

Independent component analyses for quantifying neuronal ensemble interactions

Mark Laubach *, Marshall Shuler, Miguel A.L. Nicolelis

Department of Neurobiology, Box 3209 (Room 333), Bryan Research Bldg., 101 Research Drive, Duke University Medical Center, Durham, NC 27710, USA

Received 10 July 1999; accepted 5 August 1999

Abstract

The goal of this study was to compare how multivariate statistical methods for dimension reduction account for correlations between simultaneously recorded neurons. Here, we describe applications of principal component analysis (PCA) and independent component analysis (ICA) (Cardoso J-F, Souloumiac A. *IEE-Proc F* 1993;140:362–70; Hyvarinen A, Oja E. *Neural Comput* 1997;9:1483–92; Lee TW, Girolami M, Sejnowski TJ. *Neural Comp* 1999;11:417–41) to neuronal ensemble data. Simulated ensembles of neurons were used to compare how well the methods above could account for correlated neuronal firing. The simulations showed that ‘population vectors’ defined by PCA were broadly distributed over the neuronal ensembles; thus, PCA was unable to identify independent groupings of neurons that shared common sources of input. By contrast, the ICA methods were all able to identify groupings of neurons that emerged due to correlated firing. This result suggests that correlated neuronal firing is reflected in higher-order correlations between neurons and not simply in the neurons’ covariance. To assess the significance of these methods for real neuronal ensembles, we analyzed data from populations of neurons recorded in the motor cortex of rats trained to perform a reaction-time task. Scores for PCA and ICA were reconstructed on a bin-by-bin basis for single trials. These data were then used to train an artificial neural network to discriminate between single trials with either short or long reaction-times. Classifications based on scores from the ICA-based methods were significantly better than those based on PCA. For example, scores for components defined with an ICA-based method, extended ICA (Lee et al., 1999), classified more trials correctly ($80.58 \pm 1.25\%$) than PCA ($73.14 \pm 0.84\%$) for an ensemble of 26 neurons recorded in the motor cortex (ANOVA: $P < 0.005$). This result suggests that behaviorally relevant information is represented in correlated neuronal firing and can be best detected when higher-order correlations between neurons are taken into account. © 1999 Elsevier Science B.V. All rights reserved.

Keywords: Statistics; Independent component analyses; Neuronal ensemble interactions

1. Introduction

An emerging view in systems neuroscience is that behaviorally relevant information is represented by the concurrent activation of populations of neurons that form cell assemblies distributed across multiple levels of the nervous system (see Nicolelis et al., 1997b for review). The introduction of recording methods for the acquisition of spike trains from large populations of neurons (e.g. Kruger and Bach, 1981; McNaughton et al., 1983; Reitboeck, 1983; Nicolelis et al., 1997a; see

Nicolelis, 1999 for review) has intensified interest in the investigation of how distributed cell assemblies process behaviorally relevant information. A critical step for achieving this goal is the introduction of methods for data analysis that could identify functional neuronal interactions within the high dimensional data sets collected with these new recording methods.

The initial basis of this study was the realization that multivariate statistical methods for dimension reduction might be able to reconstruct functional interactions between multiple neurons by providing a bin-by-bin estimate of correlated neuronal firing. Classically, dimension reduction techniques have been used to transform a large set of observed variables into a smaller set of arbitrary variables, or factors, that are derived from

* Corresponding author. Tel.: +1-919-684-4581; fax: +1-919-684-5435.

E-mail address: laubach@neuro.duke.edu (M. Laubach)

some property of the original data set. In the case of principal component analysis (PCA), the goal is to find a small set of linear functions that account for the covariance structure of the original high dimensional data set (see Chapin, 1999 for review). By contrast, recently developed methods for independent component analysis (ICA) attempt to isolate factors that are derived from higher-order correlations within the dataset, i.e. the higher moments describing the distribution of the variables, such as kurtosis (see Comon, 1994 for review).

Here, we report an application of ICA that allows for the visualization and quantification of behaviorally relevant information represented through correlated neuronal firing. To further validate this new approach, ICA was compared with a classical multivariate statistical method, principal component analysis, using data from (1) simulated ensembles of Poisson spike trains and (2) simultaneously recorded spike trains ($n = 14\text{--}32$) recorded in the rat motor cortex during the performance of a reaction-time task (Laubach et al., 1998). The main goal of this study was to investigate how PCA and ICA detect patterns of correlations across large ensembles of neurons. In addition, we show that these methods provide a quantitative representation of correlated neuronal firing that can be further analyzed with other methods, including techniques for statistical pattern recognition.

2. Multivariate statistical methods

2.1. Principal component analysis

PCA is a classical method for multivariate data analysis that describes interactions within a given data set that are derived from the matrix of correlations between all pairs of variables. PCA has been used for many years as a standard method for preprocessing multivariate data sets to reduce their dimensionality. The goal of PCA is to find a reduced set of new variables that, in a decreasing order, account for the largest portions of covariance across the dataset. Further details about PCA can be found in textbooks on multivariate analysis (e.g. Johnson and Wichern, 1992; Reyment and Jöreskog, 1993).

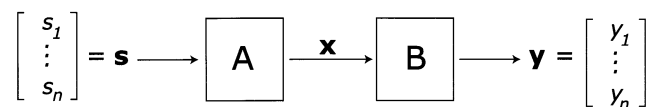


Fig. 1. In the blind source separation problem, one assumes that a set of measured signals, x , reflects the linear sum of some set of sources, s . The problem would be solved in theory if one could determine a set of weights, \mathbf{A} , that are responsible for mixing the sources over the signals, i.e. the ‘mixing’ matrix. In practice, this is done by finding a different set of weights, \mathbf{B} , that attempt to produce a new set of signals, y , that are as independent statistically as possible.

Neurophysiological applications are described in Nicolelis and Chapin (1994) and Chapin and Nicolelis (1999).

By design, individual principal components are unable to detect epochs of correlated neuronal firing between subsets of a neuronal ensemble. This is because the principal components are derived from the (second-order) correlation matrix for the neuronal ensemble. Thus, the first principal component is chosen so as to maximize the correlations between all neurons. Typically, this population vector is broadly mapped over the collection of neurons and is highly related to the average response of the neuronal ensemble (Nicolelis et al., 1998). Subsequent population vectors are derived to account for maximal, orthogonal (co-)variance. However, the higher principal components (PC2 and higher), which often tend to account for the activity of subsets of neurons, are always dependent on the choice of the first component. Because the first component reflects the overall activity level of the neurons, the higher principal components are never able to clearly resolve independent sources of excitation from the overall activity of the ensemble. We illustrate this point below using data from simulated neuronal ensembles and from neurons recorded in the motor cortex of behaving rats.

Varimax rotation (VMAX) is a modification of basic PCA that attempts to simplify the weight structure derived from PCA so that each component has a few variables with large coefficients and many with near-zero coefficients (Kaiser, 1958). The goal of VMAX is to adjust the coefficients for each principal component such that they are all near zero or near one. The results of this rotation is that the variance accounted for by each component is near maximal. A good account of the algorithm for VMAX is given in Reyment and Jöreskog (1993).

In this paper, PCA was carried out using Matlab (The Mathworks, Natick, MA) and the Matlab statistics tool box. VMAX was done using code for Matlab as described in Reyment and Jöreskog (1993).

2.2. Independent component analysis

Several methods have been developed recently for ICA. These methods have been used mostly for the ‘blind source separation’ (BSS) problem (Fig. 1), where signals from a number of sources are mixed together in some visual or auditory dataset and the goal is to ‘un-mixed’ and recover the sources. The BSS problem is typically represented as:

$$x(t) = \mathbf{A}s(t), \quad (1)$$

where $x(t)$ represents the n observed signals (i.e. neurons) over t time-points (i.e. bins), $s(t)$ represents the n source signals, and \mathbf{A} is the ‘mixing matrix’, that must

be resolved for source separation (see Comon, 1994 and Cardoso, 1998 for review). The main assumptions of the BSS problem is that the sources are independent and that they can be deduced using only the observed signals. This problem is reformulated to the following:

$$y(t) = \mathbf{B}x(t), \quad (2)$$

where $y(t)$ represents the statistical estimates of the sources that are derived from analyses such as ICA and \mathbf{B} is the ‘unmixing matrix’ that is provided by a given algorithm for ICA. In general, ICA has been used under the assumption that the number of sources are equal to the number of signals. In dimension reduction applications, this assumption is no longer made. Instead, one assumes the number of sources is less than the number of signals.

