Published in final edited form as:
J Neurosci Methods. 2009 June 30; 181(1): 27–35. doi:10.1016/j.jneumeth.2009.04.014.

New approaches to eliminating common-noise artifacts in recordings from intracortical microelectrode arrays: inter-electrode correlation and virtual referencing

Kunal J Paralikar[1], **Chinmay R Rao**[2], and **Ryan S Clement**[1]

1 *Department of Bioengineering, The Pennsylvania State University, University Park, Pennsylvania*

2 *Department of Electrical Engineering, The Pennsylvania State University, University Park, Pennsylvania*

Abstract

Intracortical microelectrode arrays record multi-unit extracellular activity for neurophysiology studies and for brain-machine interface applications. The common first step is neural spike detection; a process complicated by common-noise signals from motion artifacts, electromyographic activity, and electric field pickup, especially in awake/behaving subjects. Often common-noise spikes are very similar to neural spikes in their magnitude, spectral, and temporal features. Provided sufficient spacing exists between electrodes of the array, a local neural spike is rarely recorded on multiple electrodes simultaneously. This is not true for distant common-noise sources. Two new techniques compatible with standard spike detection schemes are introduced and evaluated. The first method, *virtual referencing* (*VR*), takes the average recording from all functional electrodes in the array (represents the signal from a virtual electrode at the array's center) and subtracts it from the test electrode signal. The second method, *inter-electrode correlation* (*IEC*), computes a correlation coefficient between threshold exceeding candidate spike segments on the test electrode and concurrent segments from remaining electrodes. When sufficient correlation is detected, the candidate spike is rejected as originating from a distant common-noise source. The performance of these algorithms was compared with standard thresholding and differential referencing approaches using neural recordings from unanaesthetized rats. By evaluating characteristics of mean-spike waveforms generated by each method under different levels of common-noise, it was found that *IEC* consistently offered the most robust means of neural spike-detection. Furthermore, *IEC*'s rejection of supra-threshold events not likely originating from local neurons significantly reduces data handling for downstream spike sorting and processing operations.

Keywords

Spike detection; multi-unit recording; common-noise; inter-electrode correlation; common average referencing

Corresponding Author: Ryan S. Clement, Assistant Professor, Department of Bioengineering, The Pennsylvania State University, 205 Hallowell Building, University Park, PA 16802, Phone: 814-865-5190, Email: E-mail: rclement@psu.edu.

1. Introduction

The study of multi-unit spike activity from neuronal assemblies is essential to further our understanding of neurophysiology involved in learning and plasticity (Schwartz et al., 2006). It is also required in developing algorithms for effective brain machine interfaces (BMI). Intra-cortical microelectrode arrays offer the spatial and temporal resolution to record spike activity *in vivo*. This ability to record from population of neurons using electrodes implanted in the cerebral cortex was demonstrated as early as 1958 (Strumwasser, 1958). Since then, a variety of electrodes have been developed. They range from inexpensive, hand-fabricated microwires to micromachined planar and 3-D electrodes with silicon, ceramic or polymeric substrate (Hoogerwerf and Wise, 1994; Nordhausen et al., 1996; Rousche et al., 2001; Stieglitz et al., 2005). Each of these arrays is implanted in the brain following standard craniotomy. Once implanted, they transduce extracellular spike activity into voltage signals that are amplified and stored for further analysis.

Neural spiking events are typically detected using a pre-defined threshold on the recorded electrode followed by template-matching, Principal Component Analysis(PCA), clustering techniques, or Bayesian filters (Chapin et al., 1999; Isaacs et al., 2000; Lewicki, 1998; Rennaker et al., 2005; Suner et al., 2005) to select events with neuronal characteristics. Some of these techniques as well as others have been employed in spike sorting applications with goal of associating individual neurons with behavioral tasks (Lebedev and Nicolelis, 2006; Schwartz et al., 2006). Although considerable sophistication has been achieved in spike-sorting techniques(Vargas-Irwin and Donoghue, 2007), little attention has been given to the operation of spike-detection that precedes sorting. Experience with chronically implanted subjects performing behavioral tasks suggests that in addition to multi-unit neural spike activity, microelectrodes also record electromyographic activity (EMG) from muscles, especially mastication signals (EMG generated by animal chewing), and relatively large signals generated by abrupt animal movements, or interference with the recording setup by awake and active subjects (Gilmour et al., 2006; Musial et al., 2002; Paralikar and Clement, 2008; Sasaki et al., 1983). These non-neural signals can be similar to neural signals in their spectral and temporal characteristics (Figure 1). Hence spike-detection schemes that involve threshold-based neural-spike detection on an electrode by electrode basis may suffer from false-positive detection thereby negatively impacting downstream spike-sorting operations as well as increasing power and processing requirements for wireless systems.

Recent modeling studies have determined that signal amplitude of layer V pyramidal cells drops substantially with increase in distance between the firing cell and recording site (Gold et al., 2006; Moffitt and McIntyre, 2005). Hence it can be argued that similar neural-like signals appearing concurrently on distantly (≥250μm) spaced electrodes may not be originating from pyramidal (neural) cells. Given the computational attractiveness of threshold-based detection technique, it is desirable to employ equally simple “signal cleaning” schemes that may enable improvements in spike detection outcomes. Traditionally, differential recording between electrode of interest and a “quiet electrode” implemented in hardware (Sasaki et al., 1983) or post-recording inter-electrode subtraction is employed to ameliorate the situation. However, slight temporal shifts in non-neuronal signals across electrodes can cause asynchronous cancelations thereby resulting in artificial spike-like outputs. Moreover presence of neural spikes on the reference electrode may affect differential recording outcomes. Hence other techniques need to be investigated.

Here, two such techniques that use information from other electrodes of an array are introduced and compared with simple thresholding (*ST*) and differential referencing (*DR*) for their ability to remove non-neuronal spike-like segments from real data acquired from unanesthetized rats. The first method, “virtual referencing” (*VR*), seeks to obtain a representation of the common-

noise recorded across the entire array via an ensemble-average of all electrode signals. This can be thought of as virtual-electrode located in the center of the array that can serve as the differential reference. The second method termed “inter-electrode correlation” (*IEC*) also utilizes data from other electrodes of an array in deciding whether or not a given candidate spike (threshold exceeding event) likely originated from a local neuron. A high degree of correlation between a candidate spike and concurrent segments on other electrodes is considered indicative of a non-neuronal artifact originating some distance from the array.

Results demonstrating the statistically significant benefit of using *IEC* incorporated detection schemes in real-data having a high degree of common-noise has been previously presented (Paralikar et al., 2008). The primary purpose of this study was to evaluate the effectiveness of the *IEC* and *VR* algorithms operating on real datasets with varying degrees of common-noise contamination and comparing performance with the more common *ST* and *DR* techniques. The efficacy of the techniques was evaluated objectively by comparing features of the mean-spikes generated from the ensemble-average of all valid spike events identified by the different algorithms. The results indicate that relative to the other methods, *IEC* offers a robust solution whose outcome is not highly dependent on the quality of data. Moreover, *IEC* significantly reduces the number of false-positives spike events that must be processed by subsequent steps like spike-sorting, especially when common-noise is particularly high. Finally, the IEC approach, which does not alter the actual signal recorded at the electrode, lends itself to experiments designed to track the state of the neural interface over time.

