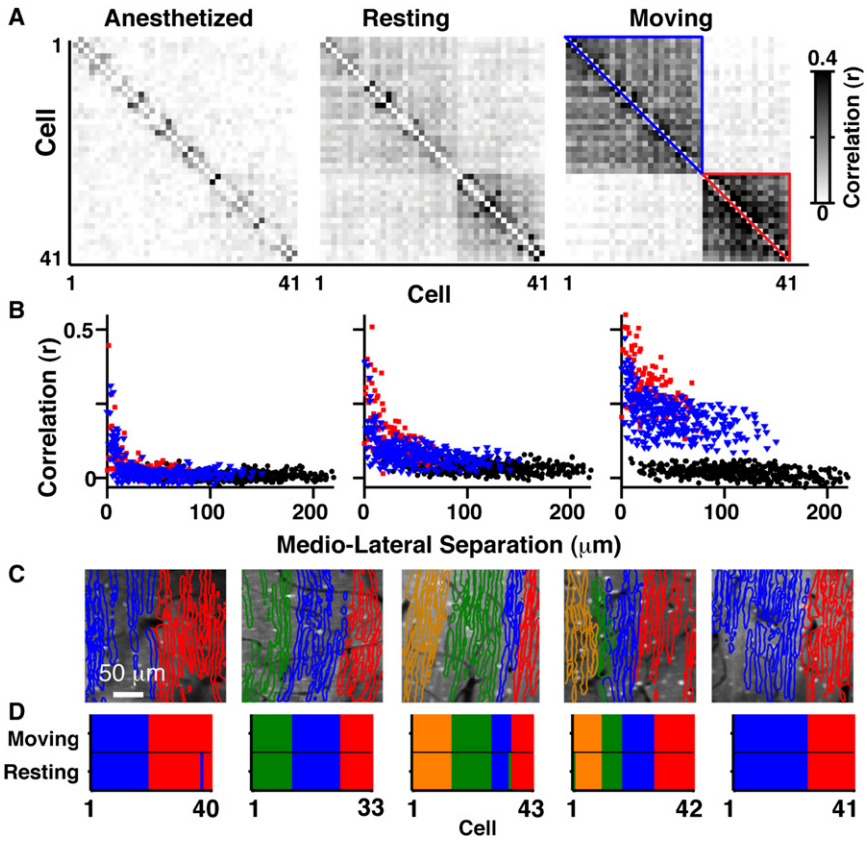

Figure 7. Cerebellar Microzones Exhibit Changes in Correlated Activity but Stable Anatomical Organization across Different Behavioral States

(A) Matrices of correlation coefficients of the spike trains recorded from each pair of cells in a population of 41 Purkinje cell dendritic trees. Cells are arranged by location in the mediolateral dimension. Correlation matrices are computed separately for three 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, red) marked in (A) and in different (black) microzones as a function of the cells' mediolateral separation for each behavioral state.

(C) Outlines of Purkinje cell dendrites identified by automated cell sorting superimposed on the average fluorescence image in five example mice from our set of >50 experiments. Colors show the microzones identified by cluster analysis respect the mediolateral ordering of cells and have sharp boundaries. The example of (A) and (B) is at far right.

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

signed rank test). Thus, gross rates of complex spiking depended on behavioral state, consistent with recent optical studies (Flusberg et al., 2008). Prior electrophysiological studies of locomotion in decerebrated cats have been ambiguous on this point (Andersson and Armstrong, 1987).

In mice studied in both awake and isoflurane-anesthetized conditions, we occasionally saw adjoining groups of Purkinje cell dendritic trees that were identified in the full data set but had little or no complex spiking (rate < 0.01 Hz) during anesthesia. For example, in Figure 6B, dendritic trees labeled 1–8 and 35–42 by mediolateral 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 cessation of anesthesia, all Purkinje cells (1–42) in view resumed Ca^{2+} spiking (Figure 6B). This example highlights the combined utility of automated sorting and Ca^{2+} imaging by revealing phenomena in dense cellular networks inaccessible to electrical recordings.

