

## Editing trains of action potentials from multi-electrode arrays

Richard B. Stein\*, Douglas J. Weber

*Centre for Neuroscience and Department of Physiology, 513 Heritage Medical Research Centre,  
University of Alberta, Edmonton, AB, Canada T6G 2S2*

Received 15 January 2003; received in revised form 12 April 2003; accepted 21 November 2003

### Abstract

When recording from multi-electrode arrays, only a short period around the time of a threshold crossing is generally saved for later analysis. Then, waveforms are often sorted automatically to identify templates of spikes from individual neurons near an electrode. As spikes sum from different neurons and noise is present, some spikes may be missed and others erroneously accepted. This paper describes methods for identifying and correcting errors in recorded spike trains to recover the pattern of spikes from each neuron as faithfully as possible. These methods are complementary to, but distinct from methods to reconstruct waveforms that arise from summation of individual templates that overlap one another. Our methods are based on the local statistics of the firing rates or inter-spike intervals and the methods work best for neurons that fire regularly (small standard deviation relative to the mean interval). First, we test whether accepting more spikes, whose waveforms are close to the templates that have been identified, will increase the regularity or smoothness of the firing rates. Then, after accepting spikes that increase regularity, we test whether individual intervals are sufficiently longer (or shorter) than their neighbors to identify spikes that have been omitted (or accepted) erroneously. The methods are tested on simulated spike trains, where spikes have been inserted or deleted at random, and on spike trains recorded from multi-electrode arrays in dorsal root ganglia of cats walking on a treadmill.

© 2003 Elsevier B.V. All rights reserved.

*Keywords:* Neurons; Spikes; Electrode arrays; Intervals

### 1. Introduction

Increasingly, researchers are recording from multiple electrodes to understand the population coding of groups of nerve cells better (e.g., Maynard et al., 1999; Taylor et al., 2002; Wessberg et al., 2000). A common problem associated with this method is that each electrode may record from several nerve cells and the spikes from these nerve cells can sum in any possible temporal sequence. Also, the noise associated with these recordings further distorts the recorded wave shapes. In behaving animals, the signals may be further contaminated by EMG. Thus, if there are several spikes, all the possible summation patterns of units and noise may be extremely difficult to analyze.

Programs and devices are available that identify clusters of waveforms in a multidimensional space (Fee et al., 1996; Gray et al., 1995; Harris et al., 2000; Jansen and Ter Maat, 1992; Kaneko et al., 1999; Lewicki, 1998; Shoham et al.,

2001). The selection of inclusion boundaries may be difficult because of the presence of noise and the superposition of waveforms, as described above. Independent criteria are needed to decide if particular waveforms should be accepted or not. One goal of this paper is to develop such criteria.

Problems of identifying spikes exist for recordings with single electrodes as well, but an individual electrode can often be moved slightly to accentuate one or two spikes from others that might contaminate the recording. Multi-electrode arrays are typically inserted as a group and are not independently manipulated to select for individual waveforms (Hoogerwerf and Wise, 1994; Rousche and Normann, 1992; but see Hoffman and McNaughton, 2002). In addition, the full time record is generally not recorded for multi-electrode arrays because of the large amount of storage space required. Instead, thresholds are set for each channel and only short periods of data around the threshold crossings are recorded.

Fig. 1 shows simulated examples of the problems that arise. Two distinct waveforms (units 1 and 2) are shown. These simulated waveforms can superimpose so that a much larger waveform results (units 1 + 2). Note that the triggering level will then be crossed at a different time than would

\* Corresponding author. Tel.: +1-780-492-1618; fax: +1-780-492-1617.

*E-mail address:* Richard.stein@ualberta.ca (R.B. Stein).

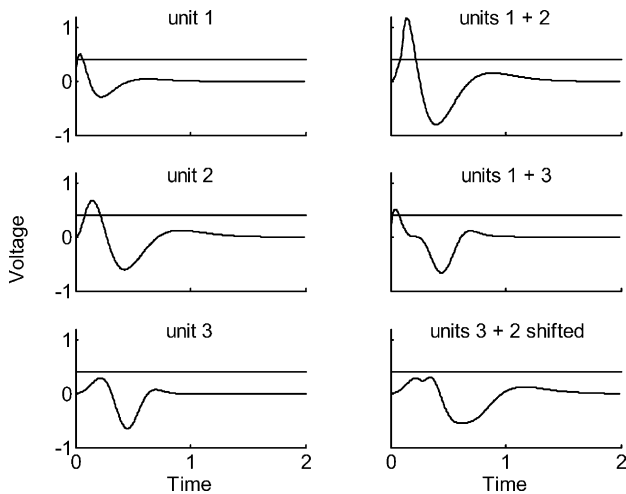


Fig. 1. Three simulated units (left side), two of which reach a set threshold (horizontal line). The units can sum to produce larger units (1 + 2), units with quite different waveforms (units 1 + 3) or prevent a unit from reaching threshold (units 3 + 2 shifted). Further description in the text.

have occurred with either spike alone. Thus, to analyze various possibilities the waveforms must be shifted with respect to each other for all possible intervals. High-speed computers can do this easily enough, but other problems exist. There may be subthreshold units such as 3 that can sum with units such as 1 and 2, and produce apparently new waveforms (units 1 + 3) or prevent spikes from units 1 and 2 from reaching threshold (units 3 + 2 shifted).

The variability in the shape of the waveforms can lead to missed spikes or extra spikes in the record of neural activity. However, additional information about missed spikes or extra spikes is available from analyzing the time series of measured spikes (Fee et al., 1996). A second goal of this paper is to derive statistical criteria for deciding if extra spikes have been mistakenly accepted and if spikes have been missed. These methods are most useful for neurons that fire regularly and change their frequency slowly. A neuron is simulated in Fig. 2A that fires at a rate near 20 Hz. A missed spike occurs at about 250 ms, causing the instantaneous frequency to drop suddenly to half its value near 300 ms. The instantaneous frequency is the inverse of the time since the last spike (inter-spike interval) and the same predictive information is available from the fact that the inter-spike interval suddenly doubles. Fig. 2B shows the opposite example, where an extra spike is included at 220 ms. If the extra spike occurs exactly in the middle of the interval, two intervals will be generated with values of instantaneous frequency double that of neighboring intervals. However, if the extra spike occurs shortly before or after a legitimate spike, one inter-spike interval will be much less than half and another will be much greater than half the neighboring intervals. Similarly, one instantaneous frequency will be much more than double and the other will be much less than double the neighboring values. Clearly, if the firing of a neuron is more random, as is true of many cortical neurons, information about the inter-spike