2. Material and Methods

2.1 Subjects and Surgical Procedures

The implant procedure, electrode fabrication details and recording conditions are described in detail in (Paralikar and Clement, 2008). All animal procedures followed National Institutes of Health (NIH) Guidelines for the care and use of animals and were approved by the Penn State Institutional Animal Care and Use Committee (IACUC) committee. Subjects were anesthetized with an initial dose of ketamine/xylazine/acepromaxine (50:5:1 mg/kg) with additional anesthesia given to maintain areflexia. Craniotomy was performed 2-4mm anterior to the bregma and lateral to the midline in order to target the motor cortex. Following excision of dura-mater, the electrode array (2×4 tungsten microwire array, Φ50μm, 250μm inter-electrode spacing, electrode impedance: 254 ± 90 kΩ *in vivo*) was introduced at a constant speed of 10μm/sec. Final implant depth was 1mm to target layer V of rat cortex. Bone screws that acted as electrode ground and mechanical anchors were placed in the cranial plates and a dental acrylic headcap was created on top of the skull to house the electrode connectors.

2.2 Data Acquisition and Selection

Each dataset consisted of five-minute recording samples from chronically implanted microelectrode arrays in three unanesthetized rats. Neural recordings were obtained using a commercial multichannel acquisition system (Tucker-Davis Technologies Inc, Alachua, FL). Acquired analog signals were digitized at 25kHz, band-pass filtered between 300Hz - 5kHz, downsampled at 12kHz and then stored for offline analysis. Subsequent processing and analysis was carried out using custom developed scripts in MATLAB (The Mathworks, Natick, MA). The goal of this study was to compare the ability of different techniques in improving spike detection outcomes in data having varying degree of common-noise present simultaneously across multiple electrodes of an array. The degree of common-noise present on a given electrode was calculated in two ways (Figure 2). In the first method, the average inter-electrode correlation coefficient of the entire recording block (5 minutes) was calculated between the signal on the electrode of interest and remaining functional electrodes of the array (Figure 2: blue diamonds). In the second method, an average signal was obtained from all the

functional electrodes of an array (virtual reference). The percentage difference in the root mean square value (RMS) of the recorded signal on the electrode of interest and that of the average signal was calculated (Figure 2: pink squares). Eighteen data-segments exhibiting a range of RMS differences and degree of correlation were selected. These were equally distributed in three groups: designated as low, medium and high common-noise as shown in Figure 2. The signal-to-noise ratio of each data set is also plotted in Figure 2. Signal-to-noise ratio is defined as the peak-to-peak magnitude of the mean-spike over the noise level for the electrode (estimated by multiplying the standard deviation of the electrode signal after candidate spikes were removed by a factor of two) as described elsewhere (Paralikar and Clement, 2008).

2.3 Signal Processing Techniques for Spike Detection

Four algorithms were used to process raw recordings obtained from arrays with 250μm inter-electrode spacing in order to identify "valid" spike segments. These algorithms included Simple Thresholding (*ST*), Differential Referencing (*DR*), Virtual Referencing (*VR*) and Inter-electrode Correlation (*IEC*). Each of these algorithms are described below and depicted pictorially in Figure 3.

For the *ST* algorithm, spikes were detected by threshold-crossing (set to -3 times the standard deviations (σ) of raw signal (Figure 3C) and extracted as 3ms waveform segments (39 sample points: 10 sample points preceding local minima and 29 sample points following the local minima). For the *DR* algorithm, the raw signal on a functional electrode with lowest RMS value was subtracted from the signal on the electrode of interest (test electrode). Spikes were detected in the differential referenced signal similarly to *ST*. In *VR*, a reference signal generated from the ensemble average recording from all functional electrodes of an array was subtracted from the signal on the electrode of interest. The signal so generated was then processed for spike detection similarly to the *ST* and *DR* algorithms. It is important to note that the term "valid spikes" will be used throughout the remainder of the paper to refer to all spikes detected as valid by a particular algorithm. "Valid" does not necessarily mean a particular spike was necessarily caused by an action potential from a local neuron since this cannot be definitively known without confirmation from intra-cellular recording.

2.4 Inter-Electrode Correlation (IEC)

In the implementation of the *IEC* algorithm, signal segments that exceeded the -3σ threshold were identified and extracted from the raw recording as candidate spikes similar to the *ST* algorithm. For each candidate spike identified on the test electrode, concurrent signal segments from the other electrodes of the array were also extracted (Figure 3D) and stored in a signal matrix. Correlation coefficients were then computed between the candidate spikes and these concurrent segments. Thus for an eight electrode array, this resulted in seven correlation coefficients. If any of the coefficients exceeded a pre-determined threshold (described next), the spike was rejected as not originating from a local neuron and hence eliminated from subsequent analysis.

The value of the inter-electrode correlation threshold was selected after exhaustive analysis utilizing the expertise of a researcher (unbiased rater) with experience in spike detection and sorting. In this assessment, the maximum of the correlation coefficients was evaluated as a discriminant for a selected training set of candidate waveforms (segments exceeding -3σ) from data having different levels of common-noise (similar to Figure 2). A selection of candidate waveforms were presented to the unbiased rater who was charged with categorizing the spike as neuronal or not. Two probability density plots were obtained for the values of the maximum correlation coefficient for evaluated events classified by the rater as neural spikes, and those that were not. This was followed by plotting of Receiver Operating Characteristic (ROC) curves that plotted the true-positive rate (determined by rater) vs. false-positive rate as a function of

the correlation coefficient threshold. The correlation threshold was then selected to be a value that was around the knee of the ROC curves in all three cases of common-noise. This ensured an optimal trade-off between false negatives and false positives (i.e. probability of equal error).

2.5 Mean-Spike Generation and Feature Analysis

For each of the algorithms, a mean-spike waveform was generated by taking point-wise average (after aligning to the minima) of all the valid spike segments identified by the specific algorithm, as shown in Figure 3E. Recent modeling studies (Gold et al., 2006; Moffitt and McIntyre, 2005) have shown that spikes with sources (neurons) closest to the recording electrode have highest peak-to-peak amplitude and lowest depolarization phase duration. Similarly, exhaustive subjective analysis undertaken in the author's lab (Paralikar et al., 2008) has shown that mean-spikes generated from signal segments identified to be neuronal by experienced spike-sorters have a higher amplitude and shorter depolarization phase duration as compared to mean-spike waveforms generated from non-sorted spikes detected by simple-thresholding (*ST* approach). Thus it was assumed that the algorithm that produced mean-spike features which most closely matched these results (high peak-to-peak amplitude and shorter depolarization phase) would contain a higher proportion of true neural spiking events, and therefore a lower proportion of nonneuronal supra-threshold events. With these assumptions in mind, a number of mean-spike features were identified and calculated for the purpose of making quantitative comparisons of the outcomes from the different algorithms. These features (presented later in Table I) included peak-to-peak amplitude (*P2P*), the durations of the depolarization and repolarization phases (*DepPhase* and *RepPhase*, respectively; defined as width of mean-spike at 10% phase minima/maxima), number of valid spikes (*#Spikes*), and the *P2P/DepPhase* ratio. One-way analysis of variance (ANOVA) test followed by Dunnett's test was used to assess whether potential differences in mean-spike features resulting from the *DR*, *VR*, and *IEC* algorithms were statistically significant compared to the *ST* algorithm (also designated as Reference Group for Dunnett's test). *P*-values less than 0.05 were considered significant.

