ISAAC N. BANKMAN

# DETECTION AND CLASSIFICATION OF TRANSIENT SIGNALS: SORTING NEURAL WAVEFORMS

The study of biological neural networks requires reliable extracellular recordings that provide data on the collective behavior of neurons. Extracellular recordings contain the transient waveforms generated by several neurons located at the tip of the electrode. These neural spike waveforms have to be detected and classified to determine the firing times of different neurons. Many recording conditions result in high noise levels that increase the difficulty of neural spike sorting. This article describes the generation of the neural waveforms that are recorded, presents a brief survey of methods available for spike sorting, and describes the two methods found to be best. Development of an optimal spike sorting system to provide a valuable tool for neuroscience is in progress at the Applied Physics Laboratory.

## INTRODUCTION

The expanding applicability of neural networks to diverse engineering problems has increased the significance of neurophysiological investigations that study the collective behavior of neuronal assemblies. It seems clear that significant advances in areas such as pattern recognition, fault tolerance, memory storage and retrieval, speech processing, computer vision, and control will be possible by a better understanding of biological neural systems. This promise has been recognized in the congressional resolution designating the 1990s as the "decade of the brain" and in the subsequent presidential proclamation encouraging appropriate programs. Furthermore, the National Academy of Sciences indicated that neuroscience, considered now a mature discipline, stands at the threshold of a significant expansion because of the clarity of propitious avenues and the availability of adequate approaches. Indeed, guiding concepts, experimental techniques, and analytical tools are available for progress in many fields ranging from ion channels to neural systems studies. As neuroscientists at the National Institutes of Health, the National Science Foundation, the Johns Hopkins Medical School, and other leading universities agree, one of the prerequisites is instrumentation for obtaining reliable recordings of the physiological activity of neuronal assemblies. The Applied Physics Laboratory, in collaboration with the Neuroscience Department of the Johns Hopkins Medical School, is developing an automated system for optimal recognition of neural waveforms, funded by a grant from the National Institutes of Health.

In neurophysiological experiments, the concurrent activity of neurons is recorded with an extracellular electrode that collects data from several neurons in the vicinity of the tip. The main advantage over an intracellular electrode is the ability to record from more than one neuron at the same time, but the extracellular electrode also allows recording without damaging the neurons and enables longer recording periods. The cost of these benefits is the requirement for sorting the interleaved neural spike trains to determine the firing instants of individual neurons. Because of differences in their geometry and the impedances connecting them to the electrode, the depolarization of different neurons is manifested with different transient waveforms in the recording. Typically, the waveform of a given neuron preserves its general shape during a recording period; therefore, the activity of individual neurons can be determined by sorting the different types of neural waveforms. An additional challenge in extracellular recordings is the relatively low signal-to-noise ratios (S/N) that can often occur. The background noise arises mainly from the activity of a large number of distant neurons, resulting in a considerable overlap between the spectra of waveforms of interest and noise. The problem of neural waveform recognition typifies the problem of detection and classification of transient patterns embedded in colored noise.

On-line neural spike sorting is preferable because immediate feedback on the recording allows better control of experimental conditions and recording quality, reducing the time requirement for the neuroscientist and the animal subject. On-line sorting implies a relatively high computational power, especially for implementations of robust algorithms. Further increase in the data collection capacity can be obtained with multiple electrodes such as the Hopkins probe recently developed at APL[1] using standard integrated circuit techniques. The large amount of data that must be analyzed on-line precludes the use of common sorting techniques that require human supervision. Available integrated circuits for digital signal processing now allow the implementation of an automated system based on effective but computationally intensive algorithms that were not considered feasible a decade ago. This article describes the neural data to be analyzed, presents a brief survey of various methods that have been

suggested, and elaborates on two promising methods: principal components and template matching.

## NEURAL DATA

Sensory or motor information is processed by the nervous system in the form of a distributed representation supported by a large number of nerve cells that are interconnected with excitatory and inhibitory synapses. The functional activity of an individual neuron depends on the strength of the synapses that provide excitation or inhibition from other neurons, and on the activity of those neurons. The fundamental unit of activity is the action potential.

In a nerve cell at rest, the distribution of anions and cations is in a steady state that maintains a resting potential of about –40 to –75 mV on the intracellular surface of the cell membrane, referred to the extracellular surface. This potential difference (polarization) results from a constant diffusion process dictated by the unequal permeability of the cell membrane to $Na^+$ and $K^+$ ions, by the higher concentration of $K^+$ inside and $Na^+$ outside the cell, and by an active $Na^+$–$K^+$ pump that retrieves $Na^+$ from the cell while injecting $K^+$ (Fig. 1). When the overall synaptic effect of the other neurons is excitatory, positive charges are injected into the neuron, and if the membrane potential at the trigger zone (decision-making component of the neuron) exceeds a critical threshold, the permeability of the cell membrane in the trigger zone changes momentarily as a result of an active process (see the boxed insert). This permeability change causes a fast depolarization followed by a fast repolarization; the intracellular membrane potential rapidly rises to about 50 mV and returns to its resting level, producing the action potential across the membrane of the trigger zone (Fig. 2). The duration of this potential spike across the cell membrane is on the order of 1 ms. Under constant excitation, a neuron can fire repeated action potentials at a rate that does not exceed about 500 impulses per second (ips); typical rates are in the range of 10 to 100 ips.

A neuron processes and transmits information to other neurons by generating and conducting action potentials. The action potential, generated at the trigger zone, is conducted along the axon both passively and actively. During the depolarization, the $Na^+$ ions that rush into the trigger zone passively travel a small distance through the conductive intracellular material and excite the membrane of a region adjacent to the trigger zone. The resulting depolarization, in turn, excites a further part of the membrane, causing an active propagation of depolarization that ensures lossless conduction of the action potential. This process generates a current flow away from the synapse, through the axon, and toward the synapse outside the cell (Fig. 3). The current flow outside the cell generates small, transient potential differences across the extracellular fluid in the vicinity of the cell. These potential spikes, typically in the range of 1 to 400 $\mu$V, can be recorded with an extracellular electrode.