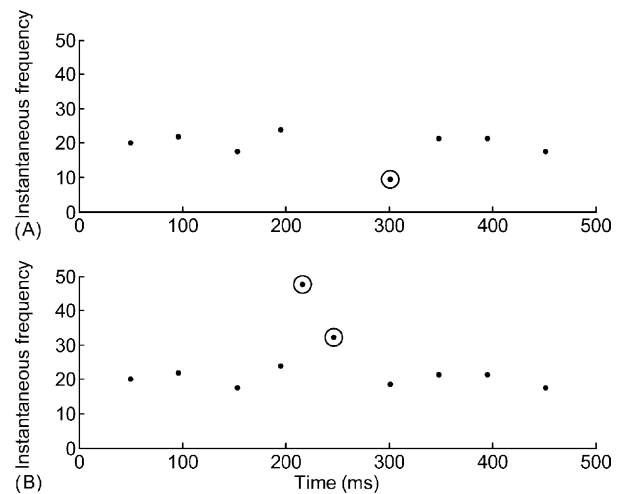


Fig. 2. Simulated spike train in which: (A) a spike is missed or (B) an extra spike is accepted. For each spike the instantaneous frequency (the inverse of the inter-spike interval, in impulses/s) is plotted. The frequency will be approximately half, if a spike is missed (circle). An extra spike will create two short intervals (two circles). Methods are described for identifying these missed or extra events and correcting them.

intervals will be less useful. Similarly, if the instantaneous frequency is changing rapidly, deciding if a spike was added to or omitted from the spike train may be difficult.

The goal of this paper is to develop statistical tests to determine if the deviation from the running average is significant or not. When properly implemented, these tests can be used to identify erroneously accepted or missed spikes in records of neural activity. This paper is organized into four parts: first, we develop methods for automatically editing the spike trains based on statistical criteria applied to inter-spike intervals. Then, we describe statistical analysis of recorded waveforms to identify candidate spikes that should be included in the spike trains. Next, we model the issues involved in inserting and deleting spikes, when the intervals are variable and change over time. Finally, we apply these methods to neurons recorded with a multi-electrode array in a walking cat.

## 2. Methods

### 2.1. Automated insertion and deletion of spikes

Let  $\{sp\}$  be the set of times for all of the classified spikes in a train. If there are  $n + 1$  accepted spikes, there will be  $n$  intervals and the instantaneous frequency  $f$  for the  $j$ th interval will be  $f(j) = [(sp(j + 1)) - sp(j)]^{-1}$ . A logarithmic scale with base 2 is convenient because a halving or doubling of the interval or the instantaneous frequency will produce a change of  $\pm 1 \log_2$  units.

For the  $j$ th interval consider the instantaneous frequencies for a range of four intervals,  $h = \{g(j - 2), g(j - 1), g(j + 1), g(j + 2)\}$  and compute the mean  $m$  and standard

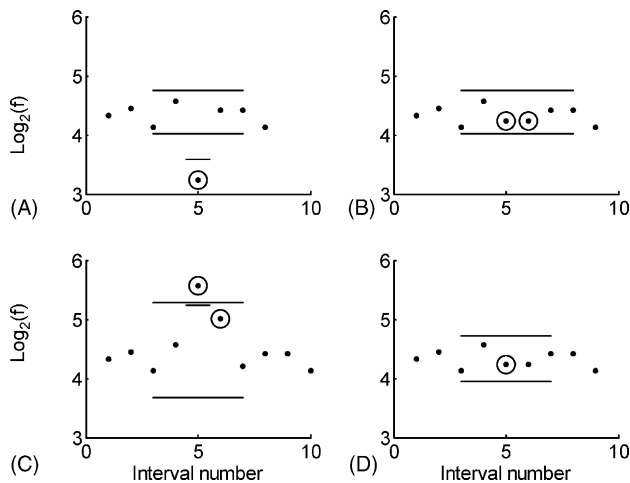


Fig. 3. (A) A missed spike (circle) will produce a lower frequency interval than the standard deviation  $s$  from the neighboring four intervals (long bars represent the mean  $m + 2s$ ). If one spike is missing the frequency should be reduced between about 0.8 (short horizontal bar) and  $1.4 \log_2$  units (see Section 2 for details). (B) Inserting a spike in the middle of the interval produces two intervals that are within the error bounds. (C) An extra spike inserted in the train can produce one or two instantaneous frequencies ( $\circ$  symbols) that are more than  $0.8 \log_2$  units above the mean (short horizontal bar). Since one of the frequencies will be included in calculating  $s$ , the error bounds will increase. (D) Removing one spike reduces the error bounds (long horizontal bars) and leaves one interval that lies within these smaller error bounds.

deviation  $s$  of this range. Note that the  $j$ th interval has been omitted from the set of intervals considered, because we wish to test whether this interval lies within the bounds set by the local statistics. The number of intervals used in the statistics must be sufficient to give reliable estimates of  $s$ , but not so large that adjacent periods are included, where the inter-spike intervals may be changing rapidly. The justification for using four intervals will be given later.

Fig. 3 illustrates the construction of error bounds for identifying and correcting: (A and B) missed or (C and D) erroneous spikes. The long horizontal bars in Fig. 3A represent  $\pm 2s$  units from the mean  $m$  of the simulated train in Fig. 2A. The fifth interval (circled) lies outside these bounds, which should only occur with a probability of  $P < 0.05$ , if the intervals are normally distributed. If one spike is missed, its frequency should be about  $1 \log_2$  unit below the mean  $m$ ; if two are missed, the frequency should be about  $\log_2 3 = 1.57$  units below  $m$ . The short bar in Fig. 3A is at  $m - 0.8$  and represents the upper bound for identifying a missed spike. This value is somewhat less negative than  $-1$  because of the variability in inter-spike intervals, as will be explained in Section 3. Since the circled interval in Fig. 3A lies between  $m - 0.8$  and  $m - 1.4$  units (the corresponding lower bound which is off the graph), we assume that one spike has been omitted. Inserting a spike at the halfway point in the interval now produces two intervals (Fig. 3B) that lie within the error bounds. If the frequency is between 1.4 and 1.9 log units below  $m$ , two spikes are inserted at the  $1/3$  and  $2/3$  points in the interval. If the interval is still longer (the frequency

is lower), more spikes could be inserted. However, the unit may also have paused firing for a period, so the automated procedure does not insert more than two spikes.

