You might find this additional information useful...

Supplemental material for this article can be found at:
http://jn.physiology.org/cgi/content/full/01310.2007/DC1

This article cites 31 articles, 16 of which you can access free at:
http://jn.physiology.org/cgi/content/full/100/1/495#BIBL

This article has been cited by 2 other HighWire hosted articles:

Functional Clustering of Neurons in Motor Cortex Determined by Cellular Resolution Imaging in Awake Behaving Mice
D. A. Dombeck, M. S. Graziano and D. W. Tank
J. Neurosci., November 4, 2009; 29 (44): 13751-13760.
[Abstract] [Full Text] [PDF]

Reliable Coding Emerges from Coactivation of Climbing Fibers in Microbands of Cerebellar Purkinje Neurons
I. Ozden, M. R. Sullivan, H. M. Lee and S. S.-H. Wang
J. Neurosci., August 26, 2009; 29 (34): 10463-10473.
[Abstract] [Full Text] [PDF]

Updated information and services including high-resolution figures, can be found at:
http://jn.physiology.org/cgi/content/full/100/1/495

Additional material and information about *Journal of Neurophysiology* can be found at:
http://www.the-aps.org/publications/jn

This information is current as of January 26, 2010 .

Identification and Clustering of Event Patterns From In Vivo Multiphoton Optical Recordings of Neuronal Ensembles

Ilker Ozden,[1,2,]* H. Megan Lee,[2,]* Megan R. Sullivan,[1,2] and Samuel S.-H. Wang[1,2]

[1]*Department of Molecular Biology and* [2]*Princeton Neuroscience Institute, Princeton University, Princeton, New Jersey*

Ozden I, Lee HM, Sullivan MR, Wang SS-H. Identification and clustering of event patterns from in vivo multiphoton optical recordings of neuronal ensembles. *J Neurophysiol* 100: 495–503, 2008. First published May 21, 2008; doi:10.1152/jn.01310.2007. In vivo multiphoton fluorescence microscopy allows imaging of cellular structures in brain tissue to depths of hundreds of micrometers and, when combined with the use of activity-dependent indicator dyes, opens the possibility of observing intact, functioning neural circuitry. We have developed tools for analyzing in vivo multiphoton data sets to identify responding structures and events in single cells as well as patterns of activity within the neural ensemble. Data were analyzed from populations of cerebellar Purkinje cell dendrites, which generate calcium-based complex action potentials. For image segmentation, active dendrites were identified using a correlation-based method to group covarying pixels. Firing events were extracted from dendritic fluorescence signals with a 95% detection rate and an 8% false-positive rate. Because an event that begins in one movie frame is sometimes not detected until the next frame, detection delays were compensated using a likelihood-based correction procedure. To identify groups of dendrites that tended to fire synchronously, a *k*-means-based procedure was developed to analyze pairwise correlations across the population. Because repeated runs of *k*-means often generated dissimilar clusterings, the runs were combined to determine a consensus cluster number and composition. This procedure, termed meta-*k*-means, gave clusterings as good as individual runs of *k*-means, was independent of random initial seeding, and allowed the exclusion of outliers. Our methods should be generally useful for analyzing multicellular activity recordings in a variety of brain structures.

INTRODUCTION

An important step in understanding dynamic principles of brain circuit function is the ability to monitor many neurons at once. Such monitoring has been made possible by in vivo multiphoton microscopy, which when combined with the application of membrane-crossing precursors of calcium indicators (Tsien et al. 1982) to intact brain tissue (Helmchen and Waters 2002; Stosiek et al. 2003), makes it possible to record from many neighboring neural processes within a field of view (Ohki et al. 2005; Sullivan et al. 2005).

Calcium imaging data have superior spatial resolution than electrophysiological recordings: a single neural structure can span many submicron pixels, while electrodes typically record from multiple units at once. However, calcium imaging data are gathered at lower bandwidth (0.01–1 kHz per location) than electrophysiology data (~10 kHz). Consequently, several new challenges in data analysis arise. First, individual structures of interest, whether single neurons or neuronal processes, must be identified by image segmentation or by their activity-dependent fluorescence changes over time. Second, the lower time resolution of frame scanning makes determination of event timing less certain. Third, patterns of activity across the whole ensemble need to be inferred with as little as a few hundred temporal samples per structure because light-induced damage and photobleaching place upper limits on the amount of obtainable useful data.

Here we present methods for analyzing movies obtained through calcium imaging from large groups of neurons. Our target for analysis is complex spike activity in cerebellar Purkinje cells (PCs) in which intercellular synchrony is prominent (Lang et al. 1999). We have developed semiautomated software tools to identify individual active structures, estimate the timing of individual events, and find clusters of structures that are synchronously active. To determine the level of synchrony between neurons more reliably, we present a correction method for event times detected from movies in which the relative timing of events in multiple structures can be incorrectly assigned when the structures are scanned at different moments. Finally we present a clustering method for detecting synchronous subgroups that uses pairwise correlations between dendritic event times and *k*-means, an algorithm widely used in gene expression analysis. By combining many results of *k*-means, we can robustly identify cluster membership and the likely number of clusters within a data set. Until now this clustering problem has not been adequately addressed in multineuronal in vivo recordings.

We demonstrate our methods on recordings from tens of dendrites of PCs, the sole output neurons of the cerebellar cortex. In PCs, dendritic calcium transients are triggered by the activity of climbing fiber (CF) axons (Ross and Werman 1987), which originate from the inferior olive. The CF projection is spatially organized so that each CF innervates many PCs in a thin parasagitally oriented plane, and neighboring inferior olive neurons innervate PCs that are neighboring in the mediolateral direction (Ruigrok and Voogd 2000; Sugihara et al. 2001). Functional structure in the mediolateral direction has not been well investigated because multielectrode methods only sample at intervals of several hundred micrometers. Our analysis uses imaging data taken from all PCs at once to identify synchrony on a fine scale of <100 micrometers.

METHODS

Animal preparation

Experimental procedures were approved by the Princeton University Institutional Animal Care and Use Committee. Mice (P24–P42) were deeply anesthetized with intraperitoneally injected ketamine HCl (80 mg/kg body wt) and xylazine (12 mg/kg body wt). Deep anesthesia was maintained throughout the experiment as confirmed by the lack of a pinch foot-withdrawal reflex and a lack of whisking, and body temperature was maintained at 37°C. A small craniotomy of diameter 2 mm was made over crus IIa. Purkinje neurons and other structures were loaded with calcium indicator Oregon Green bis-(o-aminophenoxy)-N,N,N',N'-tetraacetic acid (BAPTA)-1/AM (Invitrogen, Carlsbad, CA) as described in Sullivan et al. (2005).

Electrophysiology

Extracellular single-unit recordings from PCs were made using glass micropipettes pulled to 6–10 MΩ with a P-2000 puller (Sutter Instrument, Novato, CA) and filled with artificial cerebrospinal fluid containing (in mM) 135 NaCl, 5.4 KCl, 5 NaHEPES, 1 MgCl$_2$, and 2 CaCl$_2$ (pH 7.3 with HCl). Signals were acquired using a NeuroData IR-283A amplifier (Cygnus Technology, Delaware Water Gap, PA), amplified 10–100 times, band-pass filtered at 0.3–10 kHz with a Brownlee Model 440 amplifier (Brownlee Precision, San Jose, CA), and saved to a personal computer equipped with a data-acquisition board (PCI-MIO-16E-4 from National Instruments, Austin, TX) and custom MATLAB software.