The shape of the waveform recorded depends on the position of the electrode relative to the neuron; an electrode near the trigger zone (current sink) records a negative deflection that peaks and decays to zero rapidly with

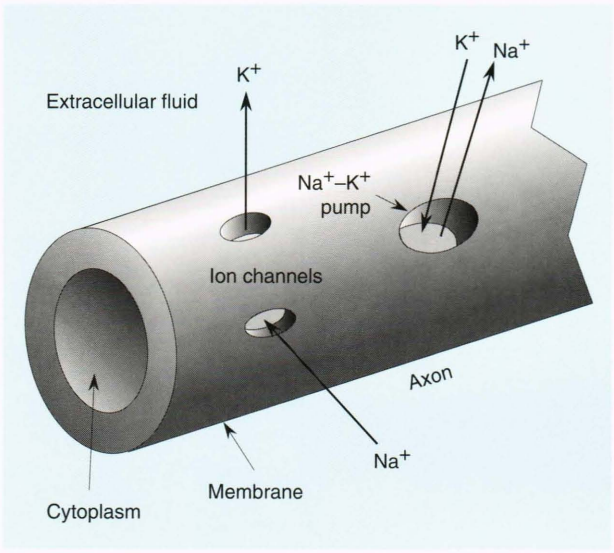



**Figure 1.** When the nerve cell is at rest, the $Na^+$ ions diffuse into the cell and $K^+$ ions diffuse out, as the result of the electrochemical gradients. These passive fluxes are balanced with an active transport of $Na^+$ and $K^+$ ions in opposite directions by the $Na^+$–$K^+$ pump.

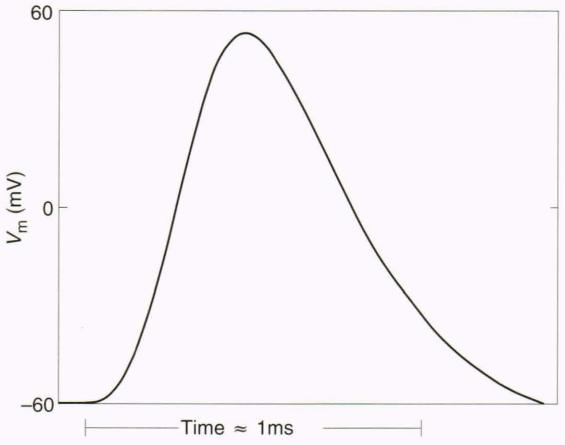



**Figure 2.** The action potential ($V_m$) waveform.

a possible overshoot, whereas an electrode placed at a site sufficiently away from the trigger zone but close to the axon (current source) records a positive peak that subsequently decays to zero with a possible undershoot (Fig. 3). Recording sites close to the dendritic tree produce more complex spike waveforms. Figure 4 shows some examples of extracellularly recorded neural spikes. The voltage level of neural spikes is inversely proportional to a power of the distance from the neuron. Therefore, the recorded data are produced primarily by nearby neurons and, to a lesser extent, by more distant ones. Ideally, the electrode would be placed at a location that provides high-amplitude spikes from several neurons in the vicinity and low-amplitude interference from other neurons. In practice, such adjustments are difficult and

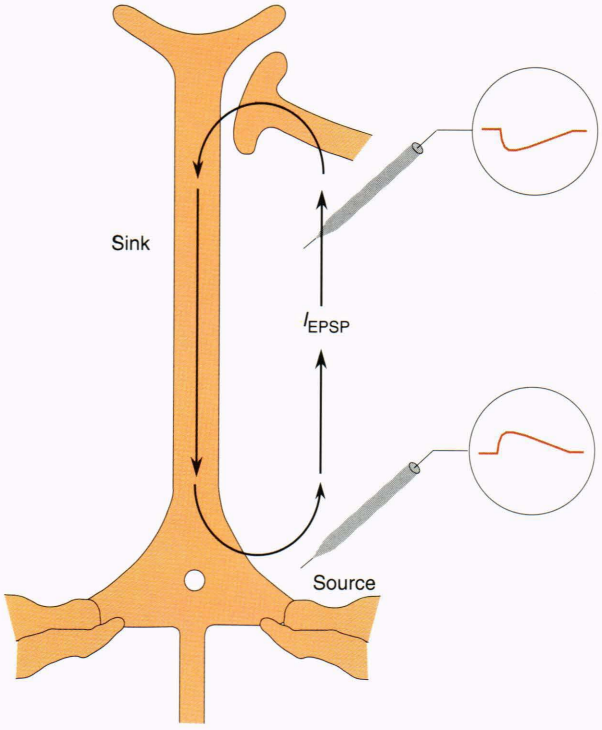



**Figure 3.** Current flow ($I_{EPSP}$) in and around a nerve cell. The extracellular electrode at the top is near the current sink and the extracellular electrode below is near the current source. The recorded waveforms are shown in the inset circles.

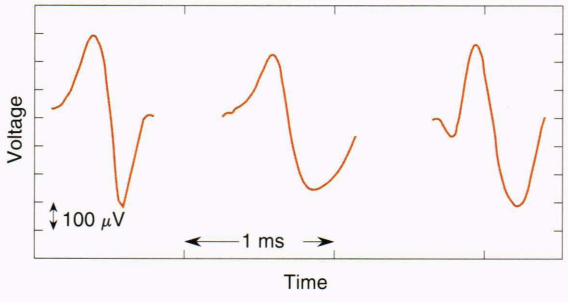



**Figure 4.** Examples of extracellularly recorded neural spike waveforms.

time-consuming, and recordings frequently have very poor signal-to-noise ratios.

Since the values of the membrane capacitance and the axoplasmic resistance do not change considerably from neuron to neuron, the time constant of the depolarization is of the same order of magnitude in most neurons. Therefore, the spectral content of neural spike waveforms is similar in most neurons, and the spectra of signal and noise in recordings overlap to a large extent. Further, since the noise process is primarily made of the accumulation of low-amplitude spikes, the autocorrelation function of noise has significantly high coefficients at lags as large as the average duration of a neural spike (about 1 ms).