To summarize, for each interval  $j$  ( $j = 3: n-2$ , where  $n$  is the number of intervals) the automated procedure inserts one spike at the midpoint of the interval, if two conditions are met:

$$m(j) - g(j) > c_0 s(j), \quad c_1 < m(j) - g(j) < c_2 \quad (1)$$

where  $c_0 = 2$ ,  $c_1 = 0.8$  and  $c_2 = 1.4$  in Fig. 3. These values will be justified later (e.g., Figs. 9 and 10). Since the calculated mean and standard deviations are based on two intervals on either side of the one being tested, two intervals at the beginning and end are not tested and the limits for  $j$  are set to 3 and  $n - 2$ . Two spikes are inserted at one-third and two-thirds of an interval, if:

$$m(j) - g(j) > c_0 s(j), \quad c_2 \leq m(j) - g(j) < c_3 \quad (2)$$

and  $c_3$  is 1.9.

The process described above will only insert new times to fill in missed spikes. Additional testing is required to identify spikes that were included incorrectly (Fig. 2B). If an extra spike was accepted, the instantaneous frequency of at least one interval should be  $\geq 1 \log_2$  unit above the mean. However, because the intervals vary, a number  $< 1.0$  should be chosen. Also, an extra spike will produce two intervals with increased instantaneous frequencies and one of those intervals will be included in the mean. This will erroneously increase the computed mean somewhat and the error bars more markedly (Fig. 3C). A value of 0.8 was used to identify extra spikes, as will be justified in Section 3. To decide which spike should be deleted, three sets of intervals need to be considered:  $h_0 = \{g(j-2) : g(j+3)\}$ ,  $h_1 = \{g_1(j-2) : g_1(j+2)\}$ , and  $h_2 = \{g_2(j-2) : g_2(j+2)\}$ . The set  $h_0$  is the set of frequencies for six intervals, including the two possibly erroneous intervals. The set  $h_1$  contains the five frequencies that would result, if spike  $j$  were deleted, and the set  $h_2$  contains the frequencies, if spike  $j+1$  were deleted. Then, the means ( $m_0:m_2$ ) and the standard deviations ( $s_0:s_2$ ) are calculated for the three sets. Generally, the set with the smallest standard deviation will be the one that does not contain any spurious intervals. In the example of Fig. 3C,  $s_2$  has the lowest value so the spike  $j+1$  is removed. The resulting interval (Fig. 3D) now lies within the bounds calculated for  $s$  in the local region.

To summarize, three conditions are required to delete a spike.

1. Its frequency should be more than  $0.8 \log_2$  units above the mean of its neighbors.
2. The spike to be deleted should minimize  $s$  in the local region.
3. The resulting frequency should be within the error bounds calculated for the local region, after the spike has been deleted.

## 2.2. Data collection

To develop and test algorithms for editing spikes, we used both simulated and actual neuronal spike trains. There are obvious advantages to using simulated spike trains. First, simulations allow for direct specification of the statistical properties of the spike generation (see Section 2.3 below). Second, the simulation provides full control over the corruption of spike trains by adding and deleting spikes. This allowed us to determine and optimize the accuracy of the spike-editing algorithms.

The spike-editing algorithms were also tested with sensory neuronal recordings to validate their performance on actual spike data. The neural data were collected using multi-electrode arrays known as Utah Electrode Arrays and a data acquisition system from Bionic Technologies LLC (Salt Lake City, UT). In the examples provided,  $9 \times 4$  arrays (36 electrodes) were inserted chronically into the L7 dorsal root ganglion of a cat. Experimental details are available elsewhere (Aoyagi et al., 2002). The cats were trained to walk on a treadmill and the recordings were made during normal walking, as well as under anesthesia. The stored waveforms were classified using an automatic spike classifier (SAC, Shoham et al., 2001). The programs for editing the spike trains are written in Matlab (Math Works, Inc., Natick, MA) and used these classified spikes. The programs can be easily adapted to other systems for analyzing neural data. Once this paper is published, the programs will be made available free of charge on the internet.

SAC sorted the recorded waveforms from each electrode using a clustering algorithm developed by Shoham et al., 2001, which projects the recorded waveforms into a parameter space defined by the group's principal components (PC). Action potential waveforms from different neurons are classified according to their distinct clusters in the  $n$ -dimensional PC space. Note that the clustering algorithm does not classify every recorded waveform as a discriminated unit, and the remaining units are labeled as *unclassified*. For each identified unit, an average waveform (*unit template*) was calculated from the set of waveforms that SAC classified. The unclassified waveforms were compared to each of the unit templates and a similarity score was assigned, based on the root mean square (RMS) error between the unclassified waveform and the unit template. The set of unclassified waveforms was then rank ordered by similarity score and *candidate* spikes were selected in ascending order of their match to the unit template.

## 2.3. Neural model

Spike trains were generated by a neural model with specified and hence known properties, including the occurrences of missed and extra spike events. The spike generator used the Matlab function *rand* to simulate an input to the model which varied randomly on a millisecond time scale. This input was then filtered using a third-order Butterworth

low-pass filter to provide a smoothed input. The cut-off frequency was varied between 0.1 and 3 Hz to match the physiological range. For example, the frequency of walking is about 1 Hz. The gain and offset of the model could also be adjusted to give a wide range of inter-spike intervals, including periods when the firing paused, as occurs in sensory neurons. The model itself consisted of a “leaky integrator” (French and Stein, 1970) with a time constant of 10 ms. Inputs to the model were integrated until a value of one was reached. Then, a pulse (spike) was generated and the model was reset to 0 after a refractory period of 5 ms. A second random input was fed directly into the leaky integrator to simulate the intrinsic variability of neurons. The distributions of inter-spike intervals were well fitted by a gamma distribution which is also true of many neurons (Stein, 1965). When only this input was applied and its gain was increased, the inter-spike intervals varied from a regular Gaussian distribution with low gain to a Poisson distribution (except for the refractory period) with high gain. These are well known to represent the limits of the gamma distribution. The times of the output pulses were stored in an array and processed with the same methods as the neural spike trains. To simulate corrupted spike trains, individual spikes were deleted or added at random times. The corrupted spike train was then processed with the spike-editing procedures described above.