Applications of ICA to the BSS problem work because ICA resolves commonalities in the phase of the observed signals (Bell and Sejnowski, 1997), which are presumably due to the signals being influenced by common sources. Phase relations between the observed signals are usually achieved by quantifying the kurtosis of the signals and finding factors that account for aspects of the signals that are ‘commonly peaked’, i.e. synchronous signals across multiple channels. In our application of ICA, PCA is used first for dimension reduction. Then, ICA is used to rotate the weights from PCA to make each component as independent as possible. This is done by remapping the weights from PCA to maximize the kurtosis of each component. These signals invariably exhibit supra-Gaussian kurtosis, which implies that the signals are clustered around a range of values that is smaller than expected from a similar data set (i.e. with the same mean and standard deviation) generated from the Gaussian distribution. Theoretically, a remapping of PC weights that takes epochs of synchronous firing into account, should ‘unmix’ independent signals that are represented by multiple PCs and produce a set of independent components.

In the context of neuronal ensemble data, we interpret supra-Gaussian kurtosis in scores for the principal components as evidence for an independent brain signal, such as a ‘common input’ in the sense of cross-correlation analysis. For example, if there are epochs of correlated firing that occur for a subset of the neuronal ensemble, these epochs will generate large scores in any principal component that maps onto the correlated neurons. When the PC weights are remapped to account for the kurtosis of the PC scores, they will be adjusted such that a new set of weights is found that indicates which neurons tend to fire together.

Two types of algorithms have been used to solve the ICA problem. Algebraic methods, such as the joint approximate diagonalization of eigenmatrices method, or JADE (Cardoso and Souloumiac, 1993), utilize matrix transformations to find a mapping over the vari-

ables that maximize for their kurtosis. By contrast, neural network methods for ICA, such as those used in this paper (Hyvarinen and Oja, 1997; Lee et al., 1999), solve this problem by finding a set of weights that maximize the entropies of each independent component and minimize the mutual information between independent components, an idea that was proposed originally by Bell and Sejnowski (1995).

The success of ICA for finding commonalities amongst multiple, simultaneously recorded neuronal spike trains, such as correlated firing, is directly related to ICA being based on analyses of higher-order correlations. For our application of ICA, we assume that the sources are brain-derived signals that alter the excitability of multiple neurons in our simultaneously recorded neuronal ensembles. In the simulations used in this paper, this relation was studied explicitly by adding spikes from the sources to the simulated spike trains. In data collected from the motor cortex, it is assumed that commonalities in the neurons’ response properties arise from common sources of inputs either from the cortex or subcortical areas (e.g. thalamus) that project to the motor cortex. Such sources of common input could produce nearly simultaneous increases in the firing probabilities of the neurons that receive their projections. Thus, neurons that receive common inputs will have correlated responses, both around the time of the event that drives the source neurons and on a trial-by-trial basis. This point is illustrated below using simulated neuronal ensembles and neuronal ensembles recorded in the motor cortex of rats that performed a reaction-time task.

3. Methods

3.1. Simulated spike trains

Simulated spike trains were created as follows: Background discharge was created for 64 channels of data over 1000 ‘trials’. For each time-point in the spike train (an array with 1 ms bins for each time-point), a Poisson process was simulated with a rate of 1 Hz. If the generator produced a count, a spike was added to the bin. A minimum refractory period of 2 ms was used such that if a spike was placed in a given bin, then the Poisson generator skipped over the next two bins in the spike train.

Four sources of excitation were generated as above except that the rate of the Poisson generator was set to between 50 and 200 Hz (with a single rate used for a given simulation). In addition, spikes were only generated for a window of time, the ‘response window’, after the occurrence of a periodic ‘stimulus’ that occurred at 100 ms intervals. The timing of the simulated neuronal responses was based on the response properties of real

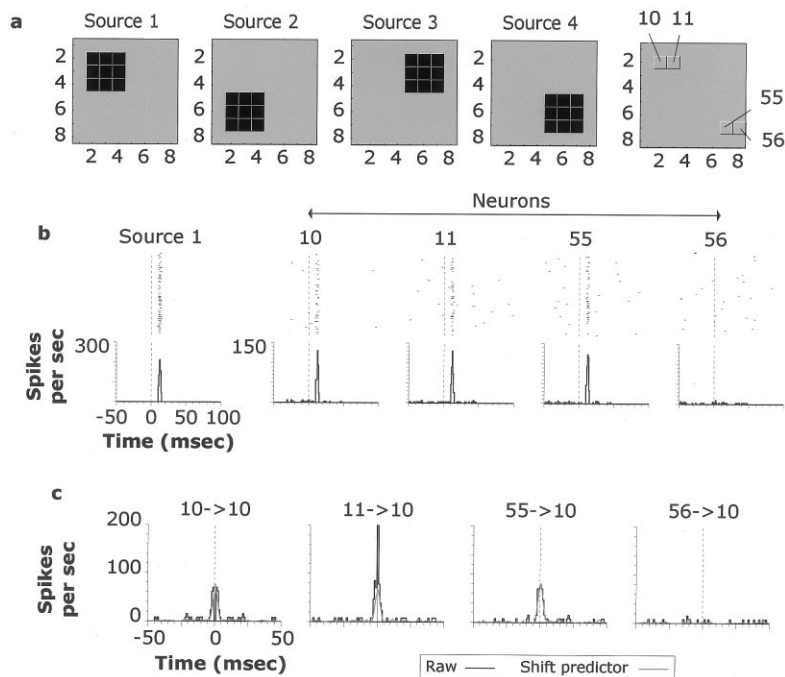


Fig. 2. An example of the simulated neuronal ensembles. (a) The distribution of four sources over an ensemble of 64 neurons over 1000 trials. (b) Response properties of source 1 and four neurons. (c) Cross-correlational analysis for these neurons.

cortical neurons studied in receptive field mapping experiments (A.A. Ghazanfar, personal communication). The response window began 8 ms after the stimulus and ended 15 ms later. The maximum probability of spiking within the window occurred at a latency of 12 ms.

Spikes were added from the sources to nine neurons in the simulations of 64 neurons. For a given simulation, the correlation between the sources and the neurons was set to 0.2–0.8 (with a single correlation used for a given simulation). Source spikes were added for each neuron independently using a pseudo-random number generator based on the uniform distribution. Spike times from the sources were randomly jittered in some simulations by adding 0, 1, or 2 ms to the times of the sources spikes. If the neuron's spike train contained a (background) spike at the time of the source spike, then the source spike was not added to the neuron's spike train. Cross-correlation analysis was then used to show that neurons that shared a common source had peaks in their cross-correlation histograms that were more than was expected from the shift predictor (Perkel et al., 1967), which was determined using Stranger (Biographics, Winston-Salem, NC). By contrast, neurons that received spikes from different sources had central peaks in their cross-correlation histograms, but these were not larger than expected from the shift-predictor test.

These points are illustrated in Fig. 2, which shows data from one of the simulated neuronal ensembles. In panel (a), the distribution of the four sources' influence over the ensemble of 64 neuron-like elements, which

were arranged in a 8×8 matrix, is shown. In addition, the relation of four of the neurons to each source is depicted in the right part of panel (a). Raster plots and peri-event histograms for source 1 and four neurons from a typical simulation are shown in panel (b). The 'trigger event' (time = 0) produced increased activity in source 1, which resulted in increased activity in neurons 10 and 11. Neuron 55 had a similar response pattern but was driven by a different source (#4). Neuron 56 was not driven by any of the four sources and so only contained random, background activity. Cross-correlation histograms for the neurons are shown in panel (c). The black lines represent the raw correlation between neuron 10 and itself (autocorrelation) and the other three neurons. The red lines represent the shift predictor, which accounts for correlations between neurons with similar response properties that are due to different sources of excitation. Note that the raw correlogram for the simulated neurons that were excited by a common source (#10 and #11) has a peak around zero lag that is larger than the peak in shift predictor. By contrast, the correlogram for neurons 10 and 55, which did not receive common inputs, is not larger than that of the shift predictor.