Cerebellar Microzones Remain Anatomically Stable across Behavioral States

Our studies of the vermis in behaving mice allowed us to reexamine the organization of parasagittal bands of correlated Purkinje cells noted previously by Ca^{2+} imaging in anesthetized animals (Ozden et al., 2008) and electrophysiological means (Andersson and Oscarsson, 1978; Lang et al., 1999). We examined whether microzones have stable anatomical boundaries (Andersson and Oscarsson, 1978) or represent flexible ensembles that vary across behavioral states (Welsh et al., 1995).

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 pairwise correlation coefficients (Figures 7C and 7D) (Ozden et al., 2008). In alert but resting mice, cluster analysis partitioned cells into microzones with higher correlations for intrazone pairs (r = 0.10 ± 0.08, mean ± SD; n = 1418 pairs in 5 mice) than interzone 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 in one cell width rather than in a gradual manner over multiple Purkinje cell dendritic trees.

Like spike rates, pairwise correlations between Ca^{2+} spike trains varied between anesthetized, alert but resting, and actively moving states. The correlations we saw among cells in each microzone grew stronger during active movement ($p < 0.001$ Wilcoxon signed rank test) for intrazone cell pairs (r = 0.20 ± 0.09), and to a statistically significant but far lesser extent for interzone pairs (r = 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 interzone

cell pairs shows there are different degrees of correlated input to intrazone versus interzone pairs and that the rise in synchrony is likely not just due to increased spike rates in independent cells. Likewise, it would be hard to argue that the rise in measured spike rates during locomotion was due to motion artifacts, since motion artifacts would not lead to precisely defined microzones of high pairwise synchrony.

Compared to rest or locomotion, isoflurane anesthesia reduced correlations dramatically (r = 0.03 ± 0.06 intrazone; r = 0.001 ± 0.02 interzone). In each microzone, 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 Ca^{2+} spike rates and pairwise synchrony within microzones increased during active movement as compared to alert rest or anesthesia.

The organization of correlated Purkinje cell complex spiking was earlier proposed to be dynamically modulated during motor behavior (Welsh et al., 1995). 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 whether microzones changed their organization across different behaviors. By comparing microzones found during locomotion and during alert rest, we found that 98.5% ± 0.8% (mean ± 68% confidence interval; n = 199 cells in 5 mice) of cells fell in the same microzone in the two conditions (Figure 7D). Our results did not reveal a modulation of microzone boundaries.

DISCUSSION

Large-scale Ca^{2+} imaging studies pose analysis challenges of three main categories: identification of cells' locations, extraction of Ca^{2+} signals, and detection of neuronal spikes. Only the third problem has received much attention, and several methods exist for detecting spikes in fluorescence traces using temporal deconvolution (Yaksi and Friedrich, 2006), template matching (Greenberg et al., 2008; Kerr et al., 2005), particle filters (Vogelstein et al., 2009), or machine learning (Sasaki et al., 2008). However, these algorithms rely on the assumption that the first two challenges have already been addressed, yielding a paired spatial filter and time trace for each cell. To attain these pairs, we developed automated procedures based on general principles that permit a decomposition of data into constituent independent signals.

Growing Need for Automated Cell Sorting

Ca^{2+} imaging data has commonly been analyzed by manual identification of cell bodies based on their morphologies in static fluorescence images (Dombeck et al., 2007; Gobel et al., 2007; Greenberg et al., 2008). Some studies have used semiautomated methods that rely on morphological filters to find cellular boundaries (Ohki et al., 2005). After finding structures with the sizes and shapes expected of cell bodies, these procedures define ROIs over which the fluorescence is averaged to extract each cell's dynamics. 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 cannot be readily delineated by morphological filtering. In areas such as neocortex, in which the neuropil displays Ca^{2+} activity (Kerr et al., 2005) that inherently contaminates somatic Ca^{2+} signals (Gobel and Helmchen, 2007a; Kerr and Denk, 2008), ROI and morphological analyses can have difficulty excluding neuropil signals from the cellular activity traces (Figure S4). For studies of cells' receptive fields or sensory tuning curves, crosstalk from neuropil has the potential to mask sharply tuned responses, since neuropil can exhibit untuned or broadly tuned signals (Ohki et al., 2005). Given these challenges, sorting approaches based on signal statistics are warranted.