## NEURAL SPIKE SORTING METHODS

Sorting neural spikes requires the detection of spikes in the record by discriminating spikes from noise. Then, spikes that have sufficiently similar shapes have to be identified and classified to indicate the firing times of the corresponding neuron. Frequently, two neurons fire at very closely spaced times and their spikes overlap; the resulting waveform may have a shape that is very different from individual spikes. Such spike superpositions need to be resolved to avoid loss of data, especially when many neurons are of interest and when they are firing at high rates.

Window discriminators, the first instruments for sorting neural spikes, are based on specialized hardware that separate different spike types by using their peak amplitudes. The availability of computers to neurophysiologists introduced a large number of algorithms for neural spike sorting.

Two pioneering interactive computerized systems[2,3] operating off-line on tape-recorded data allowed the user to form a template for each spike type by averaging several spikes of the same type. Classification was based on the mean-square difference between a candidate spike and the templates. An attempt was made in Ref. 3 to determine whether unclassified spikes were linear superpositions of two or more known spike types. In the following decades, several variants of the template matching system were built[4–13] using (1) different similarity measures such as cross-correlation, Euclidean distance, and city block distance; (2) different sampling rates ranging from 4 to 50 kHz; and (3) different numbers of samples (ranging from 5 to 69) to represent a spike.

Other methods were based on extracting a small number of features from the spike waveforms, such as peak amplitude, peak-to-peak amplitude, various duration measures, spike area, and root-mean-square value.[14–19] Such approaches may be satisfactory only when little noise is present in the data.

Feature extraction has also been applied using the principal-components method based on a set of basis functions derived from the data. An early implementation of the principal-components method used tapped delay line filters for computing the projection of the signal on the first three principal components;[20] further reports emphasized both the theoretical and practical aspects of principal components.[21–26]

Other methods of neural spike sorting that have been studied include the first eight Fourier coefficients of the spike,[27] regression coefficients in fitting a curve to the spike,[28] and estimation of the conduction velocity with multiple electrodes.[29–31]

Comparison of the performance of the suggested neural spike sorting methods has been difficult because different data were used by different investigators; a normative standard annotated database for neural spikes is not available as it is in electrocardiography. An "apparent separation matrix" was proposed for evaluating the quality of neural spike sorting techniques.[32] The classification power of a set of parameters was determined by computing a measure of separation (dissimilarity) among all spike types. The measure of separation, computed for

## CHANGES IN CELL MEMBRANE PERMEABILITY

Besides the passive ion channels, the nerve cell membrane has voltage-sensitive $Na^+$ and $K^+$ channels that can be either on or off (the density of $Na^+$ channels is 35 to 500 channels per square micrometer). When open, each channel passes a pulse of current with a variable duration but a constant amplitude. The voltage-sensitive properties of these channels cause highly nonlinear and different membrane conductances, $g_K$ and $g_{Na}$, for $K^+$ and $Na^+$ ions, respectively. Both $g_K$ and $g_{Na}$ respond to a depolarizing potential, and both have a greater response to larger depolarizations; however, they differ in their rate of onset and offset as well as in their inactivation. Experiments done by clamping the voltage across the membrane to a fixed level for a brief period show the time course of the response of $g_K$ and $g_{Na}$ to a voltage pulse that simulates depolarization (Fig. I). The rise and fall times of $g_{Na}$ are shorter than those of $g_K$. Furthermore, if the voltage pulse is short, both $g_K$ and $g_{Na}$ return to their original (rest) value when the pulse ends (corresponding to repolarization of the neuron); if the pulse is longer, $g_{Na}$ returns to a baseline value when the voltage pulse is on (during depolarization), whereas $g_K$ reaches a plateau and remains at the level for the duration of the pulse. These voltage-sensitive nonlinear dynamic properties of the membrane dictate the waveform of the action potential during depolarization.

The initial depolarization of the membrane caused by synaptic excitation first increases $g_{Na}$ and causes an inward $Na^+$ current. The current increases the depolarization (Fig. II), which results in additional inward $Na^+$ flow. This positive feedback generates the rise of the action potential. The avalanching depolarization is rapidly compensated by two factors: (1) after about 0.5 ms, $g_{Na}$ starts to decrease, which reduces the inward $Na^+$ current, and (2) the increase in $g_K$ results in an outward $K^+$ flow that reduces the depolarization. The decreasing depolarization lowers $g_K$ and $g_{Na}$ further, and the action potential gradually decreases (Fig. II). After about 1 ms, the membrane potential returns to its resting value. A hyperpolarization follows the action potential in most neurons because $g_K$ decays to its initial value gradually after the depolarizing potential is removed. The sustained outward $K^+$ flow is higher than the resting value; this condition increases the negativity of the cell, and the hyperpolarization undershoot occurs (Fig. II). The number of voltage-sensitive $K^+$ channels that remain open decays to zero in the next few milliseconds, and the neuron returns to its resting state.

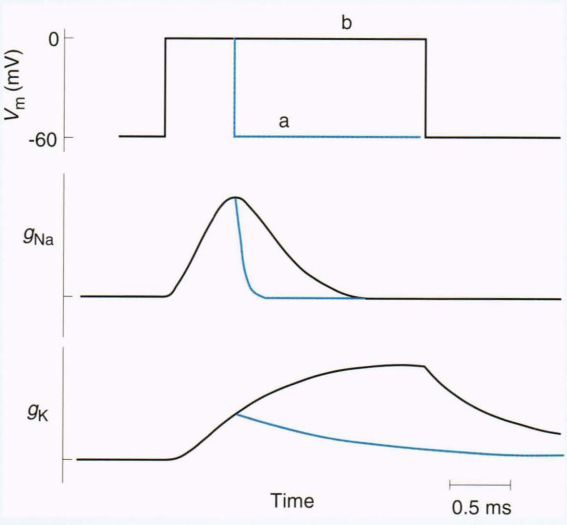



**Figure I.** When the depolarization pulse is short (a), both $g_K$ and $g_{Na}$ (blue curves) return to the initial levels at the end of the pulse (repolarization). When the pulse is longer (b), $g_{Na}$ decays to its initial value during the pulse, whereas $g_K$ reaches a plateau; $g_K$ decays to its initial level when the pulse is over. $V_m$ is the action potential, $g_{Na}$ is the membrane conductance for $Na^+$ ions, and $g_K$ is the membrane conductance for $K^+$ ions.