## 3. Results

Fig. 4 shows the analysis procedures used on neural data. Unit 17.1 (the largest unit on electrode 17 in the array) was firing steadily at about 70 imp./s. The logarithms to the base 2 of the instantaneous frequencies, shown in Fig. 4A, were computed from the 3406 spike times classified by SAC (Shoham et al., 2001). The main reason for using this logarithmic scale is that the instantaneous frequency will decrease by one unit, if there are individual missed spikes (see Section 2). Similarly, the presence of spurious spikes will produce characteristic increases in frequency. The same data as in Fig. 4A is reproduced in Fig. 4B except that 948 candidate waveforms were added to the spike train, using criteria that are discussed below. Accepting these candidates reduced the number of missing spikes, but also introduced some spurious spikes (Fig. 4B). Consideration of Fig. 4C and D will be delayed until the methods used in Figs. 5 and 6 have been presented. Fig. 5 illustrates the process of identifying candidate spikes from the set of unclassified waveforms. Fig. 5A superimposes 100 spikes that SAC classified as being from a unit 17.1. Some triggered events did not meet the statistical criteria for classification, even though their waveforms were close to those of the accepted units. These candidate spikes can be rank ordered in terms of their similarity to the average template. If these spikes are part of the actual spike train, the pattern of log frequencies will become smoother. If they are spurious spikes, the pattern

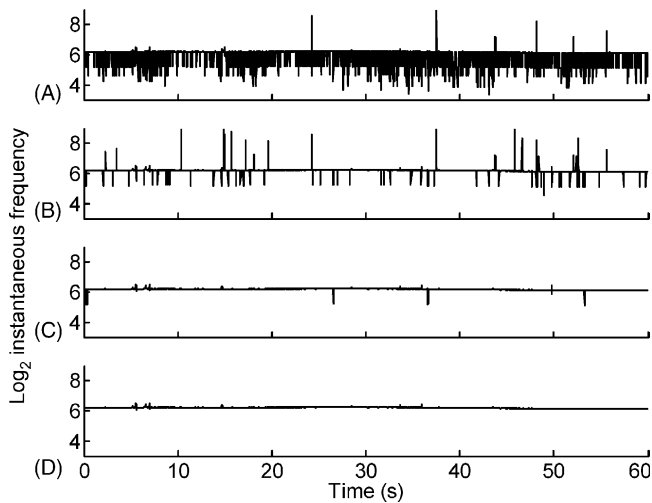


Fig. 4. (A) A neuron in the DRG (17.1) fired at a high constant rate ( $\sim 70$  Hz), which corresponds to a  $\log_2$  frequency  $> 6$ . Spikes were omitted frequently, so  $\log_2$  frequencies near 5 and lower are observed, as well as occasional, high values from accepting extra spikes (unit 17.1, 3406 spikes, SDF 0.393). (B) Accepting 64% of the candidate spikes eliminated most of the missed spikes but inserted some erroneous, extra spikes (inserted 948 candidates: SDF 0.056). (C) An automated procedure that checks the local statistics eliminated nearly all of these (see Fig. 5) (deleted 22, inserted 53 spikes: SDF 0.01). (D) The remaining errors were corrected manually by interactive, computer routines (Fig. 6) (deleted 1, inserted 10 spikes: SDF 0.007). The average running standard deviation in  $\log_2$  frequency (SDF) is indicated in each part. The method for calculating SDF is given in the text.

should become more irregular. We quantified the firing rate regularity in the neighborhood of each spike by calculating a running mean and standard deviation of the  $\log_2$  frequencies for a group of five intervals (for the  $j$ th interval we used the values for the intervals from  $j - 2$  to  $j + 2$ ). The average of these standard deviations over all values of  $j$  is a measure of how smoothly the  $\log_2$  frequencies change and will be denoted by SDF. Fig. 5B shows that the minimum SDF occurs if 64% of candidate spikes are accepted that are closest to the template. Fig. 5C shows 100 of the accepted candidate waveforms have much the same amplitude and wave shape as the classified units, but have somewhat more scatter.

Next, a procedure was applied to identify locations automatically where spikes should be inserted or deleted (see Section 2). After applying this procedure (Fig. 4C) only a few deviant values remained. Fig. 5D shows the waveforms that were deleted by this process. Some appear to have been misclassified, since they have distinctly smaller amplitudes, while others look rather similar in shape to the classified units. Despite the similarities, none of the deleted spikes exhibit the late negative phase appearing in the original and candidate spikes.

Finally, we developed an interactive procedure for inserting and deleting spikes manually. Note in Fig. 6A that the unit is firing at a constant firing of  $6.2 \log_2$  units ( $2^{6.2} = 73$  spikes/s), and a few values appear as outliers. The shaded areas are  $> 0.8 \log_2$  units above and between 0.8 and  $1.4 \log_2$  units below the running means. These are the thresholds for deleting spikes and inserting 1 or 2 spikes, as described in

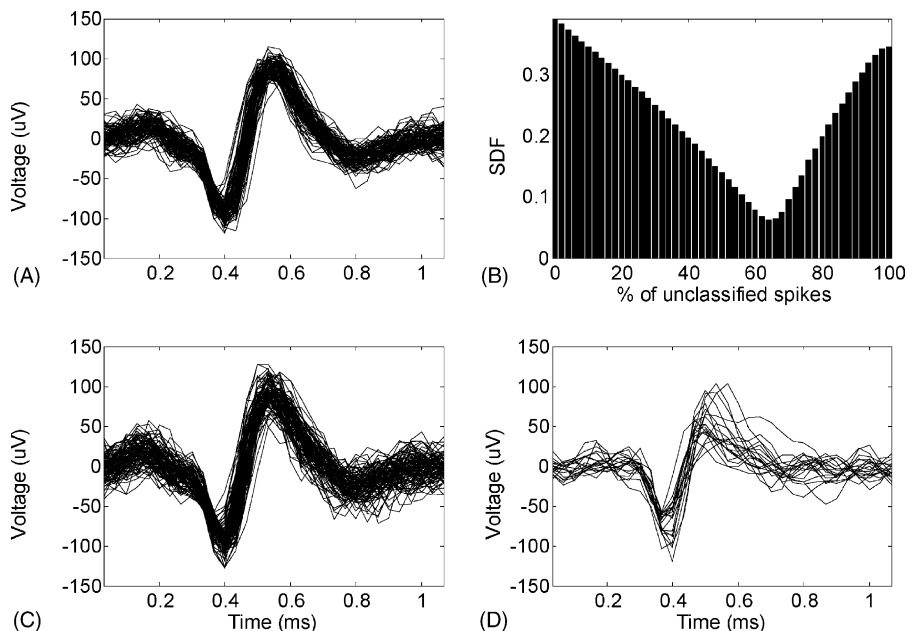


Fig. 5. (A) Waveforms of spikes discriminated by an automatic classifier (SAC; spikes classified by cluster analysis). (B) The average, running standard deviation in the  $\log_2$  frequency (SDF) reaches a minimum when 64% of unclassified, candidate spikes are accepted (optimal tolerance window size). These candidates had been rank ordered in terms of a similarity score based on the RMS deviation of each waveform from the mean template in (A). (C) Waveforms of candidate spikes that fall within this tolerance window. (D) Waveforms of spikes that were subsequently deleted by an automated process, based on the SDF (details in the text).