Multiphoton laser scanning microscopy

In vivo calcium imaging was performed using a custom-built multiphoton microscope running ScanImage software (B. Sabatini and K. Svoboda). The tissue was illuminated with a pulsed Mai Tai (Spectra-Physics, Mountain View, CA) Ti:sapphire laser (830–840 nm, 80 MHz repetition rate, 100–150 fs pulse width). Excitation light was focused onto tissue using a ×40, 0.8 NA water-immersion IR-Achroplan objective (Carl Zeiss, Thornwood, NY). Emitted light was reflected by a 720-nm long-pass dichroic mirror (750dcxru, Chroma Technology, Rockingham, VT), filtered with green BG39 Schott glass and infrared-blocking filters (Calflex X 380227034, Linos Photonics, Milford, MA), and detected using a photomultiplier tube (H7422P-40, Hamamatsu Photonics, Hamamatsu City, Japan). The movies were collected as 64 × 64 (128 ms/frame) or 128 × 128 (256 ms/frame) pixel frames or 2 ms/line scans.

Data analysis

Data analysis was performed with software written in MATLAB version 7.1 (MathWorks). All code is available at http://synapse.princeton.edu/programs. The correlation between two dendritic signals $x(t)$ and $y(t)$ was calculated as the Pearson correlation

$$ r = \frac{\sum_i x(t) \cdot y(t) - \langle x \rangle \cdot \langle y \rangle \cdot (\Delta t / T)}{\sqrt{[\langle x^2 \rangle - \langle x \rangle^2 \cdot (\Delta t / T)] \cdot [\langle y^2 \rangle - \langle y \rangle^2 \cdot (\Delta t / T)]}} $$

where T is the total recording time and Δt is the sampling time.

RESULTS

Multiphoton recording of complex spikes in PC ensembles

Multiphoton imaging data were obtained from the molecular layer of the cerebellar cortex by bulk-loading the tissue with the calcium indicator Oregon Green BAPTA-1/AM (Fig. 1A). Bulk loading of this dye stains many cellular structures in neuropil, including PCs, interneurons, and glia (Sullivan et al. 2005). After taking up the dye, the cellular structures showed fluorescence intensity changes, which were caused by calcium concentration changes. Fields of view were located at 90–120 μm above the PC layer, where the observed calcium signals were mainly from PC dendrites. PC dendrites are finely branched and fill a flat, fan-like, parasagittally oriented space ~200 μm wide and 6–10 μm thick (Fig. 1A), appearing in transverse section as tube-like structures separated by 4–8 μm (Fig. 1B) (Palay and Chan-Palay 1974). The orientation of the scan pattern was adjusted to align dendrites in the direction of fast scanning, thus establishing a precise within-frame time at which each dendrite was scanned.

Signal-based image segmentation of dendrites

In full-field scans (64 × 64 pixels at 128 ms/frame or 128 × 128 pixels at 256 ms/frame, field of view 135 or 270 μm wide), between 18 and 45 PC dendrites were visible at once (Fig. 1B; Supplementary Movie S1[1]). PC dendrites generated large (>10% $\Delta F/F$) calcium transients between 0.2 and 1 times per second as expected given the fact that CF firing evokes a calcium-based complex action potential (Eccles et al. 1966; Llinás and Sugimori 1980).

To identify each dendrite, average fluorescence intensity was not a sufficient criterion because the cell type specificity of bulk loading was low, leading to poor contrast between neighboring structures. This fact was apparent in our raw data sets (see Supplementary Movie S1), in which individual PC dendrites did not have sharp borders (Fig. 1B, *left*). Therefore to identify dendrites, we exploited the properties of complex spike-evoked calcium transients, which span the dendrite, rise in ~10 ms, and fall with a half-maximal decay time of 0.17 ± 0.05 s, and an exponential decay time constant of 0.28 ± 0.06 (SD) s (29 dendrites in 4 experiments), consistent with previous reports in vivo (Göbel and Helmchen 2007; Sullivan et al. 2005). These slow decay times make it possible to detect individual events even for a frame acquisition time of 256 ms; in the worst-case scenario, an action potential that occurs immediately after a dendrite is scanned will leave a signal in the next frame that is 30% of the peak $\Delta F/F$ amplitude. In addition to this, the low rate of calcium signals (0.48 ± 0.15 Hz) makes it possible to record at low frame rates (4–8 Hz).

Raw movies were high-pass filtered by Fourier-transforming each pixel's trace, removing low-frequency components below 0.1–0.3 Hz, and taking the real part of the inverse Fourier transform. In some cases, drift in average brightness, due largely to photobleaching, was removed by calculating a linear fit for each pixel's fluorescence time course and then subtracting it. The resulting filtered movies (e.g., Supplementary Movie S2) showed conspicuous transients against a relatively dark background. A correlation-based identification algorithm was then used to recognize individual dendrites. First the filtered movie was examined to find candidate dendrites as judged by the co-occurrence of signals in the shape and orientation of PC dendrites and by the criterion that a candidate structure fire by itself (for example, see Fig. 1B, *right*) and all at once. To find the pixels belonging to this dendrite, a region was selected slightly larger than the bright pixels

---

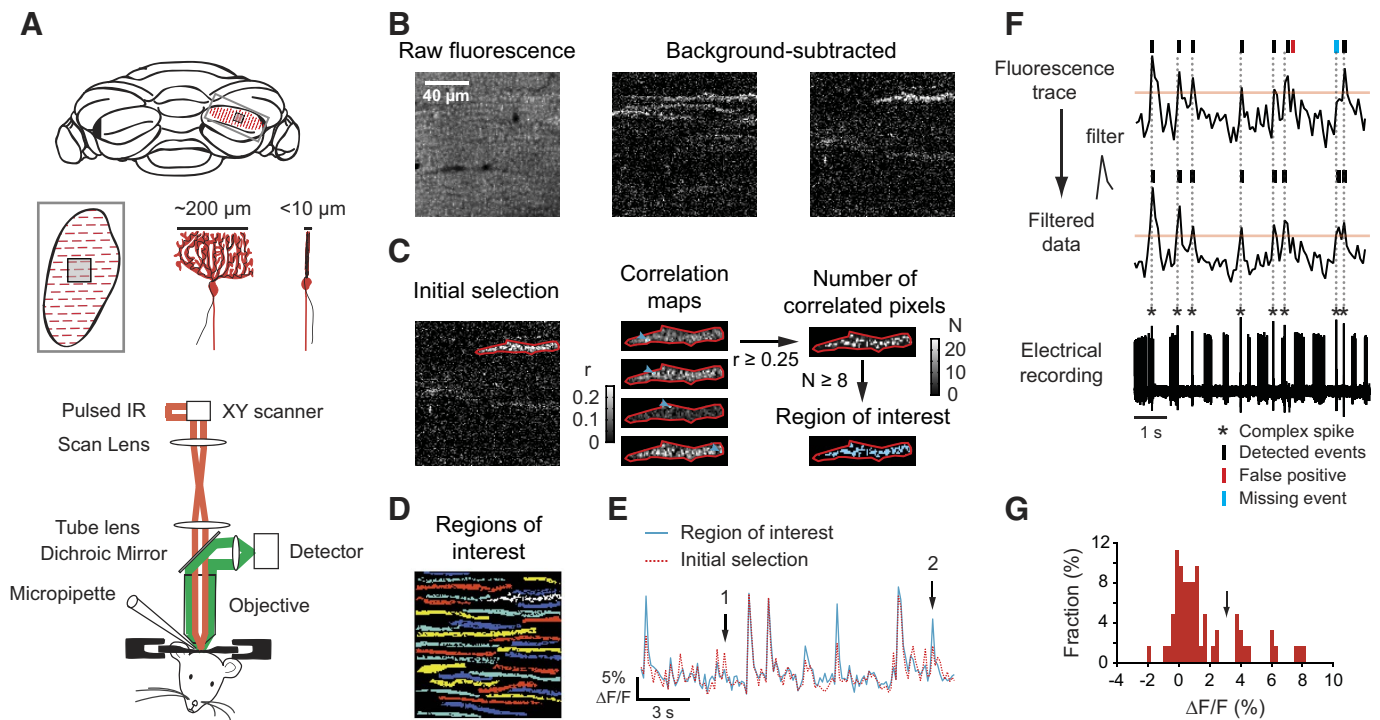[1] The online version of this article contains supplemental data.