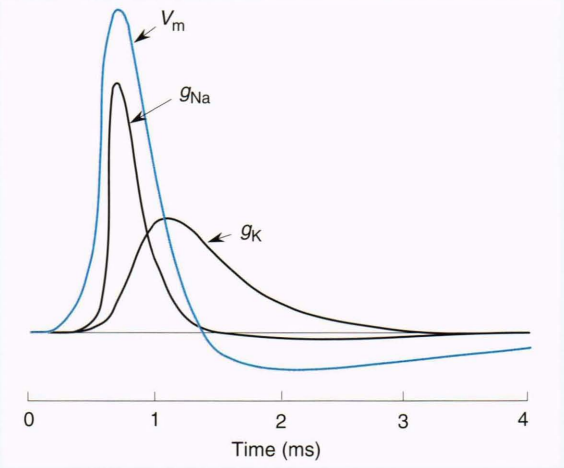



**Figure II.** Time course of $g_K$ and $g_{Na}$, and the resulting action potential $V_m$ during depolarization. $g_K$ is the membrane conductance for $K^+$ ions, and $g_{Na}$ is the membrane conductance for $Na^+$ ions.

each pair of spike types, was the rms value of the difference between the cluster centers of the two types, in parameter space. Using this evaluation approach, the classification performances of peak amplitude, the first principal component, and the conduction latency were compared.[33] In a comprehensive evaluation,[34] the classification performances of peak amplitude, conduction latency, a combination of peak amplitude and conduction latency, 32-sample template matching, and principal components were compared. The evaluation concluded that, in recordings with high values of S/N, peak amplitude and conduction latency were adequate, but robust classification with minimal noise sensitivity required principal components or template matching.

The general approach, which is common to most methods mentioned, consists of the following: (1) form-

ing a vector of parameters that represents each spike, (2) characterizing the different clusters formed by different spike types, and (3) setting decision boundaries used for classification of subsequent spikes.

In the principal-components method, the representative parameters are the projections of the data on a few optimal, orthogonal basis vectors, whereas in the template-matching method, the representative parameters are the consecutive samples digitized on a spike.

Characterization of the clusters in parameter space can be achieved by analyzing an initial segment of the data with a simultaneous unsupervised clustering method.[35,36] This approach provides reliable clusters, but it introduces a delay in the pattern-recognition process. The duration of this delay depends on the rate of arrival of patterns, the clustering algorithm, and the computer speed. If classification has to start with the first pattern, a sequential unsupervised clustering algorithm can be used to circumvent the learning delay. Under relatively high S/N conditions, and when the pattern types are not very similar, this algorithm generates the same clusters as the simultaneous-clustering method.

Decision boundaries either can be placed manually or can be implied by the use of a distance metric and an appropriate acceptance threshold. The distance and its threshold depend on the probability density of the various classes.

Theoretically, optimal classification is achieved by using the digitized samples as the representative parameters, but the principal-components method allows operation in a much lower dimensional space. Some preprocessing of data is required, however, in the principal-components method.

## Principal Components

The principal-component method, also known as the Karhunen–Loeve Transform (KLT) in the signal-processing field, uses the statistical properties of the data to determine a set of orthonormal basis vectors that can be used for effective data reduction. If each waveform is represented by an $N$-dimensional column vector $\mathbf{x}$, the covariance matrix of the waveform vectors, $\mathbf{C}$, is defined as

$$\mathbf{C} = E\{(\mathbf{x} - \mathbf{m})(\mathbf{x} - \mathbf{m})'\} , \tag{1}$$

where

$$\mathbf{m} = E\{\mathbf{x}\} \tag{2}$$

is the mean vector, $E\{\ \}$ is the expected value operator, and the prime symbol ($'$) denotes transposition. The transformation matrix $\mathbf{T}$ of the KLT is a square matrix whose rows are the eigenvectors of $\mathbf{C}$. The KLT transform $\mathbf{y}$ of a vector $\mathbf{x}$ is obtained as follows:

$$\mathbf{y} = \mathbf{T}(\mathbf{x} - \mathbf{m}) , \tag{3}$$

and the covariance matrix of the transformed vectors

$$\mathbf{L} = \mathbf{TCT}' \tag{4}$$

is a diagonal matrix whose elements are the eigenvalues $\lambda_i$ of $\mathbf{C}$. These eigenvalues are indexed in order of decreasing size, and each eigenvalue is the variance of the transformed data along the corresponding eigenvector. The inverse transform, used to reconstruct $\mathbf{x}$ from $\mathbf{y}$, is equal to the transpose of $\mathbf{T}$:

$$\mathbf{x} = \mathbf{T}^{-1}\mathbf{y} + \mathbf{m} = \mathbf{T}'\mathbf{y} + \mathbf{m} . \tag{5}$$

For an arbitrary set of data, the existence of the transformation matrix $\mathbf{T}$ is guaranteed because $\mathbf{C}$ is a real, symmetric matrix.

The effective data-reduction property of the KLT, used also in image coding, results from the fact that most of the information is often concentrated in the components of $\mathbf{y}$ that are associated with the largest eigenvalues. If a reduced transformation matrix $\mathbf{R}$ is built with only the $M$ principal eigenvectors with largest eigenvalues, the mean-square error between the original vector $\mathbf{x}$ and its approximation,

$$\mathbf{a} = \mathbf{R}'\mathbf{y} + \mathbf{m} , \tag{6}$$

is given by

$$e = \sum_{i=M+1}^{N} \lambda_i , \tag{7}$$

which is the sum of the eigenvalues of the unused eigenvectors. Since this error is minimized by selecting the eigenvectors with largest eigenvalues, the KLT transform is optimal in the least-square-error sense. Although the matrix $\mathbf{T}$ (or $\mathbf{R}$) is not separable, a fast KLT algorithm based on the fast Fourier transform has been developed.[37]

The minimal number of eigenvectors required for an adequate representation of the signal with reduced dimensionality depends on the data set. In the neural spike sorting applications, waveforms initially composed of a large number of samples (20 to 40) have been represented successfully by two coefficients for purposes of detection and classification.[25,26] In general, more than 95% of the pattern's energy is contained in the first two KLT coefficients, so as to allow very fast processing, as

well as monitoring of the data and the decision boundaries on a two-dimensional display.