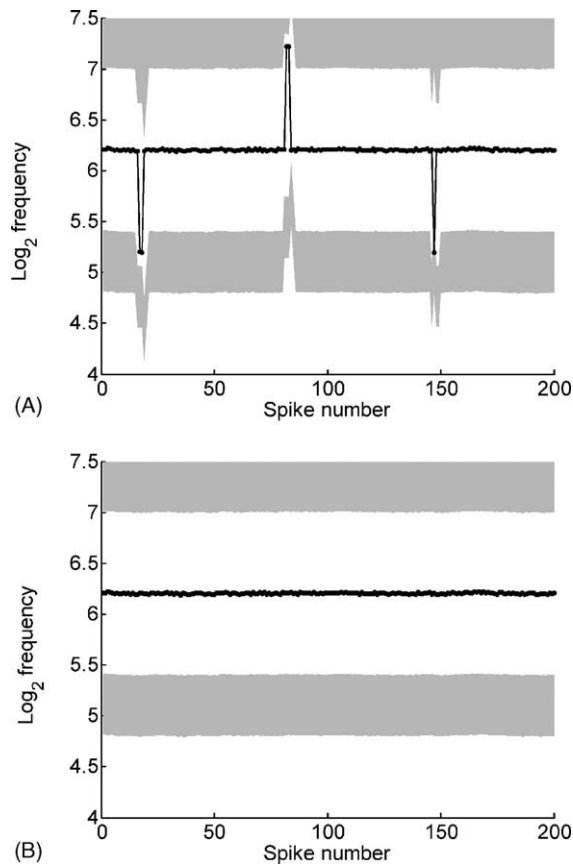


Fig. 6. Screens for manually editing spike trains. Each dot represents a single inter-spike interval. The shaded areas represent values that are  $>0.8$  above and between 0.8 and 1.4 below the running means. These are the thresholds for deleting an extra spike and inserting a missed spike by clicking on the appropriate dots. The screens are shown: (A) before and (B) after editing.

**Section 2.** If more than one spike was deleted or inserted in a small region, then the running mean may shift enough that the  $\log_2$  frequency of the spikes does not fall in the shaded areas. Thus, the automated routine will not correct these errors, but there are a number of commands available to make corrections manually. Fig. 6B shows the screen after inserting and deleting spikes manually. Finally, Fig. 4D shows the instantaneous frequencies for the whole 60 s trial after completion of the automated and manual editing procedures.

As this spike train is extremely regular, much information is contained in the inter-spike intervals or the instantaneous frequencies. The extremely regular, constant frequency of this neuron is unusual and may have resulted from pressure of the chronically implanted electrode on the neuron, resulting in an ectopic focus for eliciting spikes. The editing procedure is more prone to errors for a more irregular, time-varying pattern. To evaluate possible errors we used a model neuron (see Section 2) that fired in response to a randomly varying input (Fig. 7A). We then deleted 100 spikes from the pulse train and inserted another 100 at random times, before applying the automated procedure (Fig. 7B). Clearly, the random modification of the pulse train disrupted

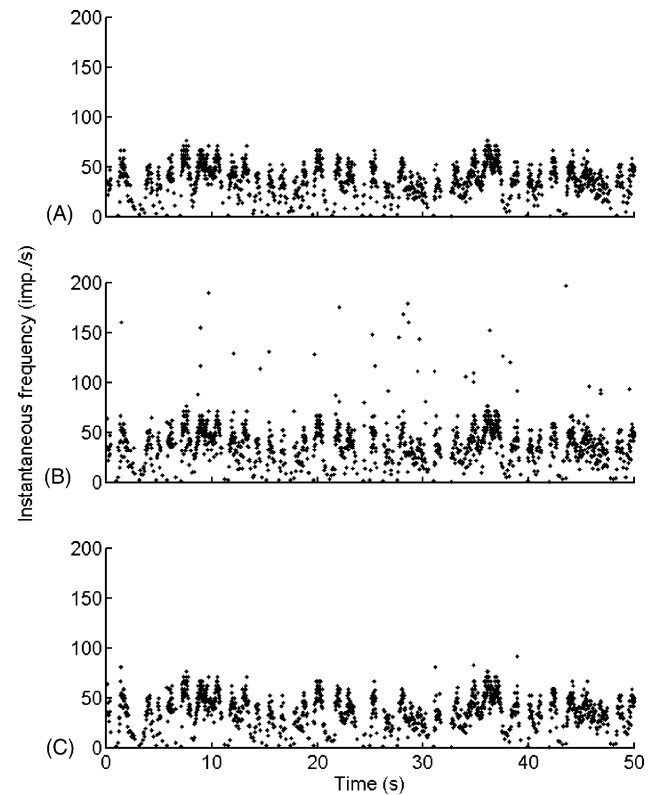


Fig. 7. (A) Frequency of pulses from a neural model (original: 1318 spikes). (B) To test the analysis 100 pulses were deleted and 100 extra pulses were inserted at random. (C) The automated editing procedure deleted most of the extra pulses and reinserted many of the deleted pulses (after editing: 88 deleted, 73 inserted; details in the text).

the pattern of activity and the automated procedure largely restored the pattern (Fig. 7C).

Fig. 8 shows the changes in more detail. Each dot represents a pulse on a second by second (line by line) basis. The + symbol represent pulses that were inserted and  $\circ$  those that were deleted. Each second is represented by two rows; the lower one is the pattern before editing and the upper one is the pattern after editing. If the editing procedure were perfect, every inserted pulse (+ on the lower row) would be matched by a deleted pulse ( $\circ$  on the upper row) and vice versa. In fact, eight of the 100 insertions and deletions occurred close to one another so that they did not change the pattern appreciably. Of the 92 remaining insertions 66 were corrected by deleting a spike in the editing procedure. In a few examples, an erroneous spike had been inserted close to an existing spike and the existing spike was deleted. These examples were considered correct since the number of spikes was correct and the instantaneous frequencies were close to the correct values. Of the 92 deletions, 40 were corrected by inserting a spike nearby. Although the automated procedure was not perfect, most of the insertions and many of the deletions were corrected.