FIG. 1. In vivo multiphoton microscopy, identification of dendrites, and event detection. *A, top*: the crus II region of the cerebellum, where loading was done. *Middle*: the orientation and size of Purkinje cell dendrites. Dendritic arbors are planar, parasagittally aligned, and nonoverlapping with one another. *Bottom*: the multiphoton microscope. *B, left*: a typical single movie frame. *Right*: frames from filtered and mean-subtracted movies showing activity in 1 or more dendrites at once. *C*: the dendrite selection procedure for the dendrite in *B*, showing initial selection of a dendrite (*left*), example correlation maps for 4 selected pixels (*middle*), the number of pixels correlated with a particular pixel with $r_{thresh} > 0.25$ (*top right*), and the final dendrite after thresholding the number of correlated pixels with $N_{thresh} = 8$ (*bottom right*). *D*: all selected dendrites, color-coded. The dendrite from *C* is shown in white. *E*: improvement of signal quality by the dendrite selection procedure. *F*: event detection. *Top trace* and raster: dendritic fluorescence recording. Thresholding of unfiltered fluorescence gives detected events corresponding to complex spikes (black bars), a false positive (red bar), and a missed event (blue bar). *Middle trace* and raster: matched filtering eliminates false and missed events. *Bottom*: a simultaneous electrical recording from the cell body showing both complex spikes (stars and vertical dotted lines) and simple spikes. *G*: the peak histogram of the trace after the application of the template. The noise component is approximately Gaussian with a median near 0% $\Delta F/F$. The arrow indicates the threshold position.

(Fig. 1*C*, *left*). In this region, a pixel was included in the final region of interest if it was correlated with at least N_{thresh} other pixels within the region with a correlation value of at least r_{thresh}.

Figure 1*C* demonstrates the procedure on one dendrite. The *left panel* shows the initial region selected as a red contour. For each pixel within this region, the correlation coefficient with each other pixel was calculated. The *middle panel* shows correlation maps for the 4 pixels indicated by cyan arrows, and the *right panel* shows a color-coded map of the number of correlated pixels, which was thresholded to determine final membership (*right panel*, blue pixels). Figure 1*D* shows all dendrites identified from the movie. In cases where a dendrite never generated a signal by itself, the other dendrites were identified first, and then a region was selected around the remaining dendrite and the correlation-based algorithm was applied as usual. Typical threshold values were r_{thresh} ranging from 0.2 to 0.3 depending on noise and activity levels, and N_{thresh} ranging from 4 to 10 pixels, 5–15% of the initial selection. Pixel selection improved the signal quality (Fig. 1*E*) by reducing contributions from other dendrites (arrow 1) while simultaneously increasing amplitudes relative to noise (arrow 2). The signal-to-noise ratio improved (14 dendrites in 6 experiments, $P < 0.02$ by paired *t*-test, from 3.42 ± 0.23 to 3.64 ± 0.29, means $\pm$ SE). In addition, pixel selection led to a moderate, near-significant increase in the true positive rate

from 93 to 95% ($P < 0.1$) and a decrease in the false positive rate from 10 to 8% ($P < 0.1$).

Event identification

After the identification of dendrites, each dendrite's fluorescence trace was converted to units of $\Delta F/F$. The raw fluorescence intensity F for a dendrite was defined as the fluorescence intensity averaged over all pixels for the whole movie. ΔF was calculated for each frame by subtracting F from the average pixel intensity for that frame. Because the calculation of F in this way includes frames in which dendrites were firing, this procedure creates an underestimate of $\Delta F/F$ of ~1–2%. However, the signal-to-noise ratio is not altered in comparison with a more elaborate procedure. Events were found by a template-and-threshold method in which a four-sample-long template was generated by normalizing and averaging the highest 10 peaks in a trace. The template was used as a matched filter by shifting it point by point along the trace and taking a dot product with the signal's change above its minimum value within the template duration.

Calcium transient events were confirmed as reflecting complex spikes by simultaneous extracellular recordings from PC cell bodies (Fig. 1*F*). Recordings also showed sodium-based simple spikes, which could be distinguished by their waveform and high firing rates (20–110 Hz). The success rate of the event

selection algorithm was tested by simultaneous extracellular recording from the same PC. Figure 1*F* shows raw (*top*) and templated (*middle*) fluorescence traces from a dendrite and electrical recording (*bottom*) from near its soma with complex spikes indicated by stars.

The example in Fig. 1*F* shows a false positive event (red bar) not associated with any complex spike as well as a missed complex spike (blue bar). In this example, template filtering removed both errors. To characterize event detection rates, we recorded 817 complex spikes from six cells in five animals. Thresholding of raw traces allowed 92% of complex spikes to be detected with a false-positive rate of 11% of all classified events. After template filtering the true-positive rate increased to 95% ($P < 0.01$, $n = 6$, paired 2-tailed *t*-test) and the false-positive rate dropped to 8% ($P < 0.02$, paired 2-tailed *t*-test).

The template-and-threshold method was used for all subsequent recordings. Events were defined when the filtered signal exceeded a threshold value determined by hand after examination of the trace and its peak-amplitude histogram (Fig. 1*G*). In our data, the standard-deviation fluctuation in dendritic fluorescence was $1.6 \pm 0.4\%$ $\Delta F/F$ (mean $\pm$ SD, 14 dendrites). Amplitude thresholds for events were determined by eye after training users on data sets with accompanying electrophysiological recordings of complex spikes. The resulting $\Delta F/F$ for events classified in this way was $7.0 \pm 2.8\%$ (817 events) in movies with 256-ms frame time, approximately one-third the size of the peak $\Delta F/F$ values observed in line scans ($19.8 \pm 6.9\%$, 29 dendrites in 4 experiments), as expected.