A recent study extracted Ca^{2+} signals from Purkinje cell dendrites by a method using temporal cross-correlations to find sets of pixels contributing to each cell's signal (Ozden et al., 2008). A user selects an area slightly larger than each dendrite and computes the cross-correlations among all pixels in this region. The final ROI contains all pixels highly correlated with many others in the selected area. This method used signal statistics to localize cells but required 2–3 hr of manual work for each data set. As the number of cells in Ca^{2+} imaging studies extends into the hundreds (Gobel et al., 2007; Ohki et al., 2005) or beyond, human selection of pixels will become increasingly prohibitive. Our procedure takes only a few minutes of user supervision to input the number of principal components and screen the results. We habitually scanned the independent components by eye, but for most data sets the selection of independent components representing cells could be automated by cluster analyses (Figures S1 and S6), especially with anesthetized mice, which had lower levels of brain displacement than behaving mice. To be cautious, users should compare the results of automated cell sorting to other data, such as from electrophysiological (Figure 5) or anatomical analyses (Nimmerjahn et al., 2009). We also recommend an initial, systematic testing of a range of μ values, to optimize the ability of spatiotemporal ICA to find cells.

Our use of ICA builds on its prior usage for analysis of optical data of other types, such as for extraction of voltage signals acquired in sea slugs by a photodiode array (Brown et al., 2001). Two studies have used PCA followed by ICA to find functional domains and hemodynamic signatures in optical imaging data (Reidl et al., 2007; Siegel et al., 2007). Signals came from large regions, and so often only a few (e.g., ~6) principal components were retained (Reidl et al., 2007). ICA has been used several times to analyze human brain activity (Beckmann and Smith, 2004; Guimaraes et al., 2007; Makeig et al., 1997; McKeown et al., 1998).

The approach here builds on and differs from previous optical studies by combining spatial and temporal statistics and by following ICA with image segmentation. We preceded ICA with PCA, to reduce the dimensionality of the data and to help ICA find global optima. PCA is not the only means of dimensional

reduction, and other approaches, such as dividing the field of view into subsets each of which is larger than individual cells, might provide viable alternatives. Our overall procedure is grounded in three suppositions: (1) cellular signals are mathematically separable into products of paired spatial and temporal components; (2) signals from different cells are statistically independent; (3) cells' spatial filters and temporal signals have skewed distributions. Interestingly, our procedure proved effective under conditions with modest deviations from these assumptions.

Cell Sorting Separates Data into Spatial and Temporal Components

The supposition that cellular signals are separable into spatial filters and time courses is based on the observation that neuronal Ca^{2+} signals in large-scale imaging data generally arise from fixed locations and do not convey details at the shortest physiological timescales regarding intracellular propagation of [Ca^{2+}] changes. For example, for Purkinje cell Ca^{2+} spikes, any delay between Ca^{2+} activation in different cell parts is generally briefer than one image frame (50–100 ms), permitting a separation of space and time in describing these events. Astrocytic Ca^{2+} waves that propagate far more slowly than neuronal Ca^{2+} spikes violate the assumption of separability (Fiacco and McCarthy, 2006). Still, our procedure extracted a useful approximation of Ca^{2+} waves that expanded over a fraction of the field of view (Figures 5A and 5B). This suggests that ICA remains a useful tool for analyzing weakly nonseparable signals. For forms of glial Ca^{2+} activation that propagate across long distances (Nimmerjahn et al., 2009), other analyses seem more suitable.

Spatiotemporal separability may also be violated due to movement artifacts. Still, ICA can often separate components representing motion artifacts from physiological signals (Figure S6). This agrees with analyses of fMRI data showing that ICA can remove moderate levels of subject motion (McKeown et al., 1998; Reidl et al., 2007). ICA cannot rescue data badly corrupted by motion. When motion artifacts were present at an intermediate level, the skewness of components representing motion were sometimes comparable to those representing cells, but motion components could still be identified based on their having kinetics inconsistent with those of the Ca^{2+} indicator (Figure S6).