## Optimal Classification

When the distribution of noise amplitude is Gaussian (as it is in neural recordings), the optimal Bayesian classification can be achieved by computing the mean of each cluster and by setting a decision boundary around each mean with a distance metric that depends on the covariance matrix of noise.

The probability density $p(\mathbf{x})$ of a multivariate Gaussian distribution is

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{N/2}|\mathbf{C}|^{1/2}} \exp[-(\mathbf{x} - \mathbf{m})'\mathbf{C}^{-1}(\mathbf{x} - \mathbf{m})/2] \ , \quad (8)$$

where $\mathbf{x}$ is the $N$-dimensional parameter vector, $\mathbf{C}$ is the covariance matrix of $\mathbf{x}$, $|\mathbf{C}|$ is the determinant of $\mathbf{C}$, and $\mathbf{m}$ is the mean vector. The quantity

$$d_{\mathrm{M}}^2 = (\mathbf{x} - \mathbf{m})'\mathbf{C}^{-1}(\mathbf{x} - \mathbf{m}) \quad (9)$$

is called the squared Mahalanobis distance between $\mathbf{x}$ and $\mathbf{m}$. In the multiclass problem, each class has its own probability density.

Let the data have $K$ different classes represented by $\omega_i$, with $i = 1, \ldots, K$. Multiclass Bayesian classification is performed with discriminant functions based on the class densities and *a priori* probabilities of the classes. A convenient choice of discriminant function is

$$g_i(\mathbf{x}) = \log[p(\mathbf{x}|\omega_i)] + \log P(\omega_i) \ , \quad (10)$$

where $g_i(\mathbf{x})$ is the discriminant function, $p(\mathbf{x}|\omega_i)$ is the probability density of class $i$, and $P(\omega_i)$ is the *a priori* probability of class $i$. The class to which a candidate pattern belongs is determined by computing the values of the discriminant function using the pattern's vector $\mathbf{x}$ for each class. The pattern is assigned to the class with the highest discriminant function value. When each class has a multivariate Gaussian distribution, the expression for $g_i(\mathbf{x})$ becomes

$$g_i(\mathbf{x}) = -(\mathbf{x} - \mathbf{m}_i)'\mathbf{C}_i^{-1}(\mathbf{x} - \mathbf{m}_i)/2$$

$$- (N \log 2\pi)/2 - (\log|\mathbf{C}_i|)/2 + \log P(\omega_i) \ . \quad (11)$$

This expression can be further simplified for the following reasons:

1. The $(N \log 2\pi)/2$ term is the same for each class and can be ignored.

2. In applications where each class of signal is corrupted by the same noise process, the covariance matrix $\mathbf{C}_i$ is the same for all classes and the $(\log|\mathbf{C}_i|)/2$ term can be dropped.

3. When the *a priori* probability of each class is known to be the same, or when it is unknown and is assumed to be the same, the $\log P(\omega_i)$ term can be ignored.

Disregarding the division by 2 in the remaining term, the discriminant function becomes

$$g_i(\mathbf{x}) = -(\mathbf{x} - \mathbf{m}_i)'\mathbf{C}_i^{-1}(\mathbf{x} - \mathbf{m}_i) \ . \quad (12)$$

Therefore, under the stated conditions, the Bayesian approach consists of classifying the candidate pattern in the class, to the mean of which the Mahalanobis distance is lowest. Constant–Mahalanobis distance contours are ellipses centered around the mean of each class; they coincide with equal-density contours on the multivariate distributions (Fig. 5A). In most applications, it is desirable to have the option of leaving some patterns unclassified or rejecting them. To leave patterns unclassified, only patterns that have a distance (or distances) below an acceptance threshold are classified, which is equivalent to setting an elliptical decision boundary around the mean of each class.

None of the reported neural spike sorting applications have used the Mahalanobis distance in classification, because of its computational burden: it requires (1) estimation of the covariance matrix and the inversion thereof, and then (for each class) (2) a matrix-vector multiplication and a dot product. The Euclidean distance, which has been commonly used, is equivalent to setting a circular decision boundary and provides suboptimal results. The extent of performance loss because the Euclidean distance is used depends on the covariance matrix of noise, the noise level of the data, and the similarity between different classes. If the noise level is relatively low and the clusters of different classes are sufficiently apart, the Euclidean distance can provide satisfactory results regardless of the covariance matrix of noise. Our recent studies, using 32-sample templates and five different neural spike classes embedded in typical neural recording noise, showed that if the Euclidean distance between the means of the two closest clusters is more than 14 standard deviations of noise, perfect classification can be obtained. But as the clusters get closer or the noise level increases, the performance drops, and the loss (referred to the optimal case) can be up to 30%. The performance drop is caused by the Euclidean distance being compatible with circular density contours (Fig. 5B), but in many applications, such as the neural spike sorting problem, the distributions are elliptical.

In the template-matching approach for neural spike sorting, since the variance of each dimension is the same, elliptical distributions result only from the significant autocorrelation in the noise process. Such elliptical distri-