Likelihood-based correction for event times

Because each calcium transient ended in a long tail lasting several tenths of a second, temporal undersampling by microscope frame scanning (128–256 ms/frame) led to the possibility that transients would not be detected until the next scan after the spike had occurred (Fig. 2). For a dendrite aligned in the fast-scan direction, a complex spike during scanning of this frame could occur after the beam had already finished scanning the dendrite, in which case the signal would not be picked up until the next frame. Thus in dendrites closer to the end of the scan (e.g., dendrite 1, Fig. 2*B*, *left*), detection is less likely to occur within-frame.

The frame-lag effect can affect the detection of synchronous spike events. When two dendrites spike at the same time, if the laser scan position is between the two dendrites the detected signal peaks will occur in consecutive frames (Fig. 2*A*). This position-dependent error can be corrected with a likelihood-based approach. If an event is detected in frame i, the a posteriori likelihood that the true event time was during the scan of frame i is $p_i = x/L$, where x is the time from the beginning of the frame for the scan beam to reach the dendrite and L is the time taken to scan the whole frame; the likelihood that the event occurred in the previous frame is $p_{i-1} = 1 - x/L$. These likelihoods suggest a statistical correction of event times in which a detected peak in sample i, which would lead to an uncorrected value of 1 in the event trace, is represented by x/L and the $(I - 1)$th time point is set to $1 - x/L$ (Fig. 2*B*). This correction reduces artifactual, apparent distance-dependent decreases of correlations when dendrites fire together but are far from one another in the field of view.
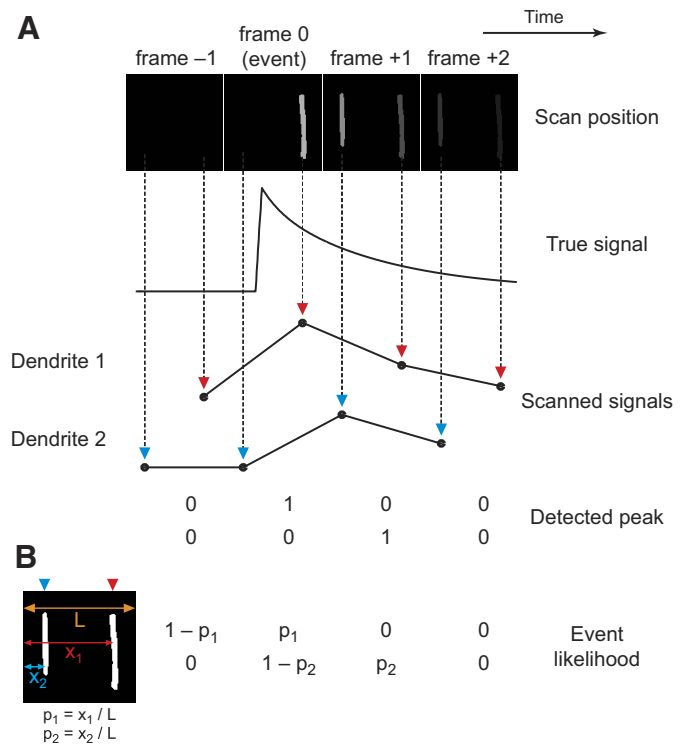

FIG. 2. Likelihood correction of event times. *A*: in standard scan patterns of fluorescence microscopes, imaged structures are scanned in succession during each frame, leading to undersampling of a smoothly varying true signal. The resulting detected peaks can be offset from the frame scan during which the true signal peak occurred. *B*: event times can be corrected by replacing an all-or-none event with 2 values representing the retrospective likelihood that the event peak occurred in the same frame, x/L, or in the previous frame, $1 - x/L$, where x is the time of scanning from the start of the frame and L is the duration of 1 whole frame.

Identification of groups of synchronously firing dendrites

The size and dimensionality of the data set produced by a population recording demand an efficient and automated technique for determining the characteristics of the data set. Because dendrites often fired simultaneously in our data sets, we were interested in grouping them according to this synchrony. However, because the patterns were widespread and varied from event to event, classification by eye was not possible.

We analyzed likelihood-corrected event traces using *k*-means, a clustering technique widely used in gene expression analysis (d'Haeseleer 2005; Le Roch et al. 2003; Tavazoie et al. 1999) to identify clusters within large data sets. The *k*-means clustering algorithm is an iterative, unsupervised learning algorithm that uses a spatial interpretation to uncover cluster structure in high-dimensional data (Hartigan and Wong 1979). The algorithm interprets the data set in a high-dimensional space and iterates to find objects that are near each other. In our case, a movie with M dendrites and N time points was represented as M points in an N-dimensional space, where each point represents the activity of a single dendrite over the entire movie, and the coordinate of the point in each dimension ranges from 0 to 1, representing the activity of the dendrite at one of the N time points. In the $M \times N$ matrix, each row represents one point, the firing pattern of a single dendrite. For example, the activity of the first dendrite is represented by

$$\vec{x}_1 = \{x_{11}, x_{12}, x_{13}, \ldots, x_{1N}\}$$

The distance between dendrites was defined as $d = 1 - r$ where r is the Pearson correlation (see METHODS). Dendrites that have similar activity during the movie have a higher correlation and thus will be closer together in space.

In k-means, for a predetermined number of clusters k, the algorithm first randomly initializes the positions of the k centroids, $\vec{c}_1, \ldots, \vec{c}_k$ (Fig. 3A). Each dendrite is then assigned to the centroid nearest to it, and the positions of the centroids are recomputed using the resulting assigned cluster membership. The algorithm then starts a new cycle of assigning dendrites to centroids, starting with the updated centroid positions. The cycle is repeated until the cluster assignments no longer change.

Determining the number and membership of clusters using meta-k-means

The k-means algorithm and its variations have been popular in gene expression studies, where k may be close to the number of functional categories (Bagirov et al. 2003; Le Roch et al. 2003; Tavazoie et al. 1999). However, unlike gene expression studies, our recordings provided no prior knowledge about the number of clusters.

A plot of the matrix of dendrite-dendrite correlations (Fig. 3C) revealed structure suggesting that dendrites were organized into groups that tended to fire together. Our initial approach to determining the number of clusters was to vary k from 3 to 8 (Fig. 3D). For each value of k, 1,000 runs of k-means were performed. We quantified the goodness of clustering (Maulik and Bandyopadhyay 2002) using Dunn's index (DI) (Fig. 3B), $\text{DI} = D_{\min}/d_{\max}$ where $D_{\min}$ is the smallest distance between two points in different clusters and $d_{\max}$ is the diameter of the widest cluster (Fig. 3B). Like other metrics of clustering goodness (Jain and Dubes 1988), Dunn's index is larger when clusters are tighter and further apart from one another. Dunn's index emphasizes the relationship between adjacent clusters and was appropriate because PC–PC correlations drop with increasing distance (Lang et al. 1999).

Maximization of Dunn's index for each value of k ($\text{DI}_{k,\max}$) yielded a single "best" clustering, usually reaching a local maximum between $k = 2$ and $k = 6$. However, deciding on an appropriate number of clusters k^* was difficult because the local maximum was often broad (Fig. 3D). Also, even for a given value of k^*, values of DI at or near $\text{DI}_{k^*,\max}$ were often associated with quite dissimilar cluster assignments, indicative of multiple locally optimal clusterings. For a single data set, k-means can give different cluster assignments because the outcome is influenced by the initial randomly chosen centroid positions. Although it is also possible for the user to initiate centroid positions manually, this is usually avoided because of the risk of bias and the difficulty of conceptualizing the high-dimensional data space.