One advantage of the neural model is that the cut-off frequency of the input, the mean firing rate and the variability in the firing rate could all be varied systematically. All three

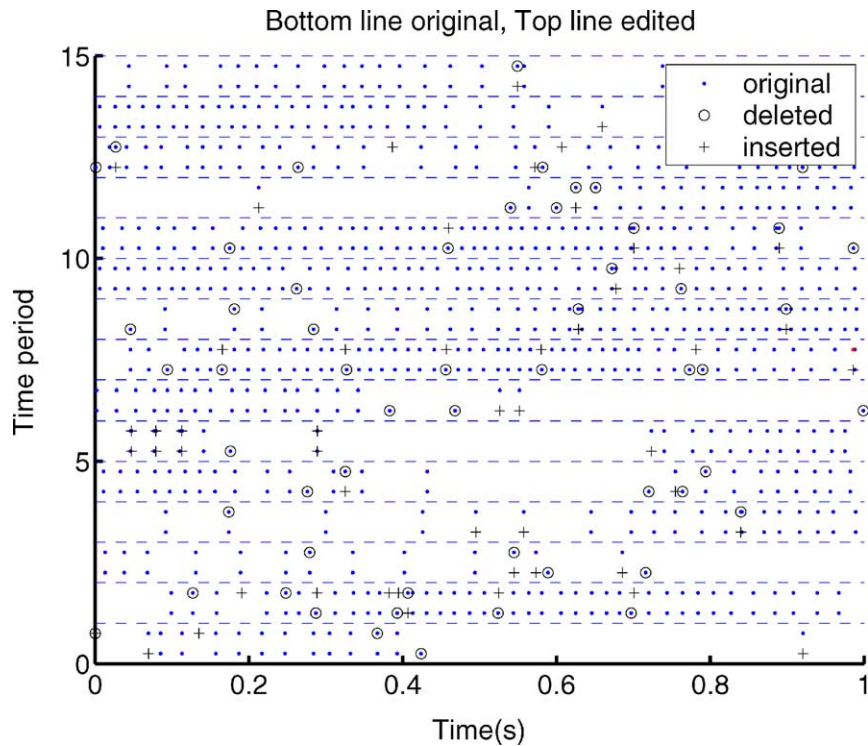


Fig. 8. Raster of pulses from the model to test the analysis procedures (bottom row original, top row edited). Original spikes are shown as blue dots. Those that were deleted ( $\circ$ ) and inserted (+) at random are shown on the lower line for each second. The insertions and deletions produced by the automatic editor are shown in the upper line for each second. If the editor worked perfectly  $\circ$  symbols on the bottom would be paired with + symbols on the top line and visa versa. Only 15 s of the 50 s trial are shown for clarity.

parameters will affect the running standard deviation of the  $\log_2$  frequency (SDF). Irrespective of which parameter was varied, the effect on the accuracy of the automated procedure was similar. For example, in Fig. 9A, the percentage

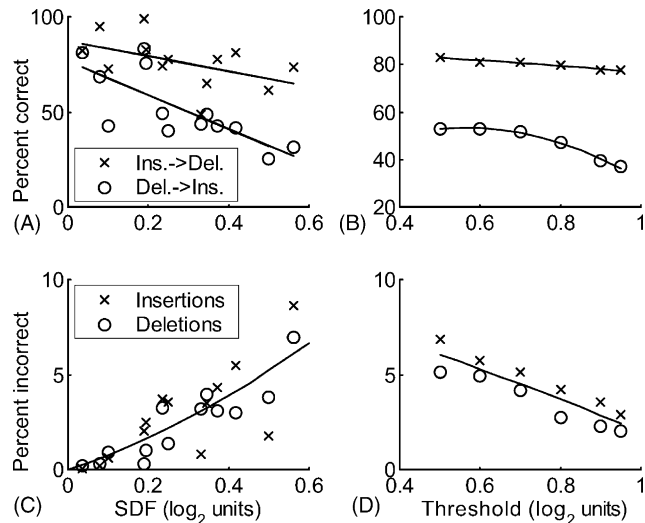


Fig. 9. The parameters of the neural model were varied so as to change the running standard deviation in  $\log_2$  frequency (SDF). (A) The percentage of correctly edited spikes decreased as SDF increased and (C) the percentage of incorrect editing changes increased. (B) Increasing the threshold for accepting spikes decreased the percentage of correct spikes found, but also (D) decreased the percentage of incorrect spikes inserted or deleted.

of the inserted spikes (x) that were deleted by the procedure declined from close to 100% when SDF was low to about 70% when SDF was large. The percentage of the deleted spikes that were reinserted fell even more steeply (to about 30% with large SDF). As expected, the percentage of incorrect insertions and deletions also increased as SDF increased (Fig. 9C), exceeding 5% when  $SDF > 0.5$ . Thus, the automated procedure will mainly be useful when  $SDF < 0.5$ . However, note that this was the value before inserting and deleting the spurious spikes. Even if the measured value of SDF is greater than this value, the automated editing procedure may bring the value down to within the acceptable range.

Fig. 9B shows the effect of varying the threshold parameter  $c_1$  for inserting a spike and the corresponding threshold for deleting a spike. The percentage of inserted spikes that were correctly deleted was quite independent of this parameter, whereas the percentage of deleted spikes that were reinserted declined fairly sharply if  $c_1 > 0.8$ . Fig. 9D shows that the percentage of incorrect insertions and deletions also decreased steadily as SDF increased. Thus, there is a tradeoff between missing as few erroneous spikes as possible, while not inserting or deleting too many incorrect ones. A value of threshold  $c_1 = 0.8$  was chosen as a suitable level. Another parameter that we varied systematically was the acceptance limit  $c_0$  in inequalities 1 and 2. Interestingly, the SDF reached a minimum (Fig. 10) at  $c_0 = 1.6$  for both the model

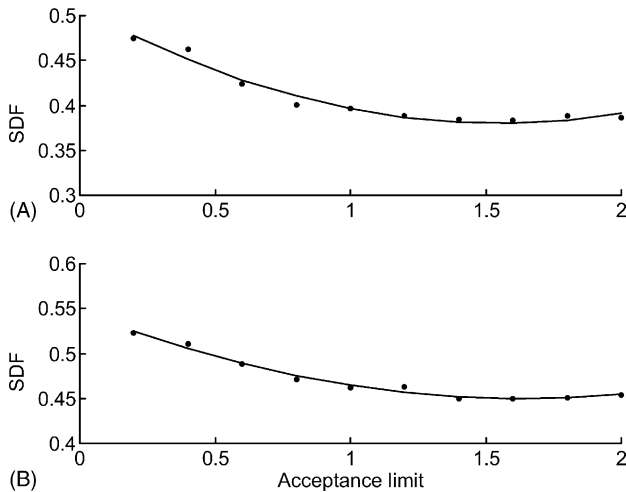


Fig. 10. SDF varies as the acceptance limit is changed. In this example, the minimum value occurred with an acceptance limit of 1.6 (i.e., log frequencies that were more than 1.6 SDF units from the local mean were tested with inequalities such as 1 and 2), both for: (A) the simulations and (B) the neural data.

and neural data. At the minimum, which typically occurred near a value of 1.6 when tested using a variety of parameters in the model, the numbers of insertions (and deletions) often approximately matched the numbers that had been deleted (and inserted).