3.2. Neurophysiological data

Adult male rats were trained to perform a simple reaction-time task. In the task, the animals were required to maintain a lever press over a variable interval (400–800 ms) and to release the lever within 1 s of the

onset of either a tone, vibration of the lever, or the combination of these stimuli. Correct performance of the task was rewarded by providing the rats with a drop of water 100 ms after each correct response. Incorrect responses occurred when the rats released the lever before the stimuli or had reaction-times longer than 1 s. These inappropriate behaviors were signaled by turning off the houselights for 2 s (i.e. time-out). After the animals demonstrated asymptotic performance in the task, arrays of microwire electrodes were implanted in the motor and premotor cortices using standard stereotaxic and aseptic procedures (Nicolelis et al., 1997a).

The behavioral data were analyzed as follows: Reaction-times (RTs) were computed as the time from the onset of the trigger stimuli to the initial movement of the response lever, which was detected by a microswitch. The upper and lower quartiles for the distributions of RTs were determined and used to partition the trials into those with short RTs (i.e. lower quartile), long RTs (i.e. upper quartile), and intermediate RTs (i.e. middle 50% of data). Quantitative analyses of the relationship between neuronal responses and RT behavior were made by comparing trials with short and long RTs using the methods for PCA and ICA described in this paper and artificial neural networks that

were trained to discriminate between trials with RTs in the lower and upper quartiles.

3.3. PCA and ICA

The steps in our data analysis procedure are depicted in Fig. 3. The organization of the data was as follows: First, a matrix was constructed from the spike trains of the neurons (Fig. 3a). Each row in the matrix represented neural activity from a single behavioral trial. Each column represented a time point (i.e. a bin of 1 ms) in a spike train of one of the neurons. Sets of columns containing spikes for the individual neurons were appended together, to create the raw data matrix. This matrix had as many columns as the product of the number of neurons and the number of bins. The spike trains were smoothed and decimated five times, using low-pass filtering with a 3-tap Hamming window and decimating the spike trains with a 32-point Kaiser window. This procedure gave rise to an effective bin size of 5 ms. Filtering and decimation was done using routines in the Signal Processing toolbox for Matlab.

The matrix of smoothed spike counts described above was then rearranged such that each neuron was represented as a column and each bin was represented as a row. Spike trains from individual trials were thus

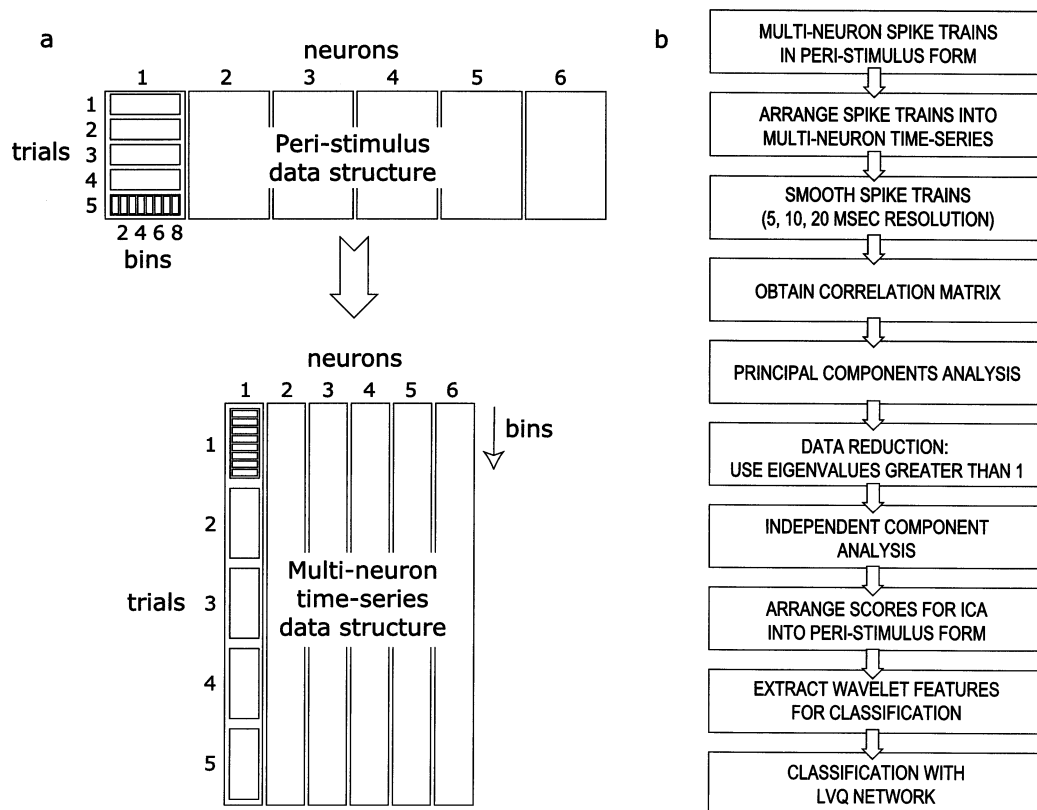


Fig. 3. The data structures and analysis procedures used in this study. (a) Data structures for applying independent component analysis to neuronal ensemble data. (b) The series of steps in data analysis used in this study.

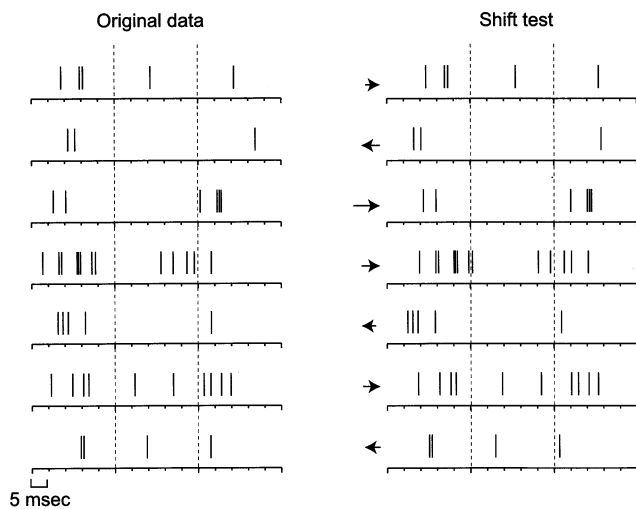


Fig. 4. Spike shifting was used to compare how well PCA and different methods for ICA accounted for correlated neuronal firing. Here, spike trains from a group of seven neurons are shown for the original condition (left) and after shifting (right). Each row in the plots represents spikes from a different neuron. Spike trains from each neuron and each trial were shifted by a random amount of time and the amount each neuron was shifted was chosen independently. The amount and direction of time shift for each neuron is indicated by the size of the arrow head to the left of the raster plots.

appended together sequentially. The mean value for each column (i.e. for each neuron) in this matrix was subtracted and each column was then normalized to have unit variance. Ensemble averages were computed as the average response over the collection of rows representing each column (i.e. time-point or bin). Principal components for the zero-mean, unit variance matrix of smoothed spike counts were extracted using the singular value decomposition algorithm. The number of major components, that indicated the number of interacting groups in the ensembles, was identified by finding the number of eigenvalues that were larger than 1 (for data with variance of 1). This value was also used to limit the number of independent components extracted from the matrix of spike counts and to determine the number of principal components that were rotated using VMAX.

Logistic (lICA) and extended (xICA) ICA were run with principal component dimension reduction and data whitening (Lee et al., 1999). Fast ICA (Hyvarinen and Oja, 1997) was done with principal component dimension reduction and with sequential and simultaneous extraction of the independent components using the following contrast functions: kurtosis (fICAk), tangsig (fICAt), and the Gaussian non-linear function (fICAg). Scores for the principal and independent components were reconstructed on a bin-by-bin basis and the resulting matrices representing the components were re-arranged into a peri-event form (i.e. each row is a trial, and each column is a bin, with a block of x bins representing a component's score with x bins per trial).

Local features from the average peri-event histograms for the single neurons, the ensemble average, and the principal and independent components were identified using the wavelet-based method, discriminant pursuit (Buckheit and Donoho, 1995). These features were then further analyzed with an artificial neural network for statistical pattern recognition. The wavelet analysis was done as follows: The average response for trials with long reaction-times was subtracted from the average for trials with short reaction-times, to give a set of 'difference vectors' for each representation of the neuronal population. The difference vectors were then decomposed into a wavelet packet table using the Daubechies 4-point wavelet. Entries in the wavelet packet table that accounted for the largest amplitudes in the difference vectors were selected as local features. Scores for these features were extracted on a trial-by-trial basis and used as input to competitive neural networks, trained with a learning vector quantization algorithm (Kohonen, 1997; see Nicolelis et al., 1999 for further details on this method), that classified single trials as having either short or long reaction-times. Leave-one-out, or N -fold, cross-validation (Mosteller and Tukey, 1977) was used to assess error rates from this classification procedure.

