Quality Metrics of Spike Sorting Using Neighborhood Components Analysis

Xinyu Liu, Hong Wan[*] and Li Shi[*]

School of Electrical Engineering, Zhengzhou University, Zhengzhou 450001, PR China

Abstract: While an electrode has allowed for simultaneously recording the activity of many neurons in microelectrode extracellular recording techniques, quantitative metrics of cluster quality after sorting to identify clusters suited for single unit analysis are lacking. In this paper, an objective measure based on the idea of neighborhood component analysis was described for evaluating cluster quality of spikes. The proposed method was tested with experimental and simulated extracellular recordings as well as compared to *isolation distance* and L_{ratio}. The results of simulation and real data from the rodent primary visual cortex have shown that values of the proposed method were related to the accuracy of spike sorting, which could discriminate well- and poorly-separated clusters. It can apply on any study based on the activity of single neurons.

1. INTRODUCTION

Studies of neuronal activity throughout the brain predominantly rely on extracellular recordings of spiking activity [1]. Since an electrode often simultaneously records spiking activity more than one neuron, spikes are always processed by sorting algorithms. Nevertheless, spike sorting algorithms are not perfect and classification errors may occur because of a number of reasons. Such as the recording hardware, electromagnetic interferences, the superimposed activity of multiple neurons, the spatially averaged activity of distant neurons and so on. Therefore, it is imperative to quantify unit quality after sorting for evaluating objectively results of spike sorting as well as improving the stability and reliability of neuron coding or decoding.

In most cases, the quality assessment was done subjectively by a human observer [2, 3], but the subjective methods don't permit comparison of data quality across different studies, and unfortunately they are predisposed to personal bias. Therefore, several objective methods have recently been proposed for measuring unit quality. These methods can be classified into four categories: (a) Methods based on signal-to-noise ratio (SNR). These approaches have often served as a quantitative measure of unit quality in traditional microelectrode techniques because their significance is clear and can be easily calculated [4]. But they are not available when multiple units are recorded by the same electrode. (b) Methods based on refractory period. If spikes are from the same neuron, then there should be no interspike intervals less than the refractory period of neuron, which for most neurons is no less than 1 ms [5]. But these

methods are only suited for checking the quality of isolation over a long collection period. (c) Methods based on statistic distribution. For example, Pouzat *et al.* [6] assumed that the noise obeys a Gaussian distribution, which used to evaluate the variability of spike waveforms. The L_{ratio} proposed by Schmitzer-Torbert *et al.* [4] used a Chi-square distribution as a distance measure of the noise events in a feature space. But, there is no general agreement about the statistic distribution of noise. (d) Methods based on distance measure. For instance, Harris *et al.* [7] and Schmitzer-Torbert *et al.* [4] introduced *isolation distance* to quantify the quality of clustering by the minimal distance where the number of spike events and noise events is equal. The drawback of these methods is that they do not have a global scale to differentiate between well and poorly isolated units. Joshua *et al.* [8] described *isolation score* algorithm based on Euclidian distance to evaluate the quality of unit. The heavy computational burden makes it be sensitive to the number of spikes.

In order to circumvent these problems, a method based on neighborhood components analysis (NCA) was proposed for evaluating cluster quality in this paper. The NCA is a learning method for clustering multivariate data into distinct classes according to a given distance metric over the data, which is always used to dimensionality reduction [9] and fast classification [10]. Here, the idea to use the NCA is not to discover new variables which could be of interest, but to use the idea of NCA to estimate the probability that every event is correctly classified, and further estimate spike's cluster quality. Functionally, the NCA serves the same purposed as the *K*-nearest neighbors algorithm, which makes no assumptions about the shape of the distribution of signals. So it can improve the estimate performance of cluster quality of spikes to some extent. In this method, each spike was firstly preprocessed through up-sampled, aligned and normalized; Then, the features of spike waveform were extracted by

*Address correspondence to these authors at the School of Electrical Engineering, Zhengzhou University, Zhengzhou 450001, PR China; Tel: +86 371 67781421; Fax: +86-0371-67783113; E-mails: wanhong@zzu.edu.cn; shili@zzu.edu.cn

wavelet transform and Kolmogorov-Smirnov test; Finally, the quality of cluster was estimated through the NCA. Simulations and real data from primary visual cortex (V1) of rats were used to demonstrate the utility of this method.

2. MATERIALS AND METHODS

2.1. Real Data

The real data were collected using a 16-channel microelectrode array (Microprobe Inc., 2×8 platinum-iridium microwire array, 250 µm inter-electrode spacing, electrode impedance at 1 kHz equals 0.5~1.0 MΩ) and recorded from experiments performed on V1 anesthetized Long Evans (LE) rats. All experimental procedures were approved by Life Science Ethical Review Committee of Zhengzhou University. Neural signals were amplified and filtered with a computer-controlled multichannel amplifier (Cerebus^{TM}, Blackrock Microsystem Inc.). Spike signals were band-pass filtered at 0.25~5 kHz by a digital two pole Butterworth filter and sampled at 30 kHz. The implant procedure and recording conditions were described elsewhere [11].

Threshold amplitude crossing was used to detect the spikes and the threshold was defined as $4\sigma_n$ [12]. $\sigma_n = \text{median}\{|x|/0.6745\}$, where x is the bandpass filtered signals and σ_n is an estimate of the standard deviation of background noise. And each spike was represented by using 2.1 ms (64 sampling points). The spikes were sorted by *WaveClus* approach. The *WaveClus*, which was proposed by Quiroga *et al.* [12], is an unsupervised spike sorting algorithm that combines wavelet transform and superparamagnetic clustering.

2.2. Synthetic Data

To test effectively the performance of proposed method, the simulated data from previously published data were used (Quiroga *et al.*, publicly available online, http://www.vis.caltech.edu/~rodri/publications) [13]. The simulated data include four sets of simulated extracellular recording signals with a sampling rate of 24 kHz and the noise level determined by background noise's standard deviation set to 0.05, 0.1, 0.15 and 0.2. Each simulation contains three distinct dominant single-unit spikes with normalized amplitude. In all simulations, three distinct spike trains have a Poisson distribution of inter-spike intervals with the mean firing rate of 20 Hz and the refractory period between spikes within the same category was set to 2 ms [12].

2.3. Neighborhood Components Analysis

Neighborhood components analysis proposed by Goldberger and colleges [10], which is a supervised learning method for clustering multivariate data into distinct classes according to a given distance metric over the data. The NCA aims at finding a distance metric that maximizes the leave-one-out error for a stochastic variant of *K*-nearest neighbors by finding a linear transformation of the input data. The key insight to the algorithm is to find a matrix *A* that corresponds to the transformation by differentiable objective function.

Setting consists of n real-valued input vectors $x_1, \cdots, x_n$ in R^D and corresponding class labels $c_1, \cdots, c_n$. And each vector x_i selects another vector x_j as its neighbor with some probability p_{ij} which was defined using a softmax over Euclidean distances in the transformed space,

$$p_{ij} = \frac{\exp(-\|Ax_i - Ax_j\|^2)}{\sum_{k \neq i} \exp(-\|Ax_i - Ax_k\|^2)}, \; p_{ii} = 0, \tag{1}$$

where $\|\cdot\|$ denotes Euclidean distances. Under this stochastic selection rule, we can compute the