This suggests an objective way of determining the value of  $c_0$  for real data sets, despite not knowing in advance which spikes are real or spurious (see Section 4). We also varied the number of intervals used to calculate SDF from 3 to 7, but this made very little difference in the result. Therefore, we have continued to use five-point statistics (two intervals on either side of the one being tested), as described above.

Finally, Fig. 11A shows the results of calculating the instantaneous frequency (the inverse of the inter-spike intervals) for a cutaneous afferent (paw) recorded on a 36-channel microelectrode array in a cat L7 DRG. The spike waveforms were classified using the automatic spike classifier described in Section 2. The cat walked on a treadmill for 60 s, but only 5 s of data are shown here for clarity. The instantaneous frequency  $f$  increased rhythmically with each step, but appeared quite erratic. First, we tested candidate spikes with waveforms of similar shape to those classified as the cutaneous afferent, as described in Section 2. In this data set only 21 were accepted (Fig. 11B). Then, each spike in the train was tested against the criteria for extra or missed spikes. The minimum SDF occurred at  $c_0 = 1.6$  (Fig. 10B), as for the simulation in Fig. 10A. The resultant discharge profile is plotted in Fig. 11C and the changes in frequency are now much smoother than in Fig. 11A. The toe moved forward during the swing phase (Fig. 11E) and back (on the treadmill) during the stance phase. In the cat, as in the human, the ankle extends at the end of stance (push off) and again at the end of swing as it is brought down toward the ground (Fig. 11D). An increase in frequency is seen each time the

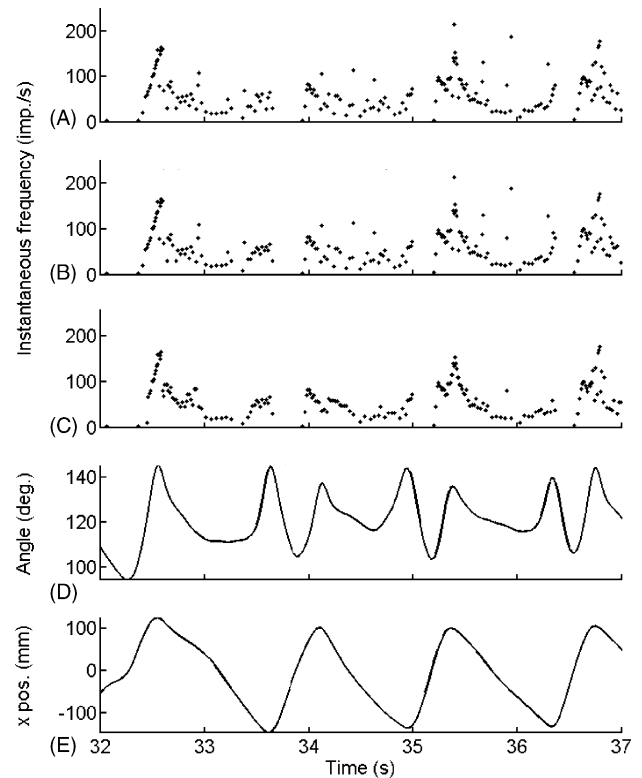


Fig. 11. Firing rate of a cutaneous afferent during walking on a treadmill. (A) Instantaneous frequency of spikes accepted by SAC. Only a 5 s portion of a 60 s trial is shown for clarity (original: unit 8.1, 2166 spikes, SDF 0.6698). (B) In this data set, only 21 candidate spikes were accepted, so inserting them hardly changed the local variability (SDF: 0.66654). (C) After using the automatic editing process the pattern of frequencies is now much smoother (SDF reduced) (deleted 213, inserted 152 spikes, SDF 0.45322). (D) The firing rate increased each time the ankle was extended. (E) The toe moved forward during swing and backward (on the treadmill) during stance (toe  $\times$  position). The ankle extends and flexes twice during each step cycle.

ankle was extended during the stance phase (Fig. 11D), which stretched the skin in its receptive field. The unit paused its firing during the swing phase, when the foot was off the ground. Note that the editing routine generally eliminated isolated high and low frequencies (compare Fig. 11C with A or B) without greatly affecting the bursts that occur during extension or the pauses that occur during rapid flexion.

#### 4. Discussion

As multi-channel recording becomes more common, so too does the need to examine the spike trains that are recorded to determine if spikes have been missed or extra spikes have been included. Particularly, when there is a substantial amount of noise or when only small segments around the time of a threshold crossing are stored, exhaustively checking all possible superpositions may be difficult or impossible. The present paper considers a complementary

method. First, spikes are examined that are close in waveform to those that have been accepted and either accepted or rejected, based on statistical criteria (Fig. 5B). Then, unusually short and long intervals (unusually high or low instantaneous frequencies) are tested to see whether they fit the criteria for extra or missed intervals (Fig. 3). The testing is done on a simple statistical basis and we have verified the process using a model neuron that simulates many of the features of our sensory neurons. In the model, we varied the frequency response of the input, the firing rates of the model neuron and its variability. In this way, the parameters of the automated process could be optimized for comparison to neural data.

In the model, a minimum error typically occurred when the acceptance limit was set to about 1.6 (Fig. 10). Whether that acceptance limit is suitable for neural data can be checked by each investigator as follows. If a single unit is recorded with good signal to noise ratio, so the investigator is confident of the discrimination of nearly every spike, spurious spikes could then be inserted and deleted at random, as we have done for the model. If the minimum error for such neurons occurs when the acceptance limit is close to a constant value, such as 1.6, then this value can be used for other neurons, which are thought to have similar discharge properties, but where the discrimination of the spikes is not as good.

Another potential application of these methods is as a metric for recording quality. The measure we introduced, the running standard deviation in log frequency (SDF) depends on the intrinsic variability of the neuron and the rate at which the inputs to it are varying. As shown in Fig. 5B, this measure will vary with the size of the tolerance window or the threshold for recording events. A test could be built into commercial recording equipment to set values that minimize SDF in a test data set. Any change from that value would give a running indication that isolation quality had changed. The feasibility of this procedure still needs to be tested on a wide range of neural data.