Instead we took the approach of aggregating the results of many runs of the k-means algorithm to determine membership more robustly and to estimate the number of clusters. It was desirable for the method to avoid forcing outlier dendrites to be clustered, to allow a dendrite to belong to more than one cluster and to give results that were independent of the randomly initialized cluster centers. The resulting approach, termed meta-k-means, is robust and versatile and produces consistent results when applied to our calcium imaging data sets.

An illustration of the meta-k-means method is shown in Fig. 4. Here we used fluorescence traces obtained from 40 dendrites (Fig. 4, *top row*) imaged in vivo. The detected events are shown as dots above each trace. The first step was to run k-means 1,000 times on the corrected event traces using $k = 3$ and random initialization of the centroid positions on each run, giving several different clusterings (Fig. 4, *3rd row*). For each pair of dendrites i and j, the number of times dendrites i and j were clustered together, counts (i, j), was determined (Fig. 4,
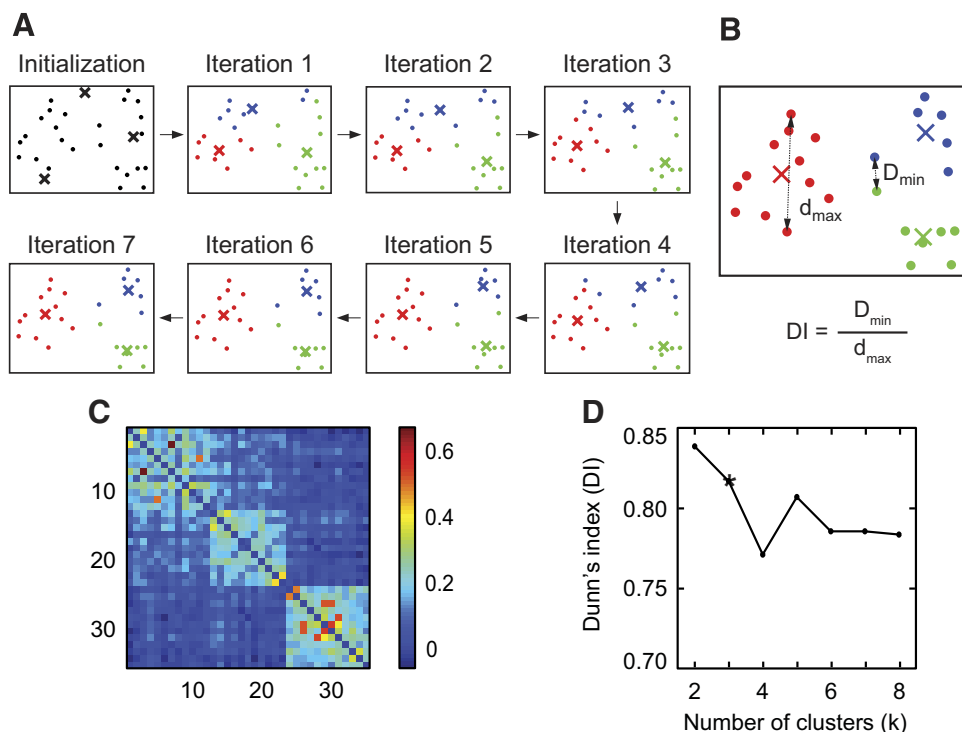

FIG. 3. The k-means clustering algorithm. A: an example showing the determination of 3 clusters in 2-dimensional space using k-means. The cluster centroids (×) and data points (●) are colored to indicate their cluster assignments. Initially, the positions of the centroids are placed randomly and data points are assigned to the nearest centroid, which is recalculated. The cycle is repeated until the cluster assignments no longer change. B: clustering goodness as defined by Dunn's index $D_{\min}/d_{\max}$, where $D_{\min}$ is the smallest distance between 2 points in different clusters and $d_{\max}$ is the widest cluster diameter in the dataset. C: color-coded correlation plot for dendritic activity. The axes labels represent dendrites. D: plot of the largest value of Dunn's index found for many individual runs at a given value of k for dendrites used to create C. This example shows that the local maximum had shallow dependence on k or disagreed with a pattern that was visible in the pairwise correlation matrix. *, the result of meta-k-means.
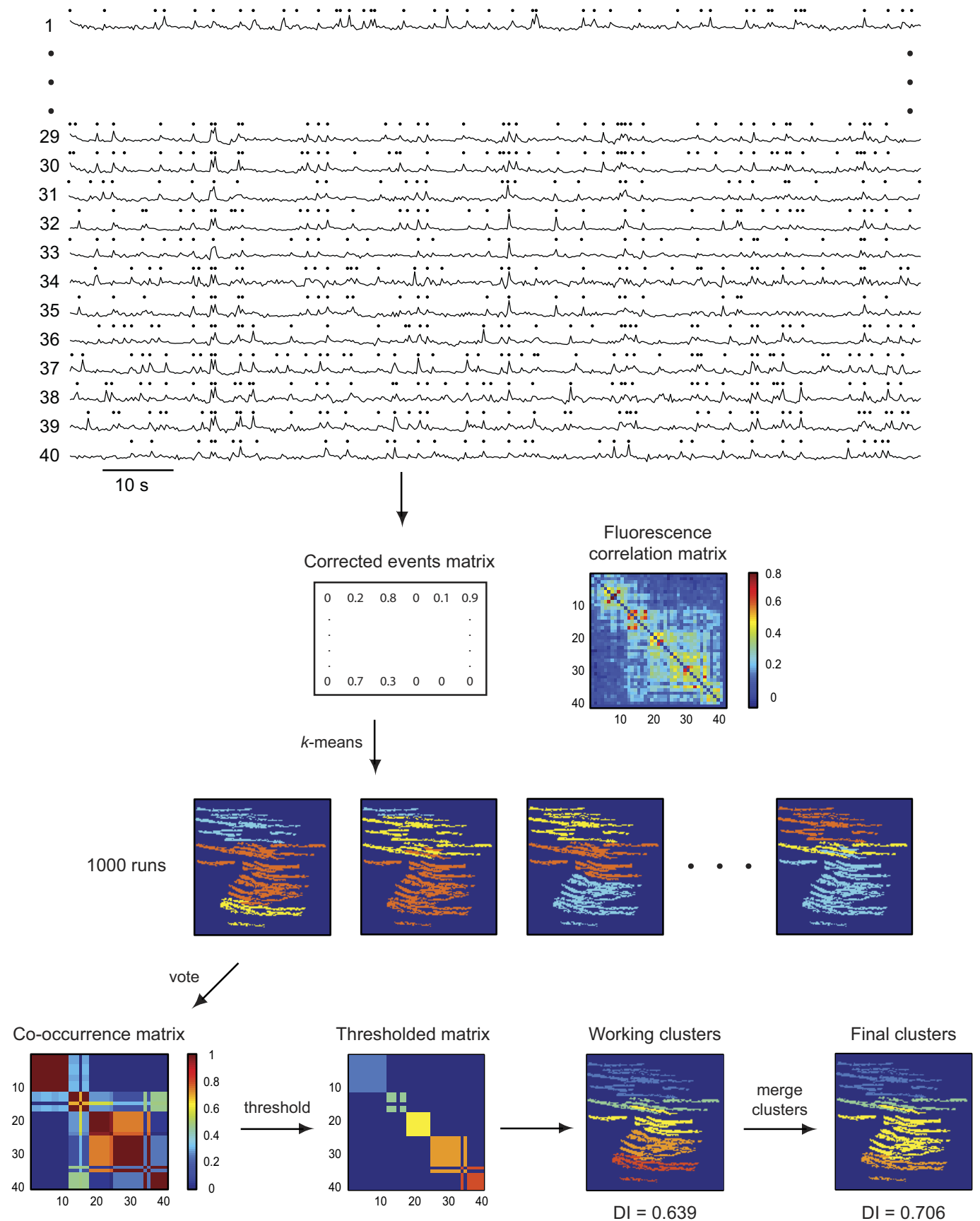
bottom row). This co-occurrence matrix always had a structure similar to that of the correlation heatmap, which shows the pairwise correlations between all pairs of dendrites in the data set (Fig. 4, *2nd row*). Working clusters were then defined as the maximum set of dendrites such that counts was greater than some threshold value, in our case 800 of 1,000, for every pair of dendrites within the cluster. This procedure usually found six to eight working clusters in a PC data set (Fig. 4, *bottom row*).