3. Results

Eighteen neural recording datasets (five minutes each) from three unanesthetized rats were used in this study. The datasets were selected for their varying degree of common-noise as discussed in Methods and Figure 2. Four algorithms (*ST*, *DR*, *VR*, *IEC*) were used to independently detect neural spiking events in raw recordings and generate the mean-spikes. Feature analysis was performed on the mean-spikes to evaluate and compare the efficacy of the different spike-detection algorithms.

3.1 Correlation Coefficient Threshold

Subjective assessments by an experienced spike sorter (blinded to the details of this study) were utilized to help determine the correlation coefficient threshold for the *IEC* algorithm. To do this, the sorter was provided three data-sets containing threshold-detected candidate spikes, one from each common-noise category (Figure 2). The rater was then asked to separate what appeared to be individual neural spike events from non-neuronal events. Next, the probability distribution of maximum inter-electrode correlation coefficients was determined for the neural events (rater-determined true positives) and non-neuronal events (rater determined true negatives). Since the distributions overlapped, any threshold chosen for the maximum correlation coefficient will likely result in a finite false-positive rate. The maximum correlation coefficient threshold was varied between zero and one and the resulting true-positive detection rate was plotted vs. false-positive rate for each threshold setting, resulting in the Receiver Operating Characteristic (ROC) curve. As shown in Figure 4A-C, the equal likelihood point (when false-positive rate equaled the false-negative rate) was between 0.75-0.82 in all three

cases. A conservative threshold to distinguish between potentially neuronal (R<0.75) and potentially non-neuronal segments (R>0.75) was chosen. While it was required to utilize a subjective spike-sorter to help determine a good value for the correlation coefficient threshold, all subsequent analysis which was performed on other data sets utilizing the *IEC* algorithm was strictly objective and automated.

3.2 Qualitative Assessment of Mean-Spike Shapes

Representative samples of mean-spike waveforms generated by the four algorithms under different common-noise conditions are presented in Figure 5. With low common-noise, the mean-spikes generated by the *ST*, *VR* and *IEC* algorithms were typically very similar in shape with relatively narrow and distinct negative depolarization phase followed by a well-defined but longer lasting positive repolarization phase, characteristic of layer V pyramidal cells. However, the *DR* algorithm yielded atypical neural mean-spike characteristics: a broad depolarization phase, a poorly-defined repolarization phase along with a tri-phasic appearance. Presence of neural spikes on the reference electrode and/or inadequate cancellation of the common-noise signal may be potential factors leading to this outcome. In instances of medium common-noise, all algorithms generated similar results: mean-spikes with relatively well-defined depolarization and repolarization phases but of smaller peak-to-peak (*P2P*) amplitude as compared to the low common-noise condition. Again the *DR* algorithm appeared to generate the least neural-looking mean-spike by its poorly-defined repolarization phase. In contrast, the mean-spike produced by the *IEC* algorithm had the sharpest depolarization and repolarization phases as well as one of the largest *P2P* amplitudes, suggestive of greater contribution from true neural spikes. Under high common-noise the mean-spikes appeared to differ significantly across the spike-detection algorithms. The shape of the mean-spike produced by *ST* had large *P2P* amplitude but exhibited a very broad depolarization phase as compared to the other algorithms, presumably due to significant contamination from non-neuronal supra-threshold events. After application of the *IEC* algorithm, which removed all spike events that were significantly correlated with signals on other electrodes, the mean-spike appeared much more neural exhibiting sharp, clearly identifiable depolarization and repolarization phases. In addition, the *P2P* amplitude of the mean-spike produced by the *IEC* algorithm was the largest among the remaining algorithms. Overall, the *IEC* algorithm seemed to consistently produce the most neural-like mean-spike waveforms regardless of the level of common-noise. Although highly qualitative, these observations point to the effectiveness of the *IEC* technique in reducing the contribution from other non-neural sources (or sources other than local neurons).

3.3 Quantitative Analysis of Mean-Spike Features

Feature analysis was done to render objective quantification to the above mentioned subjective mean-spike waveform comparisons and to evaluate results from all the datasets for statistically significant differences. The initial goal was to compare features of mean-spikes generated by different algorithms and to determine the existence of trends between mean-spike *P2P* amplitude and the duration of depolarization phase (*DepPhase*). As discussed previously, higher mean-spike *P2P* amplitudes accompanied by shorter (or sharper) *DepPhase* were assumed to result when a greater proportion of true neural spike events were present. Figure 6 shows a plot of the spread of the *P2P* amplitude of the mean-spike vs. *DepPhase* for mean-spikes generated by all the algorithms applied to each of the 18 datasets. These plots help reinforce the visual observations associated with the mean-spike shape examples of Figure 5. With the exception of the *DR* algorithm, the outcomes as indicated by the mean-spike feature plots were essentially non-differentiable in the low common-noise data sets. Consistent with the example presented in Figure 5, the *DR* algorithm on average produced the largest and most widely varying *DepPhase* in the low common-noise datasets. In the high common-noise datasets, the mean-spikes generated by the *ST* algorithm were also atypically broad and widely varying, suggesting considerable nonneuronal event contamination with this algorithm. While

the *P2P* amplitude decreased after application of the *IEC* algorithm, the sharpness of the depolarization phase was much improved on average. Overall, the analysis revealed that *IEC* generated mean-spikes consistently had the most neuronal characteristic (as indicated by the ratio *P2P* vs. *DepPhase*) amongst all the algorithms under consideration for datasets with medium and high levels of common-noise contamination. The mean-spikes generated from the *VR* method also had consistently narrow depolarization phase, but were usually smallest in mean-spike *P2P* amplitude.

A summary of the means and standard deviation for a number of performance metrics for each algorithm are provided in Table I. One-way ANOVA (followed by Dunnett's test) was carried out to detect statistically significant differences in results obtained by any of the algorithms as compared to *ST*. The *IEC* algorithm yielded the largest *P2P/DepPhase* ratio for medium and high common-noise conditions, and was only slightly smaller than *ST* for low common-noise. When compared to *ST* however, the difference in the *P2P/DepPhase* ratios were not statistically significant for any of the different common-noise conditions. This is not necessarily surprising though, considering the especially high variability in both the mean-spike *P2P* amplitude and the depolarization phase duration resulting from the *ST* algorithm. Furthermore it should be noted that the datasets with largest common-noise contamination, also had the lowest signal-to-noise ratios. Therefore it is not unexpected that the mean-spikes generated from a simple-threshold detection approach (i.e. the *ST* algorithm) would be inflated by a significant number of correlated supra-threshold events that overshadowed the relatively small neural spikes remained after *IEC* was applied. While improvements in mean-spike *P2P* amplitude were not observed, improvements in the sharpness of the depolarization phase were statistically significant for both the *VR* and *IEC* algorithms when a high degree of common-noise was present.

3.4 Comparison of Number of Valid Spike-segments

Another important finding of this study was that the number of valid spike-segments identified by each algorithm differed significantly. Figure 7 summarizes these results by normalizing the number of identified spike-segments for each algorithm to that identified by *ST*. Signal segments identified by *VR* was higher as compared to that of *ST* in all three cases. Asynchronous cancelation and the resulting unpredictable transformation of the raw-signal may have contributed to increased frequency of threshold exceeding segments. Moreover the general lowering of the spike detection threshold (due to a smaller noise-floor) may also have led to a higher number of threshold exceeding segments. Number of signal-segments identified by *DR* was comparable to *ST* in all three cases. Most noteworthy however, was that the *IEC* algorithm greatly reduced the number of spike-segments considered to be originating from local neurons. This is a predictable result since the input to the *IEC* algorithm is essentially all the spikes detected by *ST*; if correlated events were detected they would be rejected as originating from distant sources and thus lower the total number of signal segments identified by *IEC* as valid spikes. A useful outcome of this was that the reduction was statistically significant and increased with increase in common-noise (low common-noise: $p<0.103$, medium common-noise: $p<0.036$, high common-noise: $p<0.020$). On one hand, eliminating candidate spikes might at first appear to reduce the amount of information available for spike train analysis. But on the other, if a significant number of detected events are not actually produced by local neurons in the first place, use of *IEC* may help in reducing the volume of data to be considered for further processing. This may help ease the design requirements for implantable wireless interfaces by reducing transmission bandwidth.