Automated Cell Sorting Is Robust to Correlations in Cells' Activities

For cell pairs with modest correlations ($r < 0.8$), ICA was able to identify the individual cells. Since no approach to cell sorting can separate completely synchronized cells without considering geometric information, we used image segmentation following ICA. The result is a robust, hybrid procedure that handled strong levels of correlation ($r \sim 0.9$) in simulations and separated highly correlated dendrites in our cerebellar data that clearly belonged to different Purkinje cells based on their anatomical separation (Figure 4). High levels of complex spike synchrony are known to arise for distinct Purkinje cells with inputs from the same climbing fiber (Sugihara, 2005). Overall, our method's ability to find individual cells in cases of weak or strong synchrony is a key aspect of the procedure's utility.

Cell Sorting Based on Statistical Skewness

An assumption of spatial skewness is valid when individual cells occupy only a small fraction of the pixels in the field of view. In our studies of Purkinje cells, an assumption of skew temporal statistics was also satisfied, since the mean time between Ca^{2+} spikes was greater than the individual spikes' durations. However, sparse activation in time is not the only means of achieving a skewed distribution of signal amplitudes. Cells with high rates of activity that are strongly modulated in time, such as during bursts, could also have skewed distributions of fluorescence amplitudes. Thus, our approach should be well suited to neurons that exhibit irregular spiking, such as high-frequency bursts interrupted by periods of quiescence (Dombeck et al., 2007; Greenberg et al., 2008; Margrie et al., 2002; Orger et al., 2008). Both PCA and ICA ignore cells with very low activity levels, so our method fails to identify cells that are totally inactive in the data. This could lead to underestimation of cell densities and overestimation of mean activity levels in areas where many cells are largely silent (Greenberg et al., 2008). Small fields of view aided the identification of cells with low activity levels (Figure S4), so with such cells it might be beneficial to reduce the dimensionality of the raw data by subdividing the field of view, rather than by PCA.

High rates of tonic, regular Ca^{2+} activation will also lead to signal distributions that are not skew, potentially leading ICA to ignore such cells. In our recordings, we had difficulty extracting cerebellar interneurons (Sullivan et al., 2005), 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, a morphological analysis might work at the risk of increasing crosstalk.

Cerebellar Microzones as Stable Anatomical Entities

Automated cell sorting allowed us to reexamine a debate on the 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 (Llinas, 1991). Microelectrode array recordings of complex spiking in behaving rats suggested that different groups of Purkinje cells were synchronously activated at different phases of movement (Welsh et al., 1995). An opposing view holds that cerebellar microzones provide a stable architecture for controlling different parts of the body and are defined by the anatomical organization of the climbing fibers that drive the complex spikes (Apps and Garwicz, 2005). Mapping studies have shown that strips of neighboring Purkinje cells receive climbing fiber inputs with similar cutaneous receptive fields (Andersson and Oscarsson, 1978; Jorntell et al., 2000) and that these maps are reproducible across subjects (Ekerot et al., 1991). These data did not adjudicate whether adjacent microzones are precisely demarcated from one another or transition smoothly.

By imaging the Ca^{2+} spikes that represent the Ca^{2+} component of Purkinje cells' complex (Na^{+} and Ca^{2+}) spikes (Figure 5) (Flusberg et al., 2008; Ozden et al., 2008), we found that microzones' boundaries are sharply delineated in awake animals, to about the width of one Purkinje cell (Figure 7). Precise

microzones were also seen by Ca^{2+} imaging in anesthetized rodents (Ozden et al., 2008). Our results complement prior studies in awake animals using electrode arrays, which showed that complex spiking synchrony exists in bands extending millimeter-sized distances in the rostral-caudal dimension (Lang et al., 1999; Welsh et al., 1995). The 250 μm spacing between electrodes used in those studies precluded fine resolution of microzone structure in the mediolateral direction.

By extracting Ca^{2+} spike trains from large populations of Purkinje cells during locomotion and alert rest, we studied the stability of microzones. Microzone boundaries and cellular compositions were stable across the two behavioral states examined. During locomotion, Ca^{2+} spike rates rose and levels of synchronous activation increased within microzones but much less so between microzones. Thus, microzones not only retained stable boundaries across different behaviors but also increased in prominence during motor behavior. The discrepancy in synchrony between pairs of cells within and across microzones, as well as the sharpness of microzones' boundaries, indicates that our results are not due to noise, which would not lead to precisely defined regions of high pairwise synchrony. Our results are consistent with data gathered by high-speed imaging of Purkinje cell Ca^{2+} spikes in freely behaving mice, which revealed increases in spike rates and pairwise synchrony during locomotion at the level of population statistics (Flusberg et al., 2008).