The next step was an agglomerative one in which highly correlated clusters were combined to create larger clusters that improved the overall cluster structure. To combine clusters, we found the two most well-correlated clusters and compared the values of Dunn's index before and after the merge. If the merge increased Dunn's index, the two clusters were combined to form a larger cluster. This was repeated until combining the two most-well-correlated clusters no longer increased Dunn's index. After this, less correlated pairs of clusters were tested in the same way. For an example, in Fig. 4, merging two clusters increased Dunn's index from 0.639 to 0.706 and resulted in four final clusters. In some data sets, meta-k-means was able to find a clustering structure that achieved a higher Dunn's index than the most successful single run of k-means. In these cases, the improvement was due to the exclusion of outliers that were not well correlated with any other dendrite.

To test the performance of meta-k-means, we created mock data sets with predetermined clustering (Fig. 5A, *top 2 rows*). The mock data sets were generated by first creating a matrix of conditional probabilities, then placing spikes according to these likelihoods. For every pair of dendrites, the probability matrix contained the likelihood that one dendrite would spike simultaneously given that the other dendrite was firing. For each dendrite, a certain number of spikes was placed for each dendrite, and for each spike, a simultaneous partner spike was placed according to the probability matrix. Clustering was defined by making within-cluster probabilities higher than the between cluster probabilities. Within-cluster probabilities were uniform, sampled from a normal distribution, or distance dependent. For data sets with three to six clusters, meta-k-means was reliably able to find the correct clustering, including cases with outliers, in which the outliers were separated into a separate cluster (Fig. 5A, *middle row*).

We also tested the effect of varying the parameters of meta-k-means. Using different values of k, we found that $k = 3$ was most reliably able to recover the correct clustering (93.5% of 200 mock data sets with 3–6 clusters; Fig. 5A). Using $k = 4$ gave variable results when the same mock data set was analyzed repeatedly (Fig. 5A, *top 2 rows*), suggesting that although meta-k-means can lead to more than k final clusters, an excessively high choice of k may force the artifactual finding of boundaries that cannot be removed reproducibly by combining results. For $k = 4, 5,$ and 6, respectively, the correct result was found for only 68, 45.5, and 25% of mock data sets (200 mock data sets with 3–6 clusters). Although using $k = 2$ usually found the correct clusters for mock data sets, in real

data sets with visible cluster borders, the found clusters did not match the apparent structure in four of five cases (Fig. 5A, *bottom row*). In these cases, using $k = 3$ gave the expected clustering. Thus a value for k of 3 is most reliable for finding clusters in both mock and real data sets.

Although agglomerating >1,000 trials of k-means gave the same results, using <1,000 trials sometimes gave different clusterings on multiple runs of meta-k-means. The minimum number of trials required will vary depending on the consistency of a single run of k-means for that data set. For our mock and real data sets, 1,000 trials in the meta-clustering method was sufficient to generate reproducible results; in all cases, running >1,000 trials was unnecessary.

We tested whether meta-k-means could still detect clustering when pairwise correlations were artificially weakened. We shuffled events from a real data set that contained three well-defined clusters (Fig. 5, *top row*). From all within-cluster pairs of events, in some fraction of pairs one event was randomly moved to the same time point in a dendrite in a different cluster. This procedure made the distributions of intra- and inter-cluster correlation coefficients more similar to one another (Fig. 5, histograms in *center column*). The original data set shown here had different intra- and inter-cluster correlations [intra-cluster pairwise correlations 0.23 ± 0.09, inter-cluster pairwise correlations of 0.07 ± 0.05, means $\pm$ SD, a difference of 1.6 combined SDs (cSD) where cSD was defined as $\sqrt{SD_{intra}^2 + SD_{inter}^2}$]. The clustering structure was still detected by meta-k-means after randomizing 10% of events (Fig. 5, *middle row*), a case in which intra-cluster pairwise correlations 0.15 ± 0.08 and inter-cluster correlations 0.07 ± 0.05, a difference of 1.0 cSD. However, when 20% of the events were randomized (Fig. 5, *bottom row*), cluster structure began to be degraded, coinciding with the breakdown of the difference between intra-cluster pairwise correlations, 0.12 ± 0.07, and inter-cluster correlations, 0.08 ± 0.05, a difference of 0.5 cSD). Thus meta-k-means begins to fail in tandem with reductions in the inter-cluster differences within a data set.

DISCUSSION

A growing body of in vivo multiphoton optical recordings has created a need for tools to identify active cellular structures and find patterns of activity among them. In the current paper, we have described methods for identifying structures based on their activity, inferring the times of events, and finding clusters of structures that are active together. Some of the analytical tools are specifically useful for cerebellar PC populations, the firing patterns of which reflect synchrony originating in the inferior olive. Other tools are applicable to a wider variety of ensemble imaging data.