To test if precise timings between spikes across the group of neurons was an essential feature for PCA and ICA, a spike shifting procedure was used (Nicolelis et al., 1999). Spike shifting (see Fig. 4) was performed as follows: Spikes from each neuron were pseudo-randomly shifted in time between 1 and 7 bins, which gave perturbations in spike timing of between 10 and 70 ms. This was done to reduce the zero-lag coherence between neuronal spike trains on a trial-by-trial basis. These data were used as testing data for the artificial neural network that was trained with scores for the independent components obtained from the original, unshifted spike trains. The effect of the shift test was to perturb both the temporal precision of the neuronal firing patterns and the zero-lag coherence of the patterns.

4. Results

4.1. Simulated spike trains

Analyses of the simulated neuronal ensembles showed that the ICA-based methods and varimax rotation of the principal components were able to segregate the sources to the appropriate neurons (Fig. 5). By contrast, standard PCA was unable to resolve the influence of the individual sources on the ensembles. That is, VMAX and all methods for ICA weighted highly on a few neurons that were driven by a common source. By contrast, weights for PCA were broadly distributed over many neurons, with PC1 having weights for neu-

rons that were not even driven by the sources. This result indicates that the ICA methods and VMAX are better suited for detecting sparsely distributed functional relations within a neuronal ensemble. In other words, if information from one cortical area was conveyed to a restricted portion of some neuronal ensemble, then the ICA-based methods would be able to detect the input and assign it to the appropriate neurons. This same input would be represented broadly by multiple PCs, thereby masking the interactions between neurons that were driven by the input of interest.

The signals described from these weights are shown for some of the representations and are compared with the actual source signals in Fig. 6. Whereas scores for the PCs were effected by spikes in any of the four sources, scores for both JADE and xICA were more exclusively driven by the activity of a single source. This result was quantified by finding the Pearson correlation coefficients between the sources and the scores for functions derived with all methods. The maximum correlations between a single source and the components are shown on average in Fig. 7(a). Note that all methods except PCA were able to resolve at least one source (i.e. correlation coefficient large than 0.8). In addition, the ratios of the correlation between the components for each method and its two best sources are shown in Fig. 7(b). These values revealed that the ICA-based methods were more selective than either PCA or VMAX and that IICA was somewhat worse

than the other methods for ICA. These results imply that ICA-based methods identified the individual sources that drove the activity of subsets of the neuronal ensembles. By contrast, PCA represented correlations between neurons in a very different manner, which was too broad and did not reflect the nature of the underlying neuronal interaction.

The shift test showed that the temporal precision of the correlated neuronal firing was no more than 20 ms (Fig. 8). That is, the ICA-based methods were selective to a given source only when the spike trains were shifted by less than 20 ms and these relationships were lost with further destruction of the temporal precision of the correlated neuronal firing. By contrast, PCA was not able to account for the source very well on any time scale and actually showed spuriously high selectivity for the source 1 when the spike trains were shifted more than 30 ms (see Fig. 8(c)). Thus, it appears that relations between sources of excitation and neuronal responses may only be valid on fairly precise time-scales, on the order of 30 ms or less.

4.2. Neurophysiological data

Ensembles of 14–32 neurons that were recorded during behavioral experiments in which rats performed a reaction-time task were used to extend the results above to actual neuronal ensemble data (Laubach et al., 1998). The analyses were based on partitioning the

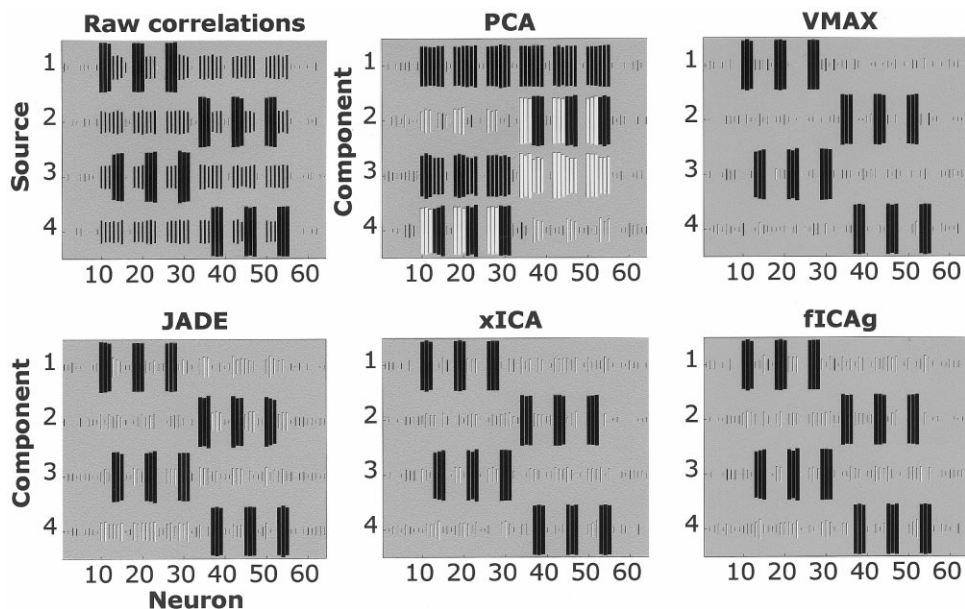


Fig. 5. Subsets of nine neurons in the simulated ensembles of 64 neurons received inputs from one of four sources (upper left plot). These distributions of source inputs were differentially identified and segregated by the PCA and ICA based methods. Note that the first principal component (PC) was broadly mapped over the entire neuronal ensemble, even on neurons that had only spontaneous background activity. The higher PCs appeared to segregate the different sources by contrasting the activity of different pairs of neuronal subsets. By contrast, the ICA-based methods, and also varimax rotation of the PCs, identified functional interrelations between the neurons and separated different sources to individual components. The plots in this figure are Hinton diagrams that depict the sign and magnitude of the coefficients using boxes of different sizes (i.e. magnitude) and color (i.e. sign). Black boxes represent positive coefficients and white boxes represents negative coefficients.

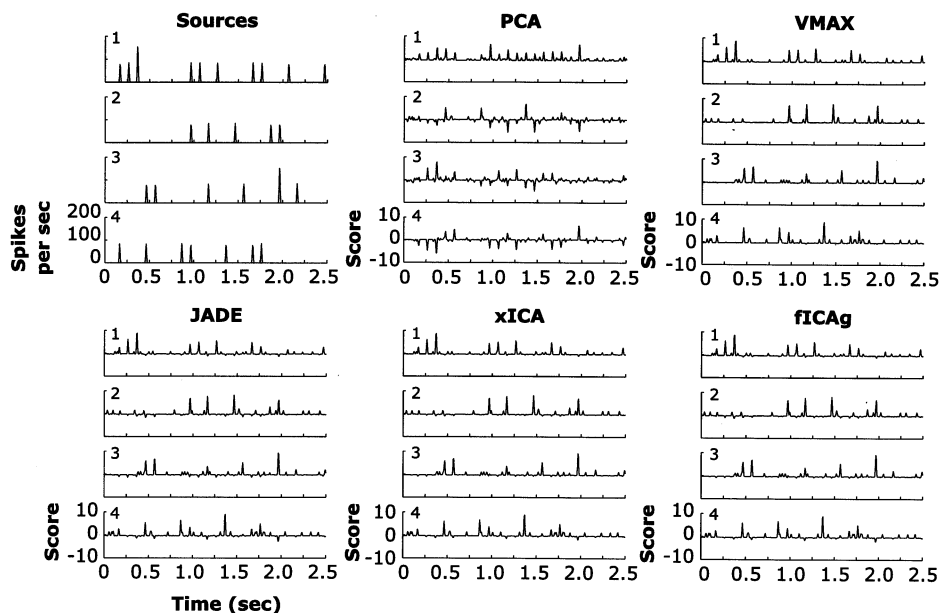


Fig. 6. ICA-based methods were better able to segregate the individual sources to individual components. The sources (upper left) and scores for three of the methods evaluated in the paper are shown over a segment of 500 ms. Note that scores for JADE (lower left) and xICA (lower right) were much more related to the single sources than were the scores for any of functions defined by PCA (upper right).

collection of reaction-times for all trials into quartiles and comparing trials in the lower quartile with trials in the upper quartile. The goal of the study was to investigate if the rats' reaction-times on single trials could be predicted by the firing of neurons in the caudal and rostral forelimb areas of the motor cortex. Neuronal activity over the 250 ms before and 100 ms after the rat released a response lever (a behavior that defined the operant response in the task) was analyzed with the various methods for PCA and ICA described above.