Our findings support the view that microzones are stable anatomical entities, for we obtained no evidence of their reorganization during motor behavior. Still, we only explored a single motor behavior and others will need to be examined. The evidence for dynamical reconfiguration of Purkinje cell assemblies (Fukuda et al., 2001; Welsh et al., 1995) should not be discounted, since the reported assemblies of cells were from larger brain areas than those studied here. An attractive possibility reconciling the two viewpoints is that correlated assemblies of different microzones, each stably defined, dynamically reconfigure in different behaviors.

Our data showed that both pairwise correlations and spike rates were higher in awake than anesthetized animals. These results contrast with recent studies of neocortex. In rat visual cortex, correlations between layer 2/3 neurons fell during periods of activity as compared to rest (Greenberg et al., 2008). Further, ketamine/xylazine anesthesia reduced spike rates but increased correlations in neocortical neurons (Greenberg et al., 2008), which contrasts with our data on Purkinje cell correlations under isoflurane anesthesia. Ketatmine/xylazine-anesthetized rats also exhibited significant correlations in Purkinje cell Ca^{2+} spiking (Ozden et al., 2008), suggesting that different anesthetics might have distinct effects on Purkinje cells' synchrony.

Technological Outlook

Refinement of automated sorting techniques should further increase the utility of optical imaging for studying dense cellular networks and allow cell sorting to assume a comparably important role as in electrophysiology. By introducing an automated method based on broadly applicable principles, we have created a framework for moving beyond heuristic and semiautomated approaches. Future refinements might incorporate a priori knowledge of various types or sparseness measures other than skewness. Other generalizations might add 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 Ca^{2+} imaging data using *ImageJ* plug-ins and custom MATLAB routines. Movies were corrected for lateral movement artifacts using *TurboReg* (Thevenaz et al., 1998). We then applied our cell sorting protocol comprising (1) PCA for dimensional reduction, (2) spatiotemporal ICA for extraction of Ca^{2+} signals, (3) image segmentation to separate highly correlated cells, and, in the case of neuronal signals, (4) temporal deconvolution and spike detection to extract spike times. See Supplemental Toolbox.

Artificial Data

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 in the mouse cerebellar molecular layer. Each simulation had 1000 time frames at 10 Hz. We added a static image to the dynamic Ca^{2+} signals to represent background fluorescence from dye-labeled interneurons, as well as unlabeled regions representing blood vessels. The dynamic and background signals defined a noiseless data set. We then introduced Poisson-distributed photon shot noise. See Supplemental Data.

Animal Procedures

Animal procedures were approved by the Stanford Administrative Panel on Laboratory Animal Care. We used male C57Bl/6 wild-type mice (5.5–15 weeks old). Each experiment had two surgeries. Several days before recordings, a custom metal head plate was attached to the skull with dental acrylic. This allowed habituation of the mice to head restraint while walking on the exercise ball and stable imaging under anesthetized or awake conditions. On the recording day, a craniotomy was opened over the cerebellar vermis or neocortex. Two-photon imaging in head-restrained mice was performed as in Nimmerjahn et al. (2009) (Supplemental Experimental Procedures).

Two-Photon Imaging

We loaded cortical tissue with the fluorescent Ca^{2+} indicator Oregon Green 488 BAPTA-1-AM (OGB-1-AM; Molecular Probes) as described (Nimmerjahn et al., 2009). We used a custom two-photon microscope equipped with an ultra-short pulsed Ti:sapphire laser tuned to 800 nm. See Supplemental Data for details.

Analysis of Cerebellar Microzones

We analyzed spiking correlations among all Purkinje cells during anesthetized, awake but resting, and actively moving conditions. We used the Pearson correlation coefficient to compare the binary Ca^{2+} spike trains for each pair of cells. We grouped cells with similar spike trains using k-means clustering (MacKay, 2003).

SUPPLEMENTAL DATA

Supplemental Data include Supplemental Experimental Procedures, six figures, a table, and a MATLAB toolbox and can be found with this article online at http://www.cell.com/neuron/supplemental/S0896-6273(09)00619-9.