Our design of clustering methods was driven by the structure of the data. A key step was the selection of a distance metric for similarity between event trains. A metric based on exact synchrony reflects the fact that strong synchrony in PC com-

FIG. 4. Meta-k-means. The steps of meta-k-means, illustrated on a cerebellar dataset. A selection of traces from 40 individual dendrites is shown in the *top row*. The dots indicate detected events. k-means is run 1,000 times on the corrected events matrix using $k = 3$ and random initialization of centroid positions, yielding several different clustering results (*3rd row*). The co-occurrence matrix (*bottom row*), determined by the number of times each pair of dendrites is clustered together, is similar to the fluorescence correlation matrix (*2nd row*), and is thresholded to give working clusters (*bottom row*). In this matrix, dendrite 15 was excluded from all clusters as an outlier. Cluster pairs are then merged until combining clusters no longer increases Dunn's index (DI). In the correlation matrix, dendrites were sorted according to their mediolateral position.
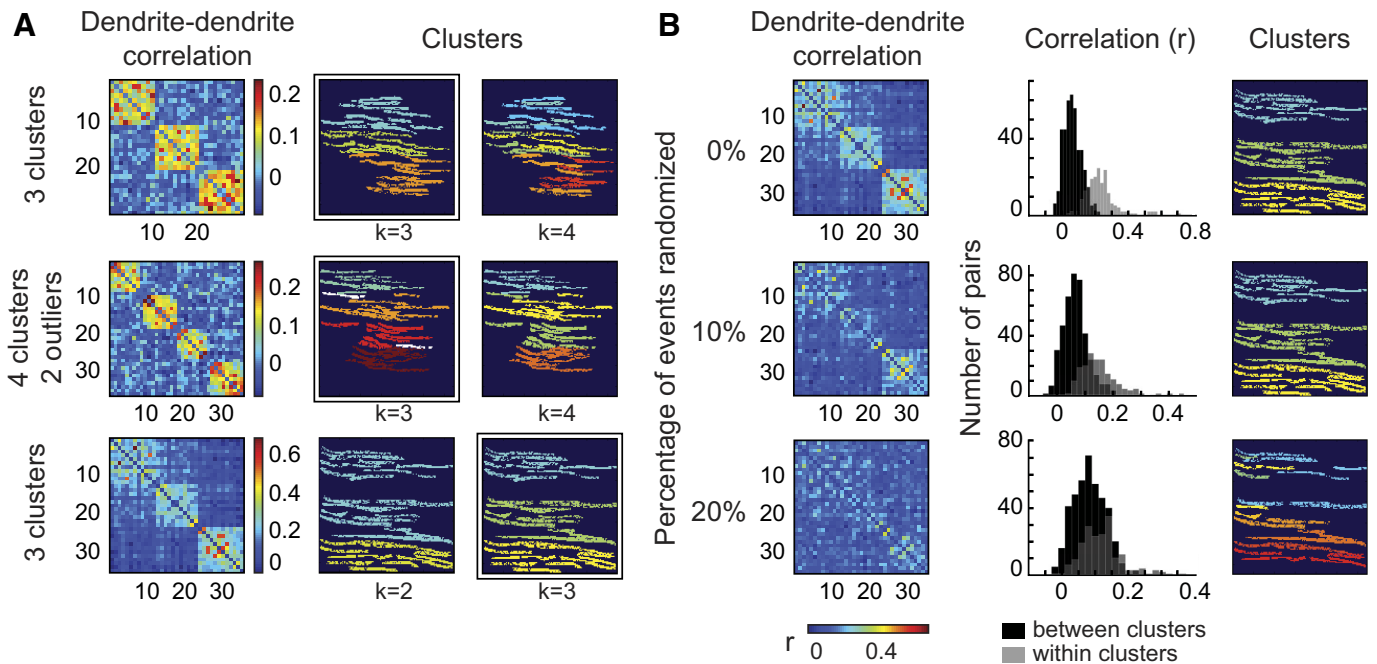
FIG. 5. Cluster determination in artificial datasets. *A*: in the mock datasets shown in the top 2 rows, meta-*k*-means found the correct clustering using $k = 3$. In the middle row which depicts a dataset with four clusters and two outliers, when $k = 4$ was used the outliers were not separated from the clusters. In the real dataset *bottom row*, the 3 clusters were found only by using $k = 3$, whereas use of $k = 2$ combined 2 of the clusters. *B, top row*: a correlation matrix for a real dataset, histograms of intracluster and intercluster dendrite-dendrite correlations, and the clustering found by meta-*k*-means. *Middle* and *bottom row*: artificial datasets generated by randomizing 10% (*middle row)* or 20% (*bottom row*) of intracluster pairs of events in the dataset. Degradation of intracluster correlations reduces the resolving power of meta-*k*-means. In this example the original three clusters were no longer identified after randomization of 20% of the events. The histograms were calculated using the original cluster assignments.

plex spike firing is driven by electrical coupling between neurons in the inferior olive, where climbing fibers originate. In the case of more complex temporal relationships, more sophisticated metrics are needed that take into account time-shift alignment (Victor 2005; Victor et al. 2007) or that otherwise relax the requirement for exact synchrony (Lindsey and Gerstein 2006). A second major choice was that of quantification of goodness of clustering (Maulik and Bandyo-padhyay 2002) using Dunn's index, which emphasizes individual cluster tightness and the distance between adjacent clusters. As signaling in the olivocerebellar system becomes further understood, the appropriateness of these metrics may need to be re-examined.

Meta-*k*-means should be useful in cases where it is desirable to identify clusters of related neurons and could be applied to both optical recordings and multielectrode data. It is mathematically similar to an approach taken to finding gene expression patterns (Le Roch et al. 2003) and is similar to previous approaches (Dudoit and Fridlyand 2003) in which the results of a less-reliable method are aggregated to generate a robust and reproducible outcome. Meta-*k*-means gives consistent results and requires less prior information compared with other partitioning approaches that have been applied to electrophysiological data such as the gravity method (Gerstein and Aertsen 1985; Lindsey and Gerstein 2006), hierarchical clustering (Eggermont 2006), and multidimensional scaling (Fukuda et al. 2001). Those methods do not explicitly address the issue of independently determining an appropriate number of clusters but instead require the number of clusters to be known beforehand. We have also shown that the parameters used in meta-*k*-means are reliable and do not require adjustment for individ-

ual mock or calcium imaging data sets, reducing bias in the results. Meta-*k*-means can be done with small data sets: reliable clustering was possible with event matrices 500 samples in duration and containing ~30 events per dendrite.

Complete analysis of a movie, from dendrite selection to final clusters, takes ~3–4 h. Most of this time is spent in selecting dendrites, which takes a trained user ~2–3 h. After dendrite selection, the event time extraction for all individual dendrites takes ~30 min. The clustering algorithm runs quite fast on computers with a 1-GHz Intel Pentium 4 processor (Intel, Santa Clara, CA) and takes a few seconds to find working clusters. The inspection and merging process to arrive at the final clustering takes a few minutes.

Several aspects of our analysis should be useful in the analysis of other neural imaging data. Likelihood-based correction of event times is important whenever the data-collection method operates at a slower rate than the physiological time scale of events. The method recovers information that would otherwise be ambiguous and is helpful for any case in which a brief event evokes a long-lasting signal. The need for likelihood-based correction can also be reduced by faster data gathering, such as by linescans, or by scanning only from regions of interest (Göbel and Helmchen 2007).

Our choice to orient dendrites parallel to the fast-scan direction represents a case analogous to that encountered in many other brain structures, where the imaged structures are typically cell bodies. Another means for potentially recovering more precise event times would be to orient the dendrites perpendicularly or obliquely to the fast-scan direction. Better disambiguation of event times will require a high signal-to-noise ratio and zooming-in of dendrites to span the field of

view, conditions that may be facilitated by improved labeling using genetically encodable probes.