The different methods produced very different accounts of neuronal interactions across simultaneously recorded neuronal ensembles from the rat motor cortex. The data set shown in Fig. 9 contained 26 neurons, with 18 neurons in the caudal forelimb area and six neurons in the rostral forelimb area. There were seven eigenvalues larger than 1 for this ensemble of neurons. In Fig. 9, coefficients are shown for PCA, varimax rotation of the coefficients from PCA (VMAX), and for two ICA-based methods, JADE and xICA. Whereas the coefficients defined through PCA and VMAX were broadly mapped over the neuronal ensemble, coefficients from JADE and xICA were much more sparsely distributed over the neuronal ensembles. The population vectors defined by these different methods were based on different levels of correlated neuronal firing. PCA-based methods accounted for broadly distributed neuronal correlations and ICA-based methods accounted for more restricted interactions between the neurons.

The ICA-based methods were able to more clearly segregate the response properties of the neurons during

the time around the lever release for trials with short reaction-times (solid, dark lines) and those for trials with long reaction-times (dashed, light lines). As is shown in Fig. 10, scores for PCA and VMAX were highly similar despite the differences in the coefficients defined by these two methods (Fig. 9). By contrast, the average scores for population vectors identified with the ICA-based methods were somewhat different. For example, while the second component from xICA was highly similar to the third principal component, only the ICA-based method shows a peak on trials with short reaction-times while PCA shows peaks for both types of trials. When scores for population vectors defined with PCA, VMAX, and the ICA methods were used to train ANNs to predict the animals' reaction-times on single trials, the methods gave differential predictions of trial outcomes. The best results were achieved using scores for independent components identified with xICA ($80.58 \pm 1.25\%$), which classified more trials correctly than scores for PCA ($73.14 \pm 0.84\%$) (Fig. 11; ANOVA: $P < 0.005$). This result suggests that the higher-order correlations between these neurons, which were best resolved by xICA, conveyed behaviorally relevant information.

Finally, when the shift test was used for these data, we observed that the variances of the scores for the population vectors were reduced (Fig. 12(a)). Moreover, the degree to which mean signals for trials with short and long reaction-times differed was reduced and certain aspects of the response properties were lost altogether (Fig. 12(b)). The time-scale of this neuronal shift was over ± 3 bins, or 30 ms. Therefore, these data

show that behaviorally relevant information could be represented in correlated neuronal firing in the rat motor cortex and that this information is based on a synchronization of the spike trains on a relatively short time-scale and not on the overall excitability of the neurons. These findings indicate that theoretical models of reaction-time performance should consider correlations between subsets of the neuronal populations in motor cortex, in addition to traditional firing rate coding, as a means of representing motor performance. In this context, our data support with recent studies that have implicated correlated neuronal activity in the sensorimotor cortex of behaving mammals in representing behaviorally relevant information (Murthy and Fetz, 1992; Nicolelis et al., 1995; Seidemann et al., 1996; Riehle et al., 1997; Donoghue et al., 1998; Hatsopoulos et al., 1998).

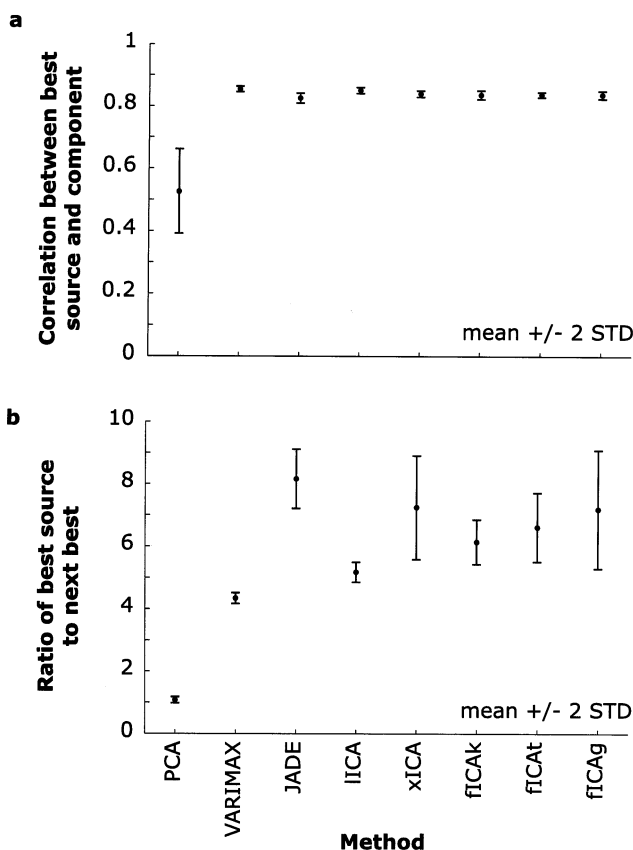


Fig. 7. The ability of the different PCA- and ICA-based methods to account selectively for a single source of common input are shown. (a) The average maximum correlations between each source and each component are shown for eight methods. All of the ICA-based methods and varimax rotation of the principal components were related to single sources with correlation coefficients around 0.8. By contrast, the maximum correlations for the principal components were much lower. (b) The ratio between the maximum correlation between each source and each component and the next largest correlation are shown. This is a measure of the selectivity of the method to resolving a single source of common input. All ICA-based methods were more selective than both PCA and varimax rotation of PCA.

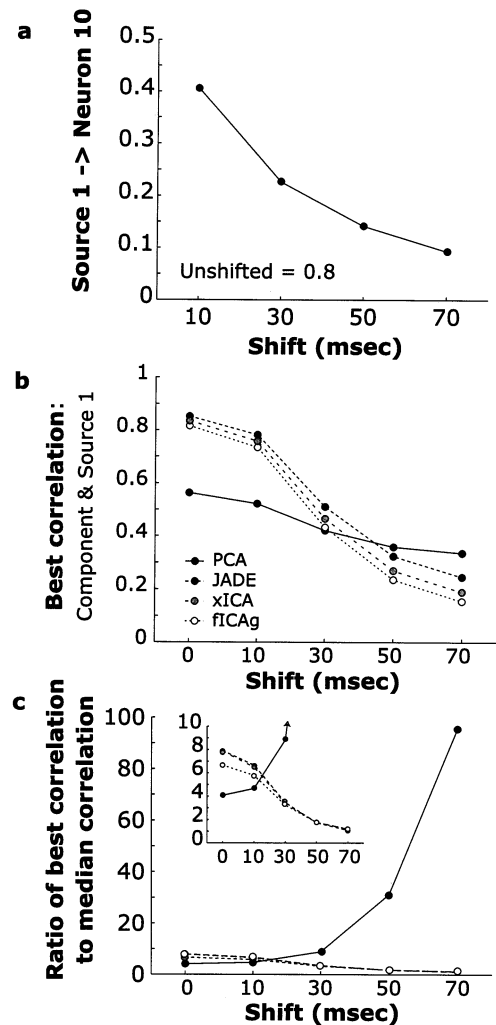


Fig. 8. The shift test produced degradations in the correlations between sources, neurons, and population vectors that were dependent on the range of time over which spikes were shifted. (a) Correlations between one of the simulated neurons, # 10, from one simulated data set and source 1 are shown for spike shifting over 10, 30, 50, and 70 ms. The correlation between the neuron and the source rapidly degraded as the time-scale of shifting was increased. (b,c) The average maximum correlation between population vectors and source 1 and for the ratio of the maximum correlation to the median correlation are shown for these same time-scales of shifting and for PCA and three ICA-based methods (JADE, xICA, and fICAg). The inset panel in (c) shows the ratio data on an expanded scale to better illustrate the differences between the three ICA-based methods.