4. Discussion

Non-neuronal signals like EMG and spurious noise generated by abrupt movements of awake and behaving animals negatively impact the neural-spike detection ability of conventional algorithms like threshold-based detection and principal component analysis. Hence the goal of this study was to evaluate different algorithms for their spike detection efficacy. Two novel approaches: virtual referencing (*VR*) and inter-electrode correlation (*IEC*) were developed and compared with traditional approaches of simple-thresholding (*ST*) and differential referencing (*DR*). It was found that the *IEC* algorithm is a robust way of improving neural spike detection and reducing false positives and in comparison, its performance was less influenced by degree of common-noise in the raw data. By identifying potential false positives, *IEC* will also be able to significantly reduce the number of neural spikes for consideration in downstream processing.

4.1 Simple Thresholding

The simplicity of threshold-based detection makes it an attractive first step in many applications. In this study, simple thresholding (*ST*) performed adequately in instances of low common-noise. But its performance deteriorated significantly in cases of medium and high common-noise as might be expected. Thus threshold detection should be supplemented with techniques that would reduce the false-detection of non-neuronal signals. Moreover the threshold selected for this type of detection is completely dependent on quality of raw data. Presence of common-noise originating from animal movement, mastication and other non-neuronal sources may artificially increase the threshold level thereby resulting in missed detections. Hence selection of a lower threshold may be investigated to reduce such occurrences followed by a cleaning algorithm such as the *IEC* technique proposed here to reduce false-positives.

4.2 Differential Referencing

Differential referencing (*DR*) represents a straight-forward solution to reduce impact of common-noise. However, the analysis carried out in this report suggests that spike detection outcomes from differential recording are highly dependent on the conditions at the differential reference. It performs poorly when neural spike activity is also present on the reference. In addition, it was found that threshold exceeding common-noise may have a slight and unpredictable lag across electrodes (possibly due to difference in tissue and electrode impedance properties) resulting in sub-optimal cancellation by differential reference subtraction. As a result, the classic *DR* algorithm seemed to have more limitations than the *VR* or *IEC* methods. Also important to consider is that differential referencing results in manipulation of the signal at its source. This limits its use in studies aimed at temporal monitoring of the true tissue-electrode interface quality since changes in the state or choice of the differential reference electrode would confound interpretations.

4.3 Virtual Referencing

Virtual referencing (*VR*) is similar in concept to the common average reference technique employed in EEG recordings (McFarland et al., 1997). The potential of spike-like signals appearing on the electrode of interest due to the presence of neural spikes on the reference electrode are greatly mitigated in this case. The *VR* method may also be more resilient to unpredictable phase lags because of the inherent spatial averaging involved. The technique helps lower the overall noise floor of the signal resulting in a lower spike-detection threshold. This was reflected in a consistent increase in the number of spikes detected after virtual reference preprocessing under all conditions of common-noise. The shape of the resulting mean-spike was also fairly neuronal but smaller in amplitude, probably due to the increase in the number of small spikes that can be detected with the decreased noise floor. Thus virtual referencing may be a good approach in tackling problems associated with threshold exceeding

non-neuronal signals. However, as in the case of differential recording, the underlying manipulation of the original signal at its source may reduce (albeit to a lesser extent) the reliability of this technique in objectively monitoring the tissue-electrode interface quality of a specific electrode sites over time.

4.4 Inter-electrode Correlation

The underlying logic behind implementation of inter-electrode correlation is supported by modeling studies for layer V pyramidal cells (Gold et al., 2006; Moffitt and McIntyre, 2005). These studies suggest a rapid change in characteristics (*P2P* amplitudes and *DepPhase*) of recorded neuronal spikes as their distance from the electrode increases. Hence it is reasonable to assume that the probability of an individual spike originating from a local neuron producing a highly correlated event across the array to be very low. By eliminating correlated segments, the *IEC* algorithm allows for true neuronal events to have greater influence in the generation of the mean-spike. Results suggest that among the four algorithms evaluated, the inter-electrode correlation (*IEC*) algorithm is least affected by the degree of common-noise in raw recordings, consistently producing mean-spikes with sharp, clearly defined depolarization/repolarization phases. Quantitatively, the mean-spike *P2P/DepPhase* ratios produced by the *IEC* algorithm for cases of medium and high common-noise were the largest among all algorithms, suggesting a higher contribution from local neural sources in generating the mean-spike. The variance in mean-spike features was also lowest with *IEC* in most instances (64 out of 72) as compared to the other methods. These results taken together suggest that *IEC* effectively enhanced the contribution of signal segments originated from similar sources (i.e. local neurons) while reducing the contamination of random common-noise artifacts.

Statistically significant improvements (over that of *ST*) in the *P2P/DepPhase* ratios were not observed as might have been initially predicted from the modeling studies. This should not however detract from the discussed benefits offered by the *IEC* algorithm since the mean-spike features themselves have limitations in predicting performance in all situations. Regardless, there are logical reasons that may explain the lack of statistical significance observed in the *P2P/DepPhase* ratios. Most of the data chosen for the medium and high common-noise cases was from electrodes implanted for more than 10 days. Hence it is possible that the spike activity and spike amplitudes may have been reduced due to reactive tissue response and associated loss of neurons (Polikov et al., 2005). The decreased *P2P* amplitude was observed with all algorithms when common-noise was high. The reduction in the mean-spike amplitudes was less for results from *ST* as compared to *IEC*. A plausible explanation is that non-neuronal signals generated by distant sources might be less influenced by the reactive tissue response close to the implant, whereas local neurons and the associated signal paths might be affected more considerably by this. Thus a lower *P2P* amplitude observed in results of *IEC* algorithm as compared to that of *ST* under high common-noise may have been simply due to the fact that the highly correlated spike events that were eliminated were in fact larger than the local neural spike signals themselves. Despite the reduced mean-spike *P2P* amplitude, the *P2P/DepPhase* ratio was still larger for *IEC* mainly due to the statistically significant reduction in the duration of the depolarization phase resulting in a much more neural-appearing mean-spike. Other confounding factors are the high variability in the mean-spike features obtained with the *ST* algorithm as shown in the spread plots of Figure 6 (for medium and high common-noise). It should also be remembered that the valid spike events detected by the *IEC* algorithm are actually a subset of the *ST* spikes and this therefore will hamper the separation of the means.

Perhaps the most surprising outcome of the study was the proportion of spike events detected by *ST* that are actually classified as highly correlated events by the *IEC* algorithm (see Figure 7). The proportion of the *ST* spikes that were highly correlated also increased significantly with the degree of common noise. If one were to process the neural recordings on a single electrode

by electrode basis, the likelihood that there would be a large number of false positive spikes (even after careful sorting) would be high. Thus *IEC* offers a simple solution to "cleaning" neural data when the conditions of the electrode-tissue interface begin to degrade. Furthermore removing highly correlated events from consideration by downstream processing would ease computational requirements and reduce the data that must be transmitted by wireless recording systems currently under development.