Our temporal correlation-based approach to image segmentation was guided by the need to resolve densely located structures in tissue labeled with low contrast. Typically, cellular units in multiphoton and other imaging data have been identified either from their high fluorescence contrast (Dombeck et al. 2007; Kerr et al. 2007) or by using cell-specific markers (Yaksi and Friedrich 2006). In these cases, the structures of interest, usually cell bodies, are sparsely located within the field of view, making it possible to select the borders of the region of interest roughly by hand. However, closely spaced structures such as PC dendrites cover the field of view (Fig. 1D) and do not have clear borders separating them. Our correlation-based approach maximizes the internal pixel correlations within the selected region and excludes the less correlated pixels, thereby increasing the likelihood that the final selected pixels belong to the same cellular structure.

Signal synchrony can also be useful for other approaches such as principal component analysis (PCA) (Jolliffe 2002), which can be used to discard higher-order uncorrelated time courses and thereby reduce noise (Wang et al. 2000). However, PCA alone is not sufficient to distinguish individual units because synchronously active units will not have independent time courses. A more satisfactory separation of signals has recently been accomplished using independent component analysis (ICA) to distinguish active cellular structures in multiphoton data (Mukamel et al. 2007). In cases where the image is less stable, such as awake animal recording, correction for tissue movement may also be warranted (Dombeck et al. 2007).

The approaches described here arose as a means of dealing with the specific qualities and limitations of currently generated data sets. In the future, advances in signal quality such as cell type-specific expression of functional indicators (Hasan et al. 2004) and the development of faster-responding indicators will reduce the need for these particular methods but also create new opportunities and problems in data analysis.

REFERENCES

Bagirov AM, Ferguson B, Ivkovic S, Saunders G, Yearwood J. New algorithms for multi-class cancer diagnosis using tumor gene expression signatures. *Bioinformatics* 19: 1800–1807, 2003.

d'Haeseleer P. How does gene expression clustering work? *Nat Biotechnol* 23: 1499–1501, 2005.

Dombeck DA, Khabbaz AN, Collman F, Adelman TL, Tank DW. Imaging large-scale neural activity with cellular resolution in awake, mobile mice. *Neuron* 56: 43–57, 2007.

Dudoit S, Fridlyand J. Bagging to improve the accuracy of a clustering procedure. *Bioinformatics* 19: 1090–1099, 2003.

Eccles JC, Llinás R, Sasaki K, Voorhoeve PE. Interaction experiments on the responses evoked in Purkinje cells by climbing fibers. *J Physiol* 182: 297–315, 1966.

Eggermont JJ. Properties of correlated neural activity clusters in cat auditory cortex resemble those of neural assemblies. *J Neurophysiol* 96: 746–764, 2006.

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

Gerstein GL, Aertsen AM. Representation of cooperative firing activity among simultaneously recorded neurons. *J Neurophysiol* 54: 1513–1528, 1985.

Göbel W, Helmchen F. New angles on neuronal dendrites in vivo. *J Neurophysiol* 98: 3770–3779, 2007.

Hartigan JA, Wong MA. Algorithm AS 136: a k-means clustering algorithm. *Appl Stat* 28: 100–108, 1979.

Hasan MT, Friedrich RW, Euler T, Larkum ME, Giese G, Both M, Duebel J, Waters J, Bujard H, Griesbeck O, Tsien RY, Nagai T, Miyawaki A, Denk W. Functional fluorescent Ca^{2+} indicator proteins in transgenic mice under TET control. *PLoS Biol* 2: e163, 2004.

Helmchen F, Waters J. Ca^{2+} imaging in the mammalian brain in vivo. *Eur J Pharmacol* 447: 119–129, 2002.

Jain AK, Dubes RC. *Algorithms for Clustering Data.* Englewood Cliffs, NJ: Prentice-Hall, 1988.

Jolliffe IT. *Principal Component Analysis.* New York: Springer, 2002.

Kerr JN, de Kock CP, Greenberg DS, Bruno RM, Sakmann B, Helmchen F. Spatial organization of neuronal population responses in layer 2/3 of rat barrel cortex. *J Neurosci* 27: 13316–13328, 2007.

Lang EJ, Sugihara I, Welsh JP, Llinás R. Patterns of spontaneous purkinje cell complex spike activity in the awake rat. *J Neurosci* 19: 2728–2739, 1999.

Le Roch KG, Zhou Y, Blair PL, Grainger M, Moch JK, Haynes JD, De La Vega P, Holder AA, Batalov S, Carucci DJ, Winzeler EA. Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science* 301: 1503–1508, 2003.

Lindsey BG, Gerstein GL. Two enhancements of the gravity algorithm for multiple spike train analysis. *J Neurosci Methods* 150: 116–127, 2006.

Llinás R, Sugimori M. Electrophysiological properties of in vitro Purkinje cell dendrites in mammalian cerebellar slices. *J Physiol* 305: 197–213, 1980.

Maulik U, Bandyopadhyay S. Performance evaluation of some clustering algorithms and validity indices. *IEEE Trans Pattern Anal Machine Intell* 24: 1650–1654, 2002.

Mukamel EA, Nimmerjahn A, Schnitzer MJ. Automated sorting of intracellular calcium signals for large-scale imaging studies of neuronal and glial populations. *Soc Neurosci Abstr* 252.2, 2007.

Ohki K, Chung S, Ch'ng YH, Kara P, Reid RC. Functional imaging with cellular resolution reveals precise micro-architecture in visual cortex. *Nature* 433: 597–603, 2005.

Palay SL, Chan-Palay V. *Cerebellar Cortex: Cytology and Organization.* Berlin: Springer, 1974.

Ross WN, Werman R. Mapping calcium transients in the dendrites of Purkinje cells from the guinea pig cerebellum in vitro. *J Physiol* 389: 319–336, 1987.

Ruigrok TJ, Voogd J. Organization of projections from the inferior olive to the cerebellar nuclei in the rat. *J Comp Neurol* 426: 209–228, 2000.

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

Sugihara I, Wu HS, Shinoda Y. The entire trajectories of single olivocerebellar axons in the cerebellar cortex and their contribution to cerebellar compartmentalization. *J Neurosci* 21: 7715–7723, 2001.

Sullivan MR, Nimmerjahn A, Sarkisov DV, Helmchen F, Wang SS-H. In vivo calcium imaging of circuit activity in cerebellar cortex. *J Neurophysiol* 94: 1636–1644, 2005.

Tavazoie S, Hughes JD, Campbell MJ, Cho RJ, Church GM. Systematic determination of genetic network architecture. *Nat Genet* 22: 281–285, 1999.

Tsien RY, Pozzan T, Rink TJ. Calcium homeostasis in intact lymphocytes: cytoplasmic free calcium monitored with a new, intracellularly trapped fluorescent indicator. *J Cell Biol* 94: 325–334, 1982.

Victor JD. Spike train metrics. *Curr Opin Neurobiol* 15: 585–592, 2005.

Victor JD, Goldberg DH, Gardner D. Dynamic programming algorithms for comparing multineuronal spike trains via cost-based metrics and alignments. *J Neurosci Methods* 161: 351–360, 2007.

Wang SS-H, Khiroug L, Augustine GJ. Quantification of spread of cerebellar long-term depression with chemical two-photon uncaging of glutamate. *Proc Natl Acad Sci USA* 97: 8635–8640, 2000.

Yaksi E, Friedrich RW. Reconstruction of firing rate changes across neuronal populations by temporally deconvolved Ca^{2+} imaging. *Nat Methods* 3: 377–383, 2006.