5. Summary and discussion

Neural assemblies are thought to arise when neurons are dynamically arrayed into groups to convey and process information in a cooperative manner. This form of neuronal processing is based on the correlated time structure of individual spike trains. Though the time scale in which neurons can form assemblies is still a matter of debate (Abeles, 1991; Shadlen and Newsome, 1994; Softky, 1995; Konig and Engel, 1996; Shadlen and Newsome, 1998; Nicolelis et al., 1999), analyses of

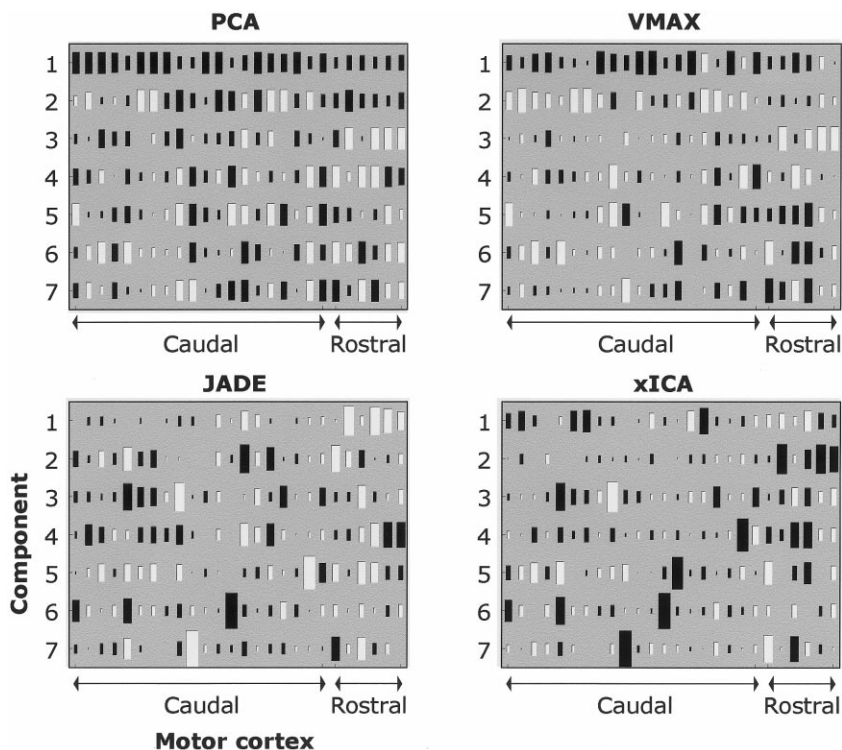


Fig. 9. Different methods for multivariate dimension reduction produced very different spatial distributions of neuronal interactions across simultaneously recorded neuronal ensembles from the rat motor cortex. This data set contained 26 neurons, with 18 neurons in the caudal forelimb area and six neurons in the rostral forelimb area. There were seven eigenvalues larger than 1 for this ensemble of neurons. Coefficients for these seven functions are shown for PCA, varimax rotation of the coefficients from PCA (VMAX), and for two ICA-based methods, JADE and xICA.

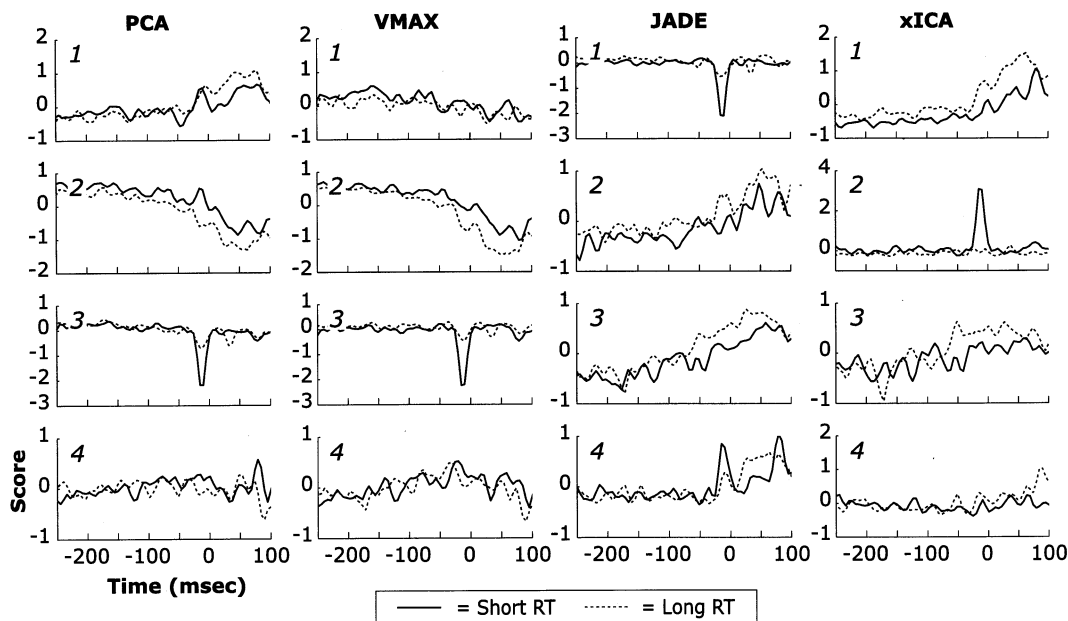


Fig. 10. These plots are scores for population vectors defined with PCA, varimax rotation of the coefficients from PCA (VMAX), and two ICA-based methods, JADE and xICA. The average scores for trials with short reaction-times are shown as solid, dark lines and those for trials with long reaction-times are shown as dashed, light lines. The reference event (time = 0) is the moment when the rats released the response lever within 500 ms of the onset of the trigger stimuli.

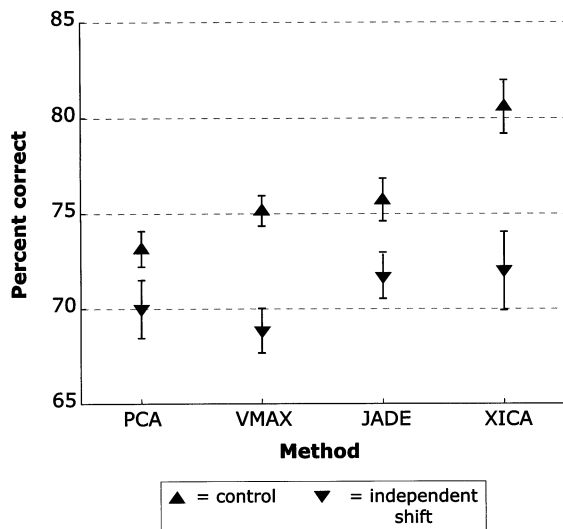


Fig. 11. More information about reaction-time performance was available from population vectors defined by the ICA-based methods than was by PCA or varimax rotations of the principal components. This figure shows the results of using scores for several of the dimension reduction methods to train artificial neural networks to discriminate between trials with short and long reaction-times based on the activity of 26 neurons in the rat motor cortex. Scores derived with xICA discriminated significantly more trials correctly than did scores derived with any other method (ANOVA: $P < 0.01$).

simultaneously recorded neurons all center on defining the correlational structure between spike trains under the null hypothesis that N neurons fire independently. After correcting for changes in individual spike rate modulations, pair-wise and population analyses seek to quantify any residual correlations that may indicate the presence of cell assemblies. A general limitation to such correlational analyses is that, while some sort of time averaging is required to estimate spiking probabilities, spike rates of individual neurons are known to be non-stationary. An increasing number of analyses applicable to neuronal ensemble recordings have been developed that address this limitation, leading to an array of methods, each of which have various strengths and weaknesses.

Cross-correlational analyses seek to establish a functional relationship between the firing times of one neuron with respect to that of another neuron; the principle method of which has been the construction of the cross-correlation histograms (CCH) (Perkel et al., 1967). In the classical CCH, spike timing differences between the 'reference' spike train and the 'target' spike train are constructed without respect to the stimulus cycle. To normalize for stimulus induced changes in firing rates between the two neurons, the 'shift predictor' is subtracted from the CCH to remove any stimulus-locked activity. Alternatively, one can use the joint peristimulus time histogram (Aersten and Gerstein, 1989) to determine the temporal structure of correlated

firing between two neurons with respect to an external event. More recently, unitary event analysis has been proposed to detect correlations between pairs of neurons that need not be time locked to the occurrence of an external event (Riehle et al., 1997 and also in this issue).