Another advantage of the *IEC* method is that it does not manipulate the original signal recorded on the electrode. It is highly automated in that it does not require any subjective inputs and it can be seamlessly integrated with any spike detection algorithm as long as multiple recording sites are present. By not manipulating the electrical signal present at the electrode site of interest it allows for an objective evaluation of the health of the tissue-electrode interface. In future the approach could be utilized in the comparison of chronic recording performance obtained from different electrode designs and treatment strategies by measuring key parameters (like amplitude) of the detected spike segments in an objective fashion.

The proposed *IEC* technique has been applied on microwire electrode arrays. Its extension to other electrode geometries and structures (especially tetrodes) is yet to be established. Weighting factors based on tissue conductance characteristics and recording site separation may be required to effectively apply correlation based identification of non-neuronal signals in those scenarios. In extremely rare cases, it is possible that true neural spikes do get recorded on two electrodes simultaneously. A refinement of this technique wherein a spike-segment is eliminated only when it is highly correlated across more than two electrodes is being explored. It is possible that neuronal signals are hidden within correlated segments. Development of subordinate techniques to extract them will be useful in further improving the spike-detection efficacy of intracortical electrodes. In those cases, application of correlation algorithm may help identify regions of elevated non-neuronal activity on which more involved processing might be required. The correlation threshold beyond which a spike is considered non-neuronal can also be adjusted to meet the specific requirements of the experiment. For example, if the goal is to have maximum detection, a higher correlation threshold can be selected whereas if the goal is to have minimal false-positive detections a lower correlation threshold can be selected.

4.5 Other Algorithms

A number of methods have been proposed in the literature for detection of neural spikes from intracortical microelectrode recordings. Many of these methods are tested on simulated data that closely resembles real-data (Chandra and Optican, 1997; Jacob Vogelstein et al., 2004; Menne et al., 2002). Although such an approach has its advantages, especially in the initial phase of algorithm development, their validation on real data that has unique noise contributions is acutely desired. Moreover very few reports use data available on other electrodes in an array to improve spike detection and sorting. Some methods like that adopted by Brier et al. (Bierer and Anderson, 1999; Snellings et al., 2006), have successfully used data from other electrodes in array processing strategies to reduce noise thereby improving the signal-to-noise ratio. A distinguishing feature of the *IEC* approach is that correlation is calculated only amongst high amplitude ($>3\sigma$) candidate neural spike segments. The two algorithms proposed here can be implemented seamlessly with minimal processing overhead. They thus aim to clean the recorded data and complement the detection capabilities of conventional algorithms. Again an added benefit of the IEC approach is that it provides a virtually unadulterated view of the recording performance of a particular electrode. The only difference between it and simple-thresholding is that highly correlated events are removed from analysis. Finally, it should be re-emphasized that the goal of these algorithms is to supplement and not to replace the conventional spike detection and sorting algorithms.

5. Conclusion

An objective comparison of different techniques aimed at removing high amplitude common-noise (non-neuronal) signals from intracortical microelectrode arrays was carried out in this study. Results suggest that effectiveness of conventional approaches like differential referencing may be highly dependent on experimental conditions. Virtual referencing scheme represents an improvement to the classic differential referencing approach but suffers from similar drawbacks. Use of inter-electrode correlation to distinguish local neuronal sources from distant, presumably non-neuronal, signals offers a very robust and reliable solution to the identification and elimination of the latter. In removing these false-positive events, it also eases the computational load on downstream processing steps as well as the data transmission burden on wireless recording systems.

References

Bierer SM, Anderson DJ. Multi-channel spike detection and sorting using an array processing technique. Neurocomputing 1999;26-27:947–56.

Chandra R, Optican LM. Detection, classification, and superposition resolution of action potentials in multiunit single-channel recordings by an on-line real-time neural network. IEEE Trans Biomed Eng 1997;44:403–12. [PubMed: 9125825]

Chapin JK, Moxon KA, Markowitz RS, Nicolelis MA. Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex. Nat Neurosci 1999;2:664–70. [PubMed: 10404201]

Gilmour TP, Krishnan L, Gaumond RP, Clement RS. A comparison of neural feature extraction methods for brain-machine interfaces. Conf Proc IEEE Eng Med Biol Soc 2006;1:1268–72. [PubMed: 17946886]

Gold C, Henze DA, Koch C, Buzsaki G. On the origin of the extracellular action potential waveform: A modeling study. J Neurophysiol 2006;95:3113–28. [PubMed: 16467426]

Hoogerwerf AC, Wise KD. A three-dimensional microelectrode array for chronic neural recording. IEEE Trans Biomed Eng 1994;41:1136–46. [PubMed: 7851915]

Isaacs RE, Weber DJ, Schwartz AB. Work toward real-time control of a cortical neural prothesis. IEEE Trans Rehabil Eng 2000;8:196–8. [PubMed: 10896185]

Jacob Vogelstein R, Murari K, Thakur PH, Diehl C, Chakrabartty S, Cauwenberghs G. Spike sorting with support vector machines. Conf Proc IEEE Eng Med Biol Soc 2004;1:546–9. [PubMed: 17271734]

Lebedev MA, Nicolelis MA. Brain-machine interfaces: past, present and future. Trends Neurosci 2006;29:536–46. [PubMed: 16859758]

Lewicki MS. A review of methods for spike sorting: the detection and classification of neural action potentials. Network 1998;9:R53–78. [PubMed: 10221571]

McFarland DJ, McCane LM, David SV, Wolpaw JR. Spatial filter selection for EEG-based communication. Electroencephalogr Clin Neurophysiol 1997;103:386–94. [PubMed: 9305287]

Menne KML, Folkers A, Malina T, Maex R, Hofmann UG. Test of spike-sorting algorithms on the basis of simulated network data. Neurocomputing 2002;44-46:1119.

Moffitt MA, McIntyre CC. Model-based analysis of cortical recording with silicon microelectrodes. Clin Neurophysiol 2005;116:2240–50. [PubMed: 16055377]

Musial PG, Baker SN, Gerstein GL, King EA, Keating JG. Signal-to-noise ratio improvement in multiple electrode recording. J Neurosci Methods 2002;115:29–43. [PubMed: 11897361]

Nordhausen CT, Maynard EM, Normann RA. Single unit recording capabilities of a 100 microelectrode array. Brain Res 1996;726:129–40. [PubMed: 8836553]

Paralikar K, Rao C, Clement RS. Automated reduction of non-neuronal signals from intra-cortical microwire array recordings by use of correlation technique. Conf Proc IEEE Eng Med Biol Soc 2008;2008:46–9. [PubMed: 19162590]

Paralikar KJ, Clement RS. Collagenase-aided intracortical microelectrode array insertion: effects on insertion force and recording performance. IEEE Trans Biomed Eng 2008;55:2258–67. [PubMed: 18713695]

Polikov VS, Tresco PA, Reichert WM. Response of brain tissue to chronically implanted neural electrodes. J Neurosci Methods 2005;148:1–18. [PubMed: 16198003]

Rennaker RL, Street S, Ruyle AM, Sloan AM. A comparison of chronic multi-channel cortical implantation techniques: manual versus mechanical insertion. J Neurosci Methods 2005;142:169–76. [PubMed: 15698656]