ACKNOWLEDGMENTS

We thank David Profitt and Stanford's Varian Machine Shop for expert machining; Andrea Lui for programming assistance and help habituating mice; Quoc-Thang Nguyen and David Kleinfeld for providing imaging software and electronics; Akirah Katoh for consultation on electrophysiological recordings of Purkinje cells; and Shaul Hestrin, Aapo Hyvärinen, Christopher Niell, Carlos Portera-Cailliau, Clay Reid, Krishna Shenoy, Haim Sompolinsky, Michael Stryker, and Byron Yu for helpful correspondence and discussions.

Our work was supported by grants to M.J.S. from ONR, NSF, NINDS, NIDCD, the UBC NIH Nanomedicine Center, the Stanford Bio-X program, and the Klingenstein, Sloan, and Packard Foundations, postdoctoral fellowships to A.N. from the Alexander von Humboldt-Foundation and the International Human Frontier Science Program Organization, and an NSF Graduate Fellowship to E.A.M.

Accepted: August 1, 2009
Published: September 23, 2009

REFERENCES

Andersson, G., and Oscarsson, O. (1978). Climbing fiber microzones in cerebellar vermis and their projection to different groups of cells in the lateral vestibular nucleus. Exp. Brain Res. *32*, 565–579.

Andersson, G., and Armstrong, D.M. (1987). Complex spikes in Purkinje cells in the lateral vermis (b zone) of the cat cerebellum during locomotion. J. Physiol. *385*, 107–134.

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

Batista, A.P., Santhanam, G., Yu, B.M., Ryu, S.I., Afshar, A., and Shenoy, K.V. (2007). Reference frames for reach planning in macaque dorsal premotor cortex. J. Neurophysiol. *98*, 966–983.

Beckmann, C.F., and Smith, S.M. (2004). Probabilistic independent component analysis for functional magnetic resonance imaging. IEEE Trans. Med. Imaging *23*, 137–152.

Beierlein, M., and Regehr, W.G. (2006). Brief bursts of parallel fiber activity trigger calcium signals in bergmann glia. J. Neurosci. *26*, 6958–6967.

Bell, A.J., and Sejnowski, T.J. (1995). An information-maximization approach to blind separation and blind deconvolution. Neural Comput. *7*, 1129–1159.

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

Csicsvari, J., Hirase, H., Czurko, A., and Buzsaki, G. (1998). Reliability and state dependence of pyramidal cell-interneuron synapses in the hippocampus: an ensemble approach in the behaving rat. Neuron *21*, 179–189.

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

Ekerot, C.F., Garwicz, M., and Schouenborg, J. (1991). Topography and nociceptive receptive fields of climbing fibres projecting to the cerebellar anterior lobe in the cat. J. Physiol. *441*, 257–274.

Fawcett, T. (2006). An introduction to ROC analysis. Pattern Recognit. Lett. *27*, 861–874.

Fee, M.S., Mitra, P.P., and Kleinfeld, D. (1996). Automatic sorting of multiple unit neuronal signals in the presence of anisotropic and non-Gaussian variability. J. Neurosci. Methods *69*, 175–188.

Fiacco, T., and McCarthy, K. (2006). Astrocyte calcium elevations: Properties, propagation, and effects on brain signaling. Glia *54*, 676–690.

Flusberg, B.A., Nimmerjahn, A., Cocker, E.D., Mukamel, E.A., Barretto, R.P., Ko, T.H., Burns, L.D., Jung, J.C., and Schnitzer, M.J. (2008). High-speed, miniaturized fluorescence microscopy in freely moving mice. Nat. Methods *5*, 935–938.

Fukuda, M., Yamamoto, T., and Llinas, R. (2001). The isochronic band hypothesis and climbing fibre regulation of motricity: an experimental study. Eur. J. Neurosci. *13*, 315–326.

Gobel, W., and Helmchen, F. (2007a). In vivo calcium imaging of neural network function. Physiology (Bethesda) *22*, 358–365.

Gobel, W., and Helmchen, F. (2007b). New angles on neuronal dendrites in vivo. J. Neurophysiol. *98*, 3770–3779.