While pair-wise analyses are sensitive to determining the functional relationship between two neurons and some event, the major limitation to such analyses arises from the combinatorial explosion of paired-combinations needed to understand the inter-relationships between even a small ensemble of simultaneously

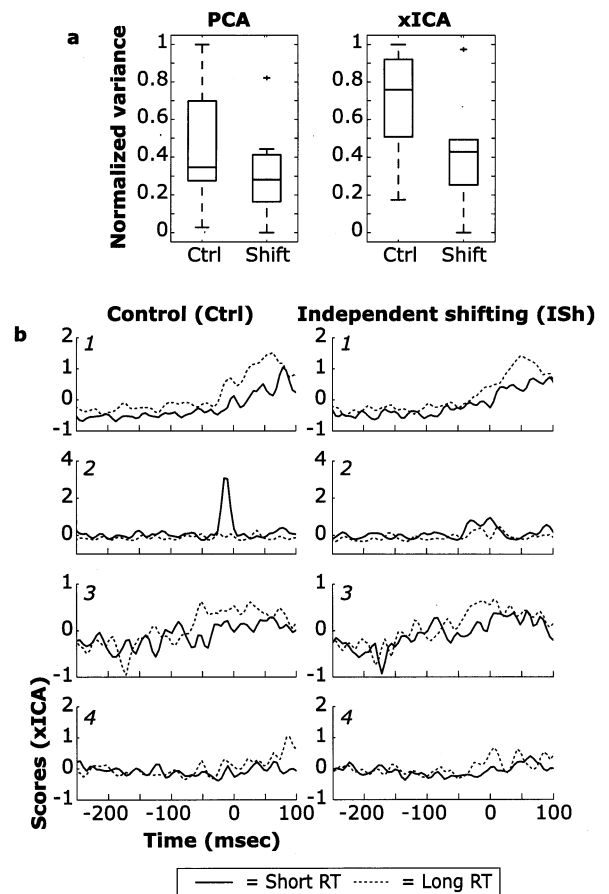


Fig. 12. Spike shifting reduced the variance of the scores for the population vectors and eliminated important aspects of the responses of the population vectors that discriminate between trials with different reaction-times. (a) The variance of the scores for the population vectors defined by ICA-based methods was reduced more than that of population vectors defined by PCA. This was due to the loss of correlated neuronal firing after shifting the spike trains. The data shown in the figure are from spike shifting over 10 ms epochs. (b) Averaged responses for population vectors defined by ICA-based algorithms discriminated between trials with different reaction-times. This discriminative information was lost after spike shifting, which produced decreases in the degree to which the average signals for trials with short and long reaction-times differed. For example, the second xICA shown in this figure exhibited an epoch of correlated firing approximately 50 ms prior to lever release on trials with short reaction-times and not on trials with long reaction-times. This peak was not apparent in the averaged histograms after spike shifting.

recorded neurons. As methods for recording from ensembles of neurons improve and the number of simultaneously recorded neurons continue to increase, the need for more appropriate methods for data analysis becomes even more apparent.

Other methods that go beyond pairwise interactions have been developed: gravitational clustering (Gerstein and Aertsen, 1985), spike pattern search algorithms (Abeles and Gerstein, 1988), and hidden Markov models (HMM) (Gat and Tishby, 1993; Radons et al., 1994; Seidemann et al., 1996). Gravity analysis is a population technique that treats all the neurons of the recorded ensemble as a single entity, clustering neurons based on synchronous firing. The typical ‘read-out’ for this method, however, is based on correlations between pairs of neurons. Spike pattern search algorithms allow for the detection of precise, reoccurring spike patterns across neurons and time to be detected and identified. But, by design, these methods do not provide trial-by-trial estimates of the detected patterns and so they do not allow for quantification of information represented in the identified patterns. On slightly broader time scales (i.e. tens of milliseconds), HMMs have been applied to detect concomitant firing-rate modulations between members of the recorded neuronal ensemble. The goal of these analyses is to determine the variety of firing patterns exhibited across an ensemble of neurons (Seidemann et al., 1996). Thus, HMMs appear to answer very different questions about correlated neuronal firing than those addressed by the other methods described above and in this paper.

Another approach to the problem of correlated neuronal firing was taken by Nicolelis and Chapin (1994) when they applied multivariate statistical methods for dimension reduction, such as PCA, to neuronal ensemble spike train data. The major advantage of this approach over the cross-correlational and pattern analysis methods described above is that the method explicitly provides a representation of correlated neuronal firing on a bin-by-bin and trial-by-trial basis. This allows for quantitative estimates of the degree to which correlated neuronal firing, as represented by PCA or ICA, may vary with the behavioral performance of an animal. Moreover, it has been shown recently that these signals can be used to control external devices such as robotic arms (Chapin et al., 1999). Thus, in addition to having theoretical applications in determining the types of neuronal codes that might arise through neuronal ensemble interactions (e.g. Laubach and Nicolelis, 1998), multivariate statistical methods have great potential for practical applications in clinical neurophysiology.

Here, we compared how PCA- and ICA-based methods detected and reconstructed epochs of correlated firing across ensembles of simultaneously recorded cortical neurons. We observed that while PCA was able to

resolve a variance-weighted version of the average signal across the ensemble, it was unable to identify sparsely distributed interactions across the ensembles. By contrast, each of the methods for ICA that was examined was able to account for sparsely distributed interactions. Therefore, it appears that PCA is an excellent method for accounting for the overall excitability and broad correlations between large groups of neurons. By contrast, the ICA-based methods used here are better suited for finding higher-order correlations (i.e. synchronous firing) between the activity of more limited portions of a neuronal ensemble.

Several methods for ICA were compared in this study. For the most part, the different algorithms produced similar results, despite the differences in their underlying nature (i.e. algebraic versus iterative) or the form of the contrast function used (e.g. kurtosis, tansig, Gaussian). The simulations showed that the selectivity of the various methods was roughly equivalent, with the major differences being in the range of variability exhibited by different methods. By contrast, our application of these methods for real neuronal ensemble data found that the xICA algorithm performed better than the other methods. This may be due to the nature in which the xICA network is updated, which is different from the other methods. Prior to xICA, the data series are randomly permuted such that the order of bins is not longer that obtained in the original data set. Thereafter, the network is ‘trained’ on randomly chosen time epochs that are arranged in frames with as many elements as there were bins on a single trial. Thus, the xICA algorithm iterates through various instances of the neuronal data in a way that should overcome any explicit episodes of non-stationarity in the data set. This is not the case for the JADE and fast ICA algorithms, which operate on the data in ‘batch’ mode, with all bins used to estimate the weights for the ICA model. It may be that the xICA procedure overcomes non-stationarities in the neuronal data that limited the success of the other methods. Given that the limited amount of data in the example used in this paper (70 trials) is actually typical of most real behavioral neurophysiological data sets, we suspect that xICA maybe the best choice for future applications of ICA-based methods for analyzing neuronal ensemble data.

Despite the advantages of ICA over cross-correlational methods, the method has a number of limitations. Our experience with ICA has shown that it will work appropriately when the following conditions are met: (1) firing probabilities are relatively low; (2) correlations within sub-ensembles are relatively high; (3) the temporal precision of the neuronal spike trains is high. For example, if several independent sources for neuronal synchronization effect activity in a given subset of a neuronal population (i.e. spatial overlap of multiple sources), then a given independent component may

represent a linear combination of the sources rather than a true independent source (Makeig et al., 1999). Such a situation is likely to arise in most network level studies of brain function, suggesting that more work is needed to further improve methodologies for analyzing the high-dimensional data sets that are routinely acquired using modern methods for neuronal ensemble recording.

The techniques described in this paper may soon be extended to address a genuinely new question for experimental neuroscience: the role of interactions *between* multiple brain areas. For example, if arrays of recording electrodes are placed in several cortical areas that are known to be connected anatomically, then the methods for ICA used in this paper could be applied to each area to find a set of components that account for correlated firing within each cortical area. Statistical pattern recognition methods could then be used to evaluate the contributions of the individual cortical areas to the collective representation of the animal's behavior, such as in a reaction-time task. In addition, other methods, such as the method of directed coherence (Sameshima and Baccalá, 1999) could be used to determine the direction of information flow between multiple cortical networks using a reduced set of variables that are provided by the independent components. The exploration of these issues, especially when applied to simultaneous recordings from multiple areas of cortex, is likely to become one of the main topics of research in systems neuroscience.