Rousche PJ, Pellinen DS, Pivin DP Jr, Williams JC, Vetter RJ, Kipke DR. Flexible polyimide-based intracortical electrode arrays with bioactive capability. IEEE Trans Biomed Eng 2001;48:361–71. [PubMed: 11327505]

Sasaki K, Ono T, Nishino H, Fukuda M, Muramoto KI. A method for long-term artifact-free recording of single unit activity in freely moving, eating and drinking animals. J Neurosci Methods 1983;7:43–7. [PubMed: 6834886]

Schwartz AB, Cui XT, Weber DJ, Moran DW. Brain-controlled interfaces: movement restoration with neural prosthetics. Neuron 2006;52:205–20. [PubMed: 17015237]

Snellings A, Anderson DJ, Aldridge JW. Improved signal and reduced noise in neural recordings from close-spaced electrode arrays using independent component analysis as a preprocessor. J Neurosci Methods 2006;150:254–64. [PubMed: 16430966]

Stieglitz T, Schuettler M, Koch KP. Implantable biomedical microsystems for neural prostheses. IEEE Eng Med Biol Mag 2005;24:58–65. [PubMed: 16248118]

Strumwasser F. Long-term recording' from single neurons in brain of unrestrained mammals. Science 1958;127:469–70. [PubMed: 13529005]

Suner S, Fellows MR, Vargas-Irwin C, Nakata GK, Donoghue JP. Reliability of signals from a chronically implanted, silicon-based electrode array in non-human primate primary motor cortex. IEEE Trans Neural Syst Rehabil Eng 2005;13:524–41. [PubMed: 16425835]

Vargas-Irwin C, Donoghue JP. Automated spike sorting using density grid contour clustering and subtractive waveform decomposition. J Neurosci Methods 2007;164:1–18. [PubMed: 17512603]

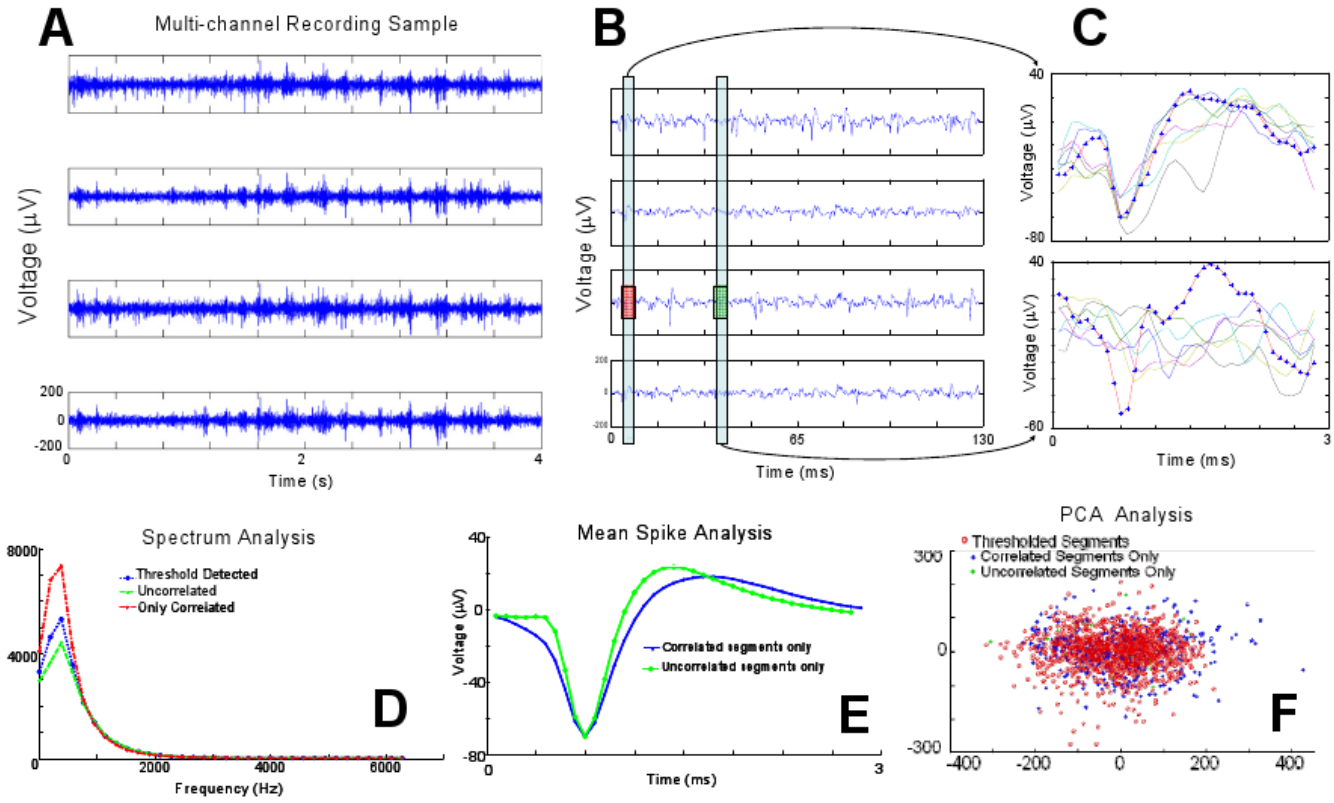

Figure 1. Problem Description

(A) Four-second raw recording segment from four microelectrodes spaced 250μm apart with evident common-noise. **(B)** Sample of same recordings on smaller time-scale (130ms) shows examples where threshold exceeding activity is recorded on all (red box; left) or only one electrode (green box; right). **(C)** Extracted signal-segments concurrently recorded at the same time points indicated in B (candidate spikes are identified in red with data point symbols visible). The appearance of correlated events (top) can closely resemble typical neural spike shape (bottom). **(D)** An FFT spectrum analysis of all threshold exceeding segments (threshold detected), segments with spikes only occurring on the electrode of interest (uncorrelated) and segments with events showing high correlation across the array, reveals that their frequency characteristics overlap. **(E)** Similarly a mean-spike generated from candidate spike-segments with high degree of inter-electrode correlation as compared with candidate spikes that appear only on the electrode of interest show similar features which could make template-based identification difficult. **(F)** The aforementioned spike segments also show over-lap in 2-D principal component space.