The process presented here works best when the extra or missing spikes are relatively sparse, since the tests are based on local intervals immediately before and after the spike in question. If errors in spike detection occur frequently, the basis for the statistical tests may not be valid. Frequent errors will increase  $s$ , so it may be more difficult to reject extra spikes or replace missed ones. Highly irregular spike trains or ones in which the rates change very rapidly also increase  $s$ . Many cortical cells are highly variable, with a discharge that approximates a random, Poisson process (Lee et al., 1998; Maynard et al., 1999; Shadlen and Newsome, 1998). From the results with the model, the present methods may be of limited value if  $SDF > 0.5$  for these neurons. Rather than merely calculating the local mean and standard deviations, more complex fitting procedures, such as polynomial regressions, splines, etc. could be used. However, if there are multiple missed spikes or extra spikes, these procedures will be more affected than a simple calculation of the local mean and standard deviation, so we have not explored them in detail.

Fortunately, it is just the neurons with regular firing patterns that most need these methods. In many multi-unit recordings the number of spikes is counted over 100 ms bins, for example, for a neuron that is firing regularly at 15 spikes/s, one or two spikes will occur in each 100 ms bin, representing rates of 10 or 20 spikes/s, rather than the actual rate of 15 spikes/s. If spikes are missed or extra spikes are added, the range will be even greater. This quantization error will be less significant, if the discharge is irregular with intervals varying from 40 to 200 ms (5–25 spikes/s), but the bins will then contain from 0 to 3 spikes (0–30 spikes/s). Some filtering of the counts in each bin is often done, but this will reduce the ability to follow rapidly changing inputs. According to the Nyquist criterion, only frequencies below 5 Hz can be studied, if 100 ms bins are used, and filtering will reduce the frequency response further. Since many neurons are able to follow cyclic inputs varying at between 10 and 100 Hz, this represents a very severe limitation. In summary, instantaneous frequencies can provide a more accurate record of neural activity than binning and filtering, if the neurons fire regularly and erroneous spikes can be edited out. Finally, the investigator can edit the spike trains manually, if additional information is available that can be used to accept or reject spikes that have been selected. However, this manual editing is not constrained by objective statistical criteria and so can lead to distortions or over-corrections. Therefore, such manual editing should be used sparingly.

Other than the manual editing, all the procedures are relatively fast to compute and proceed automatically. Since many multi-unit recordings are intended for real-time control, eventually for a neuroprosthetic application, speed may become an important issue. We are not currently doing real-time control, so this application has not been tested. A test set of data can be run, and then the parameters from these data can be applied to new data, as they are recorded. With the increasing speed and power of microprocessors and gate arrays, real-time application should be possible in the near future, if not already feasible. In conclusion, this paper describes a variety of methods to improve the accuracy of spike data obtained from multi-electrode arrays and further applications of the methods appear promising.

## Acknowledgements

This research was supported by the Canadian Institutes of Health Research. Dr. Weber is a Post-doctoral fellow of the Alberta Heritage Foundation for Medical Research. We thank Dr. P. Ellaway and Dr. S. Shoham for comments and suggestions.

## References

- Aoyagi Y, Stein RB, Weber DJ, McDonnall D, Branner A, Normann RA. Recording capabilities of a penetrating microelectrode array in

- dorsal root ganglion and its usefulness for coding of limb position. In: Proceedings of the 6th Annual Conference, IFESS; 2002. p. 109–11.
- Fee MS, Mitra PP, Kleinfeld D. Automatic sorting of multiple unit neuronal signals in the presence of anisotropic and non-Gaussian variability. *J Neurosci Methods* 1996;69:175–88.
- French AS, Stein RB. A flexible neural analog using integrated circuits. *IEEE Trans Biomed Eng* 1970;17:248–53.
- Gray CM, Maldonado PE, Wilson M, McNaughton B. Tetrodes markedly improve the reliability and yield of multiple single-unit isolation from multi-unit recordings in cat striate cortex. *J Neurosci Methods* 1995;63:43–54.
- Harris KD, Henze DA, Csicsvari J, Hirase H, Buzaki G. Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements. *J Neurophysiol* 2000;84:401–14.
- Hoffman KL, McNaughton BL. Coordinated reactivation of distributed memory traces in primate neocortex. *Science* 2002;297:2070–3.
- Hoogerwerf AC, Wise KD. A three-dimensional microelectrode array for chronic neural recording. *IEEE Trans Biomed Eng* 1994;41:1136–46.
- Jansen RF, Ter Maat A. Automatic wave form classification of extracellular multi-neuron recordings. *J Neurosci Methods* 1992;41:123–32.
- Kaneko H, Susuki SS, Odada J, Akamatsu M. Multineuronal spike classification based on multisite electrode recording, whole-waveform analysis, and hierarchical clustering. *IEEE Trans Biomed Eng* 1999;46:280–90.
- Lee D, Port NL, Kruse W, Georgopoulos AP. Variability and correlated noise in the discharge of neurons in motor and parietal areas of the primate cortex. *J Neurosci* 1998;18:1161–70.
- Lewicki MS. A review of methods for spike sorting: the detection and classification of neural action potentials. *Network* 1998;9:R53–78.
- Maynard EM, Hatsopoulos NG, Ojakangas CL, Acuna BD, Sanes JN, Normann RA, Donoghue JP. Neuronal interactions improve cortical population coding of movement direction. *J Neurosci* 1999;19:8083–93.
- Rousche PJ, Normann RA. A method for pneumatically inserting an array of penetrating electrodes into cortical tissue. *Ann Biomed Eng* 1992;20:413–22.
- Shadlen MN, Newsome WT. The variable discharge of cortical neurons: implications for connectivity. *J Neurosci* 1998;18:3870–96.
- Shoham S. Advances Towards an Implantable Motor Cortical Interface. Ph.D. Thesis, Biomedical Engineering Department, University of Utah, Salt Lake City, UT; 2001.
- Stein RB. A theoretical analysis of neuronal variability. *Biophys J* 1965;5:173–94.
- Taylor DM, Helms Tillery SI, Schwartz AB. Direct cortical control of 3D neuroprosthetic devices. *Science* 2002;296:1829–32.
- Wessberg J, Stambaugh CR, Kralik JD, Beck PD, Laubach M, Chapin JK, Kim J, Biggs SJ, Srinivasan MA, Nicolelis MA. Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. *Nature* 2000;408:361–5.