Acknowledgements

We thank Drs Terry Sejnowski and Jean-Francois Cardoso for helpful discussions regarding our application of ICA to neuronal ensemble data. Matlab scripts for the ICA methods described in this paper is available at <http://sig.enst.ft/~cardoso> (JADE), <http://www.cnl.salk.edu> (xICA), and <http://www.cis.hut.fi/projects/ica> (fICA). Supported by a grant from DARPA and ONR (N00014-98-1-0676) to MALN and NIH N510597 to ML.

References

- Abeles M. *Corticonics: Neural Circuits of the Cerebral Cortex*. New York: Cambridge University Press, 1991.
- Abeles M, Gerstein G. Detecting spatiotemporal firing patterns among simultaneously recorded single neurons. *J Neurophys* 1988;60:909–24.
- Aersten A, Gerstein G. Dynamics of neuronal firing correlation: modulation of 'effective connectivity'. *J Neurophys* 1989;61:900–17.

- Bell AJ, Sejnowski TJ. An information maximization approach to blind source separation and blind deconvolution. *Neural Comp* 1995;7:1129–59.
- Bell AJ, Sejnowski TJ. The 'independent components' of natural scenes are edge filters. *Vision Res* 1997;37:3327–38.
- Buckheit J, Donoho D. Improved linear discrimination using time-frequency dictionaries. *Proc SPIE* 1995;2569:540–51.
- Cardoso J-F, Souloumiac A. Blind beamforming for non Gaussian signals. *IEE-Proc F* 1993;140:362–70.
- Cardoso J-F. Blind signal separation: statistical principles. *Proc IEEE* 1998;90:2009–20026.
- Chapin JK. Population-level analysis of multi-single neuron recording data: Multivariate statistical methods. In: Nicolelis M, editor. *Methods for Neural Ensemble Recordings*. Boca Raton: CRC, 1999:193–228.
- Chapin JK, Moxon KA, Markowitz RS, Nicolelis MAL. Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex. *Nat Neurosci* 1999;2:664–70.
- Chapin JK, Nicolelis MAL. Principal component analysis of neuronal ensemble activity reveals multidimensional somatosensory representations. *J Neurosci Meth* 1999;94:121–40.
- Comon P. Independent component analysis — a new concept? *Signal Process* 1994;36:287–314.
- Donoghue JP, Sanes JN, Hatsopoulos NG, Gaal G. Neural discharge and local field potential oscillations in primate motor cortex during voluntary movements. *J Neurophysiol* 1998;79:159–73.
- Gat I, Tishby N. Statistical modelling of cell-assemblies activities in associative cortex of behaving monkeys. In: Moody JE, Hanson SJ, Lippmann RP, editors. *Advances in Neural Information Processing Systems*, vol. 5. San Mateo: Morgan Kaufmann, 1993:945–53.
- Gerstein GL, Aertsen AM. Representation of cooperative firing activity among simultaneously recorded neurons. *J Neurophysiol* 1985;54:1513–28.
- Hatsopoulos NG, Ojakangas CL, Paninski L, Donoghue JP. Information about movement direction obtained from synchronous activity of motor cortical neurons. *Proc Natl Acad Sci USA* 1998;95:15706–11.
- Hyvarinen A, Oja E. A fast fixed-point algorithm for independent component analysis. *Neural Comput* 1997;9:1483–92.
- Johnson RA, Wichern DW. *Applied Multivariate Statistical Analysis*, 3rd edition. Engelwood Cliffs, NJ: Prentice-Hall, 1992.
- Kaiser HF. The varimax criterion for analytic rotation in factor analysis. *Psychometrika* 1958;23:187–200.
- Kohonen T. *Self-Organizing Maps*. New York: Springer-Verlag, 1997.
- Konig P, Engel AK, Singer W. Integrator or coincidence detector? The role of the cortical neuron revisited. *Trends Neurosci* 1996;19:130–7.
- Kruger J, Bach M. Simultaneous recording with 30 microelectrodes in monkey visual cortex. *Exp Br Res* 1981;41:191–4.
- Laubach M, Nicolelis MAL. Interactions between sensorimotor cortical and thalamic neuronal ensembles are altered during the acquisition of a reaction-time task. *Soc Neurosci Abstr* 1998; 56.5.
- Laubach M, Shuler M, Nicolelis MAL. Principal and independent component analyses for multi-site investigations of neural ensemble interactions. Paper presented at CNS*98, Santa Barbara, CA, July, 1998.
- Lee TW, Girolami M, Sejnowski TJ. Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural Comp* 1999;11:417–41.
- McNaughton BL, O'Keefe J, Barnes CA. The stereotrode: a new technique for simultaneous isolation of several single units in the central nervous system from multiple unit records. *J Neurosci Methods* 1983;8:391–7.

- Makeig S, Westerfield M, Jung TP, Covington J, Townsend J, Sejnowski TJ, Courchesne E. Functionally independent components of the late positive event-related potential during visual spatial attention. *J Neurosci* 1999;19:2665–80.
- Mosteller F, Tukey JW. *Data Analysis and Regression: A Second Course in Statistics*. Reading, MA: Addison-Wesley, 1977:xvii.
- Murthy VN, Fetz EE. Coherent 25- to 35-Hz oscillations in the sensorimotor cortex of awake behaving monkeys. *Proc Nat Acad Sci USA* 1992;89:5670–4.
- Nicolelis MAL, Chapin JK. Spatiotemporal structure of somatosensory responses of many-neuron ensembles in the rat ventral posterior medial nucleus of the thalamus. *J Neurosci* 1994;14:3511–32.
- Nicolelis MAL, Baccala LA, Lin RC, Chapin JK. Sensorimotor encoding by synchronous neural ensemble activity at multiple levels of the somatosensory system. *Science* 1995;268:1353–8.
- Nicolelis MAL, Ghazanfar AA, Faggin B, Votow S, Oliveira LMO. Reconstructing the engram: Simultaneous, multisite, many single neuron recordings. *Neuron* 1997a;18:529–37.
- Nicolelis MAL, Fanselow EE, Ghazanfar AA. Hebb's dream: the resurgence of cell assemblies. *Neuron* 1997b;19:219–21.
- Nicolelis MAL, Ghazanfar AA, Stambaugh CR, Oliveira LM, Laubach M, Chapin JK, Nelson RJ, Kaas JH. Simultaneous encoding of tactile information by three primate cortical areas. *Nat Neurosci* 1998;1:621–30.
- Nicolelis MAL, Stambaugh CR, Brisben A, Laubach M. Methods for simultaneous multisite neural ensemble recordings in behaving primates. In: Nicolelis MAL, editor. *Methods for Neural Ensemble Recordings*. Boca Raton: CRC, 1999:121–56.
- Nicolelis MAL. *Methods for Neural Ensemble Recordings*. Boca Raton: CRC, 1999.
- Perkel DH, Gerstein GL, Moore GP. Neuronal spike trains and stochastic point processes. II. Simultaneous spike train. *Biophys* 1967;7:419–40.
- Radons G, Becker JD, Dulfer B, Kruger J. Analysis, classification, and coding of multielectrode spike trains with hidden Markov models. *Biol Cybern* 1994;71:359–73.
- Reyment RA, Jöreskog KG. *Applied Factor Analysis in the Natural Sciences*. New York: Cambridge University Press, 1993.
- Riehle A, Grun S, Diesmann A, Aertsen A. Spike synchronization and rate modulation differentially involved in motor cortical function. *Science* 1997;278:1950–3.
- Reitboeck HJ. Fiber microelectrodes for electrophysiological recordings. *J Neurosci Methods* 1983;8:249–62.
- Sameshima K, Baccalá LA. Using partial directed coherence to describe neuronal ensemble interactions. *J Neurosci Meth* 1999;94:93–103.
- Seidemann E, Meilijson I, Abeles M, Bergman H, Vaadia E. Simultaneously recorded single units in the frontal cortex go through sequences of discrete and stable states in monkeys performing delayed localization task. *J Neurosci* 1996;16:752–68.
- Shadlen M, Newsome W. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J Neurosci* 1998;18:3870–96.
- Shadlen MN, Newsome WT. Noise, neural codes and cortical organization. *Curr Opin Neurobiol* 1994;4:560–79.
- Softky WR. Simple codes versus efficient codes. *Curr Opin Neurobiol* 1995;5:239–47.