Gobel, W., Kampa, B.M., and Helmchen, F. (2007). Imaging cellular network dynamics in three dimensions using fast 3D laser scanning. Nat. Methods *4*, 73–79.

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

Grosche, J., Matyash, V., Möller, T., Verkhratsky, A., Reichenbach, A., and Kettenmann, H. (1999). Microdomains for neuron-glia interaction: parallel fiber signaling to Bergmann glial cells. Nat. Neurosci. *2*, 139–143.

Guimaraes, M.P., Wong, D.K., Uy, E.T., Grosenick, L., and Suppes, P. (2007). Single-trial classification of MEG recordings. IEEE Trans. Biomed. Eng. *54*, 436–443.

Harris, K.D., Henze, D.A., Csicsvari, J., Hirase, H., and Buzsaki, G. (2000). Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements. J. Neurophysiol. *84*, 401–414.

Hastie, T., Tibshirani, R., and Friedman, J.H. (2001). The Elements of Statistical Learning: Data Mining, Inference, and Prediction (New York: Springer), pp. 22.

Helmchen, F., Imoto, K., and Sakmann, B. (1996). Ca2+ buffering and action potential-evoked Ca2+ signaling in dendrites of pyramidal neurons. Biophys. J. *70*, 1069–1081.

Hoogland, T.M., Kuhn, B., Gobel, W., Huang, W., Nakai, J., Helmchen, F., Flint, J., and Wang, S.S. (2009). Radially expanding transglial calcium waves in the intact cerebellum. Proc. Natl. Acad. Sci. USA *106*, 3496–3501.

Hyvarinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. IEEE Trans. Neural Netw. *10*, 626–634.

Hyvarinen, A., and Oja, E. (2000). Independent component analysis: algorithms and applications. Neural Netw. *13*, 411–430.

Jorntell, H., Ekerot, C., Garwicz, M., and Luo, X.L. (2000). Functional organization of climbing fibre projection to the cerebellar anterior lobe of the rat. J. Physiol. *522*, 297–309.

Kerr, J.N., and Denk, W. (2008). Imaging in vivo: watching the brain in action. Nat. Rev. Neurosci. *9*, 195–205.

Kerr, J.N., Greenberg, D., and Helmchen, F. (2005). Imaging input and output of neocortical networks in vivo. Proc. Natl. Acad. Sci. USA *102*, 14063–14068.

Lang, E.J., Sugihara, I., Welsh, J.P., and Llinas, R. (1999). Patterns of spontaneous purkinje cell complex spike activity in the awake rat. J. Neurosci. *19*, 2728–2739.

Lewicki, M.S. (1998). A review of methods for spike sorting: the detection and classification of neural action potentials. Network *9*, R53–R78.

Llinas, R. (1991). The noncontinuous nature of movement execution. In Motor Control: Concepts and Issues, D.R. Humphrey and H.-J. Freund, eds. (New York: Wiley), pp. 223–242.

Llinas, R.R., Walton, K.D., and Lang, E.J. (2004). Cerebellum. In The Synaptic Organization of the Brain, G.M. Shepherd, ed. (Oxford, New York: Oxford University Press), pp. 271–310.

MacKay, D.J.C. (2003). Information Theory, Inference, and Learning Algorithms (Cambridge, UK; New York: Cambridge University Press).

Makeig, S., Jung, T.P., Bell, A.J., Ghahremani, D., and Sejnowski, T.J. (1997). Blind separation of auditory event-related brain responses into independent components. Proc. Natl. Acad. Sci. USA *94*, 10979–10984.

Margrie, T.W., Brecht, M., and Sakmann, B. (2002). In vivo, low-resistance, whole-cell recordings from neurons in the anaesthetized and awake mammalian brain. Pflugers Arch. *444*, 491–498.

McKeown, M.J., Makeig, S., Brown, G.G., Jung, T.P., Kindermann, S.S., Bell, A.J., and Sejnowski, T.J. (1998). Analysis of fMRI data by blind separation into independent spatial components. Hum. Brain Mapp. *6*, 160–188.

Meister, M. (1996). Multineuronal codes in retinal signaling. Proc. Natl. Acad. Sci. USA *93*, 609–614.