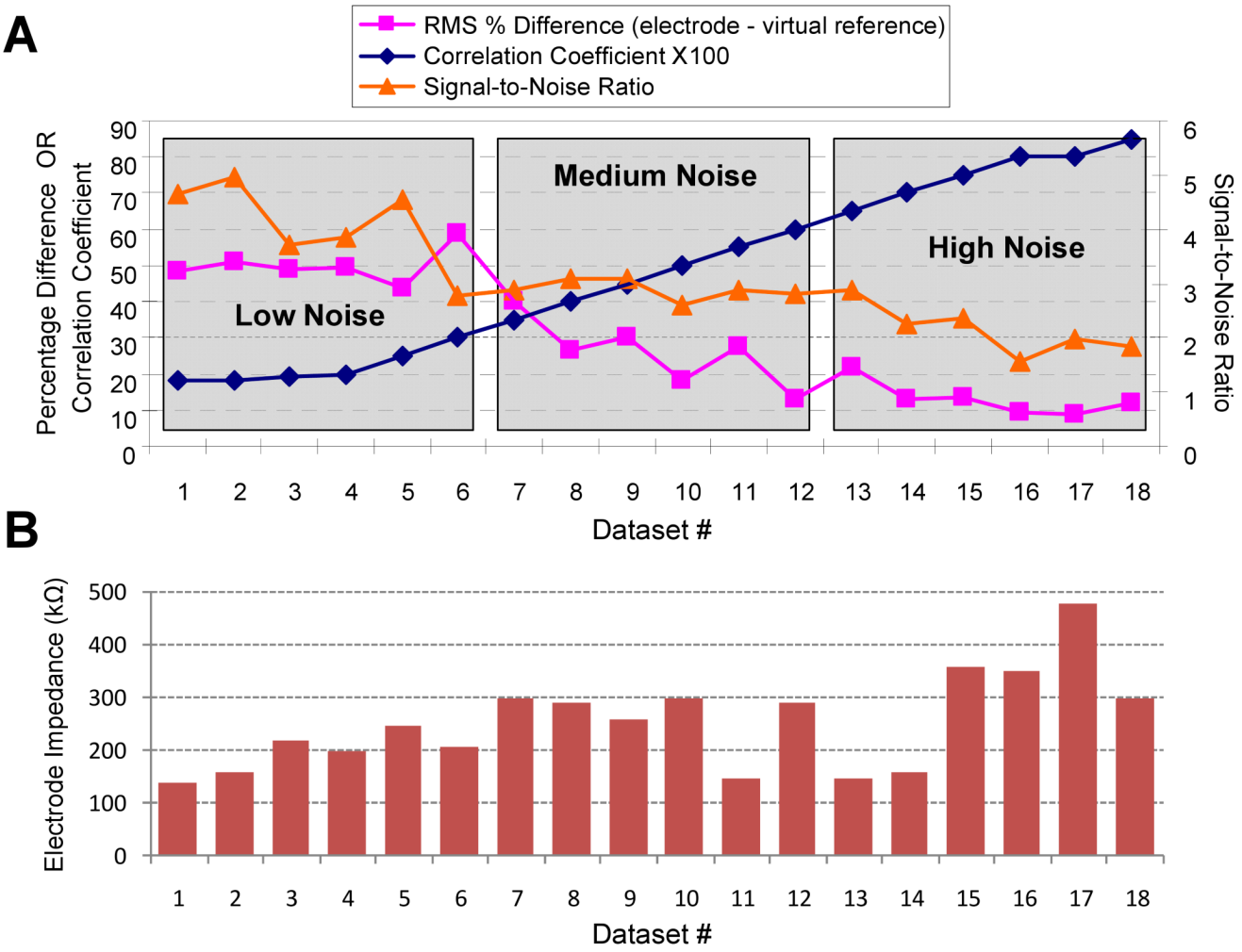

Figure 2. Objective Data Selection

(A) Data sets were selected so that algorithm performance could be evaluated for a range of common-noise content. Data sets were composed of 5-minute recordings from unanesthetized animals that received 8-channel implants. They are ranked in order of observed common-noise content (1 is lowest, 18 is highest) as determined by the average correlation coefficient among all electrode signals over the 5-minute session. The difference in the RMS energy between the virtual reference (created by averaging all signals in the array) and the analyzed electrode signals showed an opposite trend. Taken together this means that datasets with high correlation also had similar RMS energy levels across the array. Signal-to-noise ratio as defined in the Methods is also presented to illustrate the quality of recordings used in this study. **(B)** Electrode impedance magnitude (at 1kHz) of the test electrode measured the same day the neural recordings for the given dataset were obtained.

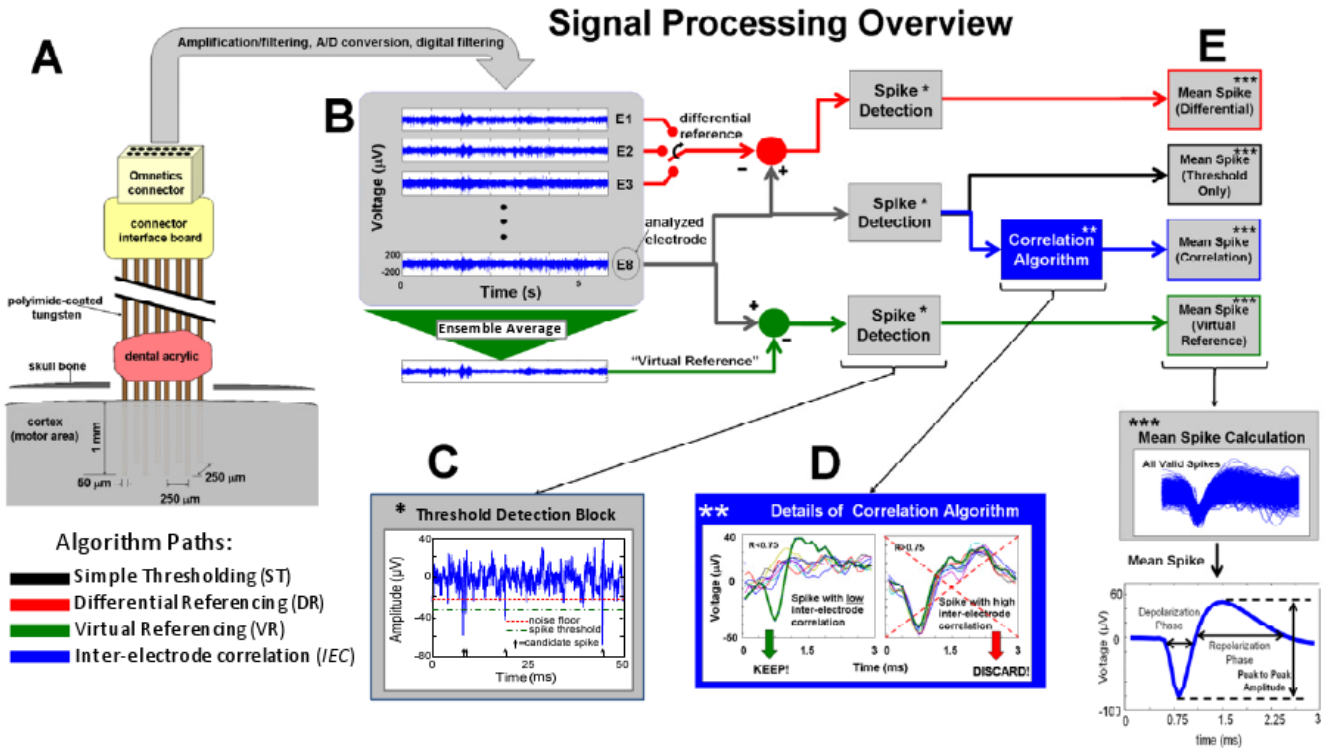

Figure 3. Signal Processing Overview

(A) Detailed schematic of the implanted electrode array with approximate dimensions. **(B)** Simultaneous recording from an eight channel array including the resultant ensemble average (used for virtual reference). **(C)** Signal segments that exceeded a threshold of 3 standard deviations of the signal (green line; dash-dot) were extracted as candidate neural spikes (3ms segments). **(D)** The inter-electrode correlation (*IEC*) algorithm seeks to eliminate candidate spikes that are highly correlated with signal segments recorded simultaneously on other electrodes in the array. The correlation coefficient between each candidate spike and the corresponding signal segment on each of the electrodes is determined. For the *IEC* algorithm, all spikes with correlation coefficient R>0.75 are classified as nonneuronal and discarded from further analysis. **(E)** At the end of processing of each data set, a mean-spike was computed from all valid spikes detected from each of the four algorithms (simple-thresholding, differential reference, virtual reference, and inter-electrode correlation; *ST*, *DR*, *VR*, and *IEC*, respectively). Several features were features from the mean-spikes as indicated for comparison purposes as indicated in the text.