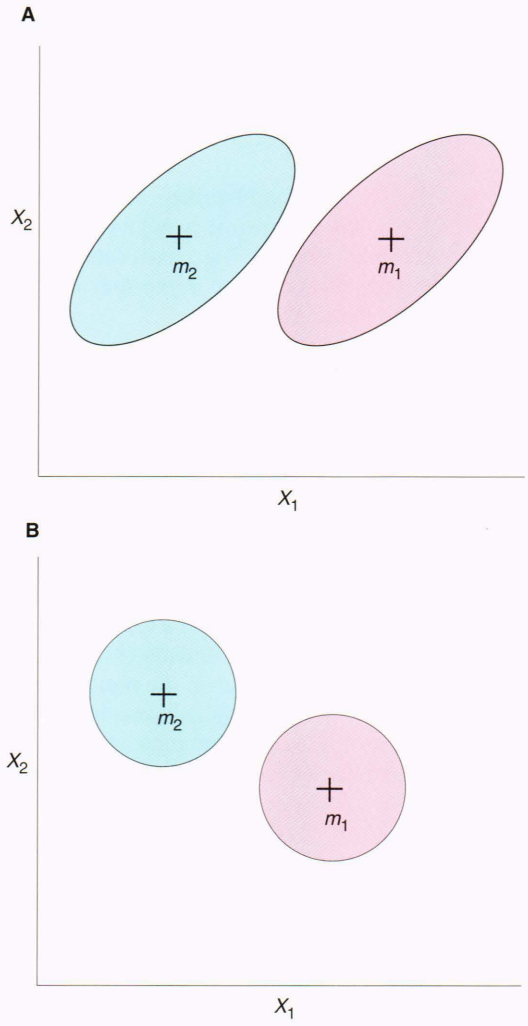



**Figure 5.** Decision boundaries in two-dimensional space spanned by variables $x_1$ and $x_2$; $m_1$ and $m_2$ are the mean vectors of the corresponding classes. **A.** Decision boundaries of a Mahalanobis-distance classifier. **B.** Decision boundaries of a Euclidean-distance classifier.

butions are reflected in the covariance matrix. The extent to which Euclidean distance causes the classification performance to deteriorate depends on the covariance matrix. With Euclidean distance, the higher the autocorrelation in noise, the higher the eccentricity of the elliptical distributions and the lower the performance.

### Whitening

If the density contours were circular, optimal classification could be achieved with the Euclidean distance. Circular density contours occur when the components of the multivariate distribution are uncorrelated, in other words, when the noise process is white. Then the covariance matrix $\mathbf{C}$ is diagonal, with the elements all equal to the variance of noise, $s^2$; its inverse $\mathbf{C}^{-1}$ is diagonal, with elements equal to $1/s^2$; and the discriminant function reduces to

$$g_i(\mathbf{x}) = -(\mathbf{x} - \mathbf{m}_i)'(\mathbf{x} - \mathbf{m}_i) , \qquad (13)$$

which is equivalent to classification with minimal Euclidean distance. We recently reported a method for transforming the correlated noise into white noise that provided theoretically expected optimal results with the Euclidean distance.[38] This method is based on an autoregressive moving-average model of noise and can be implemented on-line with a recursive filter using less than ten coefficients. The best possible classification performance can be achieved with this approach, under any level of noise.

## APPLICATION

The two main tasks in neural waveform sorting are detection of waveforms and their classification. The optimal approach significantly improves the performance of both detection and classification. Detection can be considered a two-way classification of the data as waveform or no-waveform. Therefore, one class contains all waveforms from all neurons, and the other class contains noise alone. In multidimensional space, noise segments of $N$ points each form a hyperspherical cluster around the origin when the data are whitened. When no prior information is available on waveforms, the optimal approach is to use a hyperspherical decision boundary to discriminate waveforms from noise. Waveforms can be discriminated from noise by computing the sum of squares of the $N$ samples and by comparing this power estimate to an appropriate threshold. The threshold determines the radius of the hyperspherical decision shell and can be set by using the statistics of the noise. The guiding criterion in setting the threshold is to obtain no, or very few, false positives (e.g., one per second).

Figure 6 (top) shows a segment of neural recording with a relatively low level of noise. Amplitude discrimination, a widely used detection technique, consists of setting a threshold on the recorded data; it can, in the segment shown, perfectly discriminate noise from waveform. But when the noise level is higher (Fig. 6, middle and bottom), as in many neural recordings, the detection performance of amplitude discrimination drops to about 90% and 30% correct for recordings similar to those of the middle and bottom panels of Figure 6, respectively. On the other hand, the corresponding performance of the optimal detection technique, with $N$ set to 32, is 100% and about 70% correct, respectively.

In the classification phase, a detected waveform represented by a vector of $N$ samples has to be classified in one of the $K$ different classes. Optimal classification is achieved by whitening the data and assigning the waveform to the class that yields the lowest Euclidean distance (Eq. 13) below an acceptance threshold. The classification threshold is set in a manner similar to that of detection. Classification with amplitude is 100%, 65%, and 40% correct on records such as those in the top, middle, and bottom panels, respectively, of Figure 6. The corresponding performance of optimal classification is 100% correct for all three records.

## DISCUSSION AND CONCLUSION

The optimal detection and classification approach using whitening and template matching provides the most

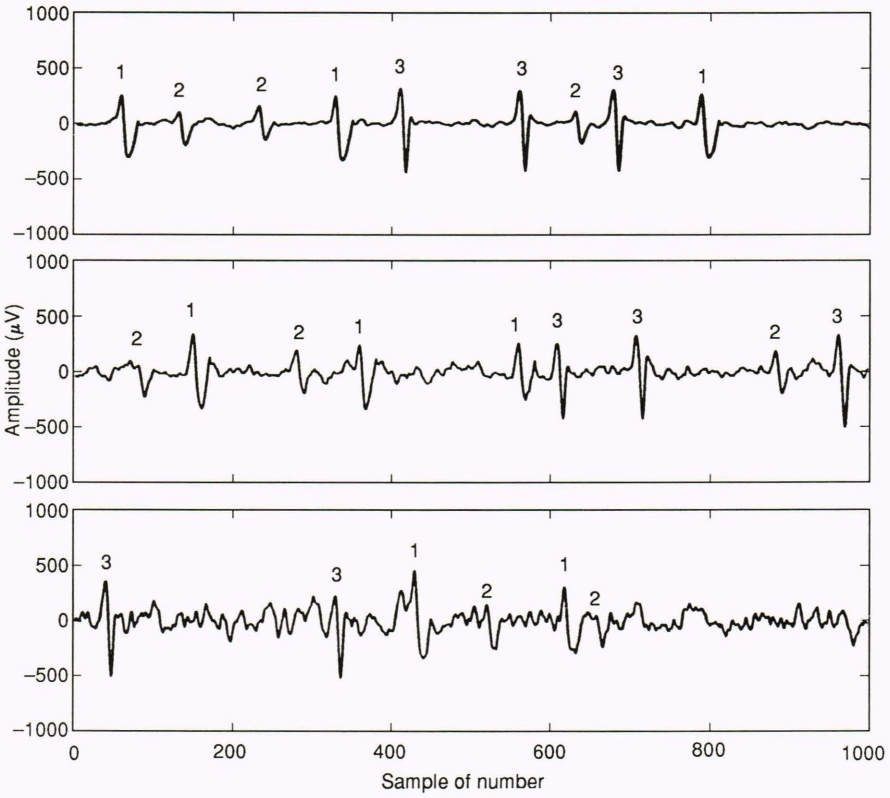



**Figure 6**. Neural recordings at three different noise levels, digitized at 32 kHz. The numbers on the waveforms indicate the neuron that fired.