Mitra, P.P., and Pesaran, B. (1999). Analysis of dynamic brain imaging data. Biophys. J. *76*, 691–708.

Mrsic-Flogel, T.D., Hofer, S.B., Ohki, K., Reid, R.C., Bonhoeffer, T., and Hubener, M. (2007). Homeostatic regulation of eye-specific responses in visual cortex during ocular dominance plasticity. Neuron *54*, 961–972.

Niell, C.M., and Smith, S.J. (2005). Functional imaging reveals rapid development of visual response properties in the zebrafish tectum. Neuron *45*, 941–951.

Nimmerjahn, A., Mukamel, E.A., and Schnitzer, M.J. (2009). Motor behavior activates Bergmann glial networks. Neuron *62*, 400–412.

Ohki, K., Chung, S., Ch'ng, Y.H., Kara, P., and Reid, R.C. (2005). Functional imaging with cellular resolution reveals precise micro-architecture in visual cortex. Nature *433*, 597–603.

Orger, M.B., Kampff, A.R., Severi, K.E., Bollmann, J.H., and Engert, F. (2008). Control of visually guided behavior by distinct populations of spinal projection neurons. Nat. Neurosci. *11*, 327–333.

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

Palay, S.L., and Chan-Palay, V. (1974). Cerebellar Cortex: Cytology and Organization (Berlin, Heidelberg, New York: Springer).

Piet, R., and Jahr, C.E. (2007). Glutamatergic and purinergic receptor-mediated calcium transients in Bergmann glial cells. J. Neurosci. *27*, 4027–4035.

Reidl, J., Starke, J., Omer, D.B., Grinvald, A., and Spors, H. (2007). Independent component analysis of high-resolution imaging data identifies distinct functional domains. Neuroimage *34*, 94–108.

Sasaki, T., Takahashi, N., Matsuki, N., and Ikegaya, Y. (2008). Fast and accurate detection of action potentials from somatic calcium fluctuations. J. Neurophysiol. *100*, 1668–1676.

Servais, L., Bearzatto, B., Hourez, R., Dan, B., Schiffmann, S.N., and Cheron, G. (2004). Effect of simple spike firing mode on complex spike firing rate and waveform in cerebellar Purkinje cells in non-anesthetized mice. Neurosci. Lett. *367*, 171–176.

Siegel, R.M., Duann, J.R., Jung, T.P., and Sejnowski, T. (2007). Spatiotemporal dynamics of the functional architecture for gain fields in inferior parietal lobule of behaving monkey. Cereb. Cortex *17*, 378–390.

Stone, J.V., Porrill, J., Porter, N.R., and Wilkinson, I.D. (2002). Spatiotemporal independent component analysis of event-related fMRI data using skewed probability density functions. Neuroimage *15*, 407–421.

Stosiek, C., Garaschuk, O., Holthoff, K., and Konnerth, A. (2003). In vivo two-photon calcium imaging of neuronal networks. Proc. Natl. Acad. Sci. USA *100*, 7319–7324.

Sugihara, I. (2005). Microzonal projection and climbing fiber remodeling in single olivocerebellar axons of newborn rats at postnatal days 4-7. J. Comp. Neurol. *487*, 93–106.

Sullivan, M.R., Nimmerjahn, A., Sarkisov, D.V., Helmchen, F., and Wang, S.S. (2005). In vivo calcium imaging of circuit activity in cerebellar cortex. J. Neurophysiol. *94*, 1636–1644.

Thevenaz, P., Ruttimann, U.E., and Unser, M. (1998). A pyramid approach to subpixel registration based on intensity. IEEE Trans. Image Process. *7*, 27–41.

Vogelstein, J.T., Watson, B.O., Packer, A.M., Yuste, R., Jedynak, B., and Paninski, L. (2009). Spike inference from calcium imaging using sequential Monte Carlo methods. Biophys. J. *97*, 636–655.

Welsh, J.P., Lang, E.J., Suglhara, I., and Llinas, R. (1995). Dynamic organization of motor control within the olivocerebellar system. Nature *374*, 453–457.

Yaksi, E., and Friedrich, R.W. (2006). Reconstruction of firing rate changes across neuronal populations by temporally deconvolved Ca2+ imaging. Nat. Methods *3*, 377–383.