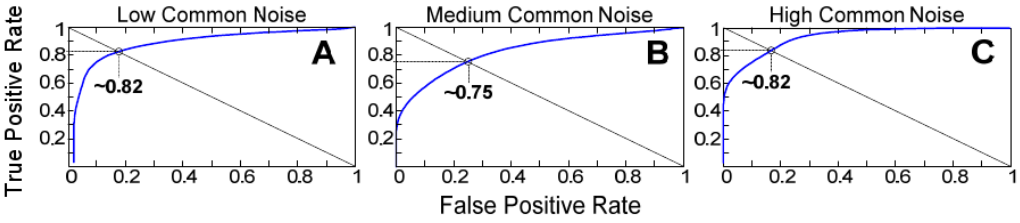

Figure 4. Receiver operating characteristic (ROC) curves for the determination of optimal correlation coefficient threshold

From left to right (A-C) the data sets analyzed increased in their common-noise content. The point of equal error (miss rate equal to false positive rate) is around 0.8. Hence selection of 0.75 as a correlation coefficient threshold for accepting/rejecting spikes was conservative and tended to favor acceptance of false-positives.

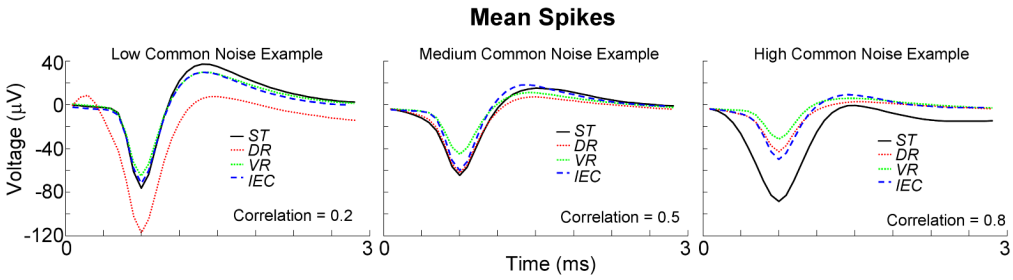

Figure 5. Examples of mean-spike shapes

Examples of mean-spikes obtained by the four different methods for varying degrees of common-noise are plotted together. Correlation numbers correspond to the average correlation observed across the entire 5-minute duration of recording between the test electrode and all other functional electrodes in the array.

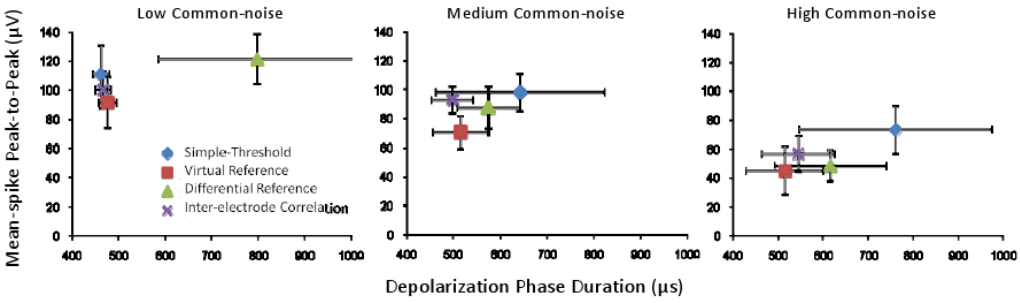

Figure 6. Mean-Spike Features

Plot of the spread of mean-spike peak-to-peak (*P2P*) amplitude versus depolarization phase (*DepPhase*) duration for all datasets considered.

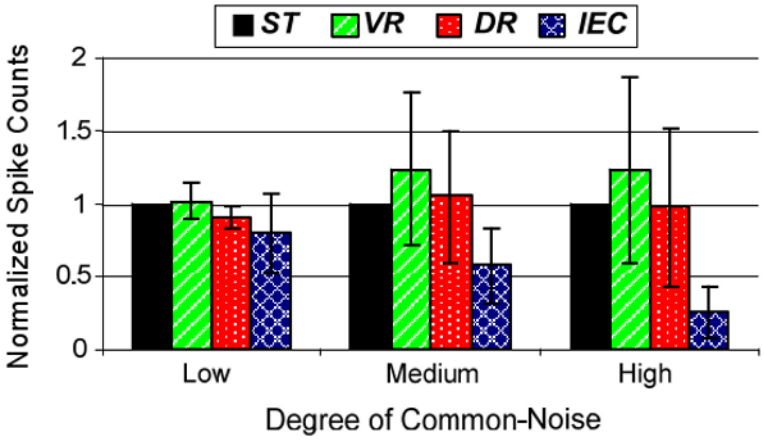

Figure 7. Comparison of Spike Detection

Comparison of number of valid spike-segments detected by the four algorithms. The original numbers are normalized with respect to those detected by standard thresholding technique. Error bars indicate the standard deviation in the data.

Table I

Summary of Mean-Spike Feature Analysis Results

			Spike-Detection Algorithm			
		Mean-Spike Feature	***ST***	***DR***	***VR***	***IEC***
Degree of common-Noise	**Low**	DepPhase (μs)	461.3±18	796.8±212.4**	475.6±19.8	465.8±17
		RepPhase (μs)	1401.7±268.6	883.6±542.9	1304±183.5	1254.2±131.7
		P2P Amp (μV)	110.6±19.8	121.2±17.2	91.5±17.7	100±12.6
		# Spikes	11055±2432	9957±1501	11211±1893	8791±1914
		P2P/DepPhase Ratio	0.24±0.05	0.16±0.04*	0.19±0.04	0.22±0.03
	Medium	DepPhase (μs)	641.9±197.4	575±72.8	515±1467.1	497.6±49.2
		RepPhase (μs)	1357.5±410.5	1355.9±205.3	1467.1±63.4	1359.2±79.6
		P2P Amp (μV)	98.2±14.1	87.6±15.5	70.4±12.2**	93±10
		# Spikes	7713±1982	8088±1979	9513±2760	4407±2688**
		P2P/DepPhase Ratio	0.16±0.04	0.16±0.04	0.14±0.03	0.19±0.02
	High	DepPhase (μs)	760.3±214.3	615.6±124.7	514.3±85.9*	544.6±81.9*
		RepPhase (μs)	1083±662.4	1600.1±222.8	1628.8±183.5	1308.4±328.5
		P2P Amp (μV)	73.1±16.5	48.3±10.8*	44.83±16.9*	56.6±12.1
		# Spikes	4735±2527	4598±3018	5823±3636	1206±1859*
		P2P/DepPhase Ratio	0.1±0.03	0.08±0.04	0.09±0.05	0.11±0.04

Note: All values are expressed as mean ± standard deviation. DepPhase = depolarization phase duration, RepPhase = repolarization phase duration, P2P Amp = peak-to-peak amplitude, # Spikes = number of valid spikes identified by algorithm, See section 2.5 and Figure 3E for more details. Six datasets were analyzed per entry (N=6).

* indicates statistically significant difference (p<0.05) compared to outcome of standard thresholding '*ST*', as obtained by performing one-way analysis of variance (ANOVA).

** indicates p<0.01.