reliable performance. Previous studies showed that the performance of the principal-components method can be close to that of template matching for neural spike sorting.[34] The advantages of principal components are considerable data reduction and speed, whereas template matching is the theoretically optimal approach. Reliable resolution of spike superpositions requires templates of individual spikes. Resolution of superpositions can be reliably achieved by using an iterative algorithm based on template matching.[38]

It is possible that basis functions other than principal components can provide successful classification. A promising candidate is the wavelet transform,[39] which is especially suited for an efficient representation of transient signals. Since its orthonormal basis functions are independent of the data, the wavelet transform does not require computation of the data covariance matrix and its eigenvectors.

Recently, neural networks have been applied to various pattern-recognition problems. Their ability to generate nonlinear and sometimes disjoint decision regions is the main contribution of neural networks to pattern recognition. In applications where each class has one cluster and when components of patterns are uncorrelated (whitened data), the distribution of each class is spherical around its centroid. In such applications, the required decision boundary is a spherical shell, and it can be implemented simply with a template and the Euclidean distance. The nonlinear discriminatory power of neural networks does not provide an advantage for such applications. If the application precludes satisfactory whitening of the data, then neural networks may contribute to the solution.

Our classification performance studies also included the evaluation of template matching with the city block distance, the sum of absolute valued differences of components of two patterns. The binary equivalent of the city block distance is the Hamming distance. The evaluation showed that the classification performance is only about 5% lower than the performance with the Euclidean distance, and that the increase in performance that results from whitening is the same for both distances. In some applications, the city block distance can be preferable for higher processing speed.

The optimal neural spike sorting system that we are developing is based on template matching with Euclidean distance and a high-speed whitening front end. This approach is applicable to a wide range of signal- and image-processing problems where theoretically optimal detection and classification under heavy noise conditions are desired. We expect the robust on-line operation and full automation of the system to contribute significantly to functional data analysis in the investigation of biological neural networks.

## REFERENCES

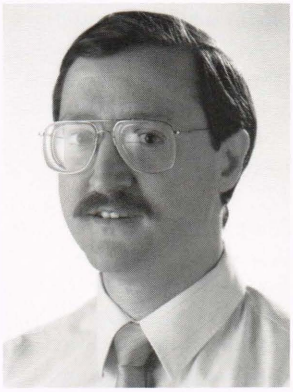
[1] Blum, N. A., Carkhyff, B. G., Charles, H. K., Jr., Edwards, R. L., and Meyer, R. A., "Multisite Probes for Neural Recordings," *IEEE Trans. Biomed. Eng.* **BME-38**, 68–74 (1991).

[2] Gerstein, G. L., and Clark, W. A., "Simultaneous Studies of Firing Patterns in Several Neurons," *Science* **143**, 1325–1327 (1964).

[3] Keehn, D. G., "An Iterative Spike Separation Technique," *IEEE Trans. Biomed. Eng.* **BME-13**, 19–28 (1966).
[4] Prochazka, V. J., "Bioelectric Signal Sorter (JULIA)," in *DECUS Program Library*, pp. 12–35, Digital Equipment Computer Users' Society, Digital Equipment Corp., Marlboro, Mass. (1971).
[5] Prochazka, V. J., Conrad, B., and Sindermann, F., "A Neuroelectric Signal Recognition System," *Electroencephalogr. Clin. Neurophysiol.* **32**, 95–97 (1972).
[6] Capowski, J. J., "The Spike Program: A Computer System for Analysis of Neurophysiological Action Potentials," in *Computer Technology in Neuroscience*, Brown, P. B. (ed.), Hemisphere Publ., Washington, D.C., pp. 237–251 (1976).
[7] Matthews, B., "Identifying Action Potential Waveforms in Neurophysiological Recordings," *J. Physiol.* **277**, 32P–33P (1978).
[8] D'Hollander, E. H., and Orban, G. A., "Spike Recognition and Online Classification by Unsupervised Learning System," *IEEE Trans. Biomed. Eng.* **BME-26**, 279–284 (1979).
[9] Cohen, A., and Landsberg, D., "Adaptive Real-time Wavelet Detection," *IEEE Trans. Biomed. Eng.* **BME-30**, 332–340 (1983).
[10] Studer, R. M., de Figueiredo, R. J. P., and Moschytz, G. S., "An Algorithm for Sequential Signal Estimation and System Identification for EMG Signals," *IEEE Trans. Biomed. Eng.* **BME-31**, 285–295 (1984).
[11] DeLuca, C. J., and Forrest, W. J., "An Electrode for Single Motor Unit Activity during Strong Muscle Contractions," *IEEE Trans. Biomed. Eng.* **BME-19**, 367–372 (1979).
[12] Feldman, J. F., and Roberge, F. A., "Computer Detection and Analysis of Neuronal Spike Sequences," *Infor.* **9**, 185–197 (1971).
[13] Dinning, G. J., and Sanderson, A. C., "Real-time Classification of Multiunit Neural Signals Using Reduced Feature Sets," *IEEE Trans. Biomed. Eng.* **BME-28**, 804–812 (1981).
[14] Harper, R. M., and McGinty, D. J., "A Technique for Recording Single Neurons from Unrestrained Animals," in *Brain Unit Activity During Behavior*, Phillips, M. I. (ed.), C. Thomas, Springfield, Ill., pp. 80–104 (1973).
[15] Radna, R. J., and Vaughn, W. J., "Computer Assisted Unit Data Acquisition/Reduction," *Electroencephalogr. Clin. Neurophysiol.* **44**, 239–242 (1978).
[16] Best, P. J., "An On-line Computer Technique for Display and Analysis of Single Cell Activity," in *Brain Unit Activity During Behavior*, Phillips, M. I. (ed.), C. Thomas, Springfield, Ill., pp. 67–75 (1973).
[17] O'Connell, R. J., Kocsis, W. A., and Schoenfeld, R. L., "Minicomputer Identification and Time of Nerve Impulses Mixed in a Single Recording Channel," *Proc. IEEE* **61**, 1615–1621 (1973).
[18] Dill, J. C., Lockemann, P. C., and Naka, K. I., "An Attempt to Analyze Multiunit Recordings," *Electroencephalogr. Clin. Neurophysiol.* **28**, 79–82 (1970).
[19] Mishelevich, D. J., "On-line Real-time Digital Computer Separation of Extracellular Neuroelectric Signals," *IEEE Trans. Biomed. Eng.* **BME- 17**, 147–150 (1970).
[20] Marks, W. B., "Some Methods of Simultaneous Multiunit Recording," in *Proc. Symp. Information Processing in Sight Sensory Systems*, California Institute of Technology, Pasadena, Calif. (1965).
[21] Friedman, D. H., *Detection of Signals by Template Matching*, Ph.D. Dissertation, The Johns Hopkins University, Baltimore, Md. (1968).
[22] Glaser, E. M., "Separation of Neuronal Activity by Waveform Analysis," in *Advances in Biomedical Engineering, Vol. 1*, Kenedt, R. M. (ed.), Academic Press, New York, pp. 77–136 (1971).
[23] Glaser, E. M., and Marks, W. B., "On-line Separation of Interleaved Pulse Sequences," in *Data Acquisition and Processing in Biology and Medicine*, Vol. 5, Pergamon Press, New York pp. 137–156 (1968).
[24] Gerstein, G. L., Bloom, M. J., Espinosa, I. E., Evanczuk, S., and Turner, M. R., "Design of a Laboratory for Multineuron Studies," *IEEE Trans. Syst. Man. Cybern.* **SMC-13**, 668–676 (1983).
[25] Abeles, M., and Goldstein, M. H., Jr., "Multispike Train Analysis," *Proc. IEEE* **65**, 762–773 (1977).
[26] Smith, S. R., and Wheeler, B. C., "A Real-Time Multiprocessor System for Acquisition of Multichannel Neural Data," *IEEE Trans. Biomed. Eng.* **BME-35**, 875–877 (1988).
[27] Bessou, P., and Perl, E. R., "Response of Cutaneous Sensory Units with Unmyelinated Fibers to Noxious Stimuli," *J. Neurophysiol.* **32**, 1025–1043 (1969).
[28] Remmel, R. S., "A Computerized Discriminator for Action Potentials," *Electroencephalogr. Clin. Neurophysiol.* **56**, 528–530 (1983).
[29] Camp, C., and Pinsker, H., "Computer Separation of Unitary Spikes from Whole-Nerve Recordings," *Brain Res.* **169**, 455–479 (1979).
[30] Roberts, W. M., and Hartline, D. K., "Separation of Multiunit Nerve Impulse Trains by a Multi-Channel Linear Filter Algorithm," *Brain Res.* **94**, 141–149 (1975).
[31] Oguztoreli, M. N., and Stein, R. B., "Optimal Linear Filtering of Nerve Signals," *Biol. Cybernet.* **27**, 41–48 (1977).
[32] Heetderks, W. J., "Criteria for Evaluating Multiunit Spike Separation Techniques," *Biol. Cybernet.* **29**, 215–220 (1978).
[33] Wheeler, B. C., and Heetderks, W. J., "Separation of Cockroach Giant Action Potentials using Multiunit Analysis Techniques," in *Proc. 7th New England Bioengineering Conf.*, IEEE Engineering in Medicine and Biology Society, New York, pp. 310–313 (1979).
[34] Wheeler, B. C., and Heetderks, W. J., "A Comparison of Techniques for Classification of Multiple Neural Signals," *IEEE Trans. Biomed. Eng.* **BME-29**, 752–759 (1982).
[35] Bow, S., "Clustering Analysis and Nonsupervised Learning," in *Pattern Recognition*, Marcel Dekker, New York, pp. 98–153 (1984).
[36] Dubes, R., and Jain, A. K., "Clustering Methodologies in Exploratory Data Analysis," *Adv. Comput.* **19**, 113–228 (1980).
[37] Jain, A. K., "A Fast Karhunen Loeve Transform for a Class of Random Processes," *IEEE Trans. Commun.* **24**, 1023–1029 (1975).
[38] Bankman, I. N., Johnson, K. O., and Schneider, W., "Optimal Detection, Classification, and Resolution of Superpositions of Neural Waveforms," *IEEE Trans. Biomed. Eng.* (in press).
[39] Daubechies, I., "The Wavelet Transform, Time Frequency Localization and Signal Analysis," *IEEE Trans. Infor. Theory* **IT-36**, 961–1005 (1990).

ACKNOWLEDGMENT: The author would like to thank Kenneth O. Johnson and Michael A. Steinmetz of the Neuroscience Department of the Johns Hopkins Medical School and Wolfger Schneider of APL for their collaboration. This work is supported by NIH Grant NS 07226, and partially by Department of the Navy Contract N00089-C-5301.

## THE AUTHOR



ISAAC N. BANKMAN holds B.Sc. and M.Sc. degrees in electrical engineering and a Ph.D. in biomedical engineering. He joined the Biomedical Engineering Department of The Johns Hopkins University as a postdoctoral fellow in 1985 and was a research associate between 1987 and 1990. In June 1990, he joined APL's Eisenhower Research Center as a senior research scientist. His field of interest is signal and image processing, including optimal signal detection and classification, image pattern recognition, neural-network applications, modeling of neural systems, and algorithms for the analysis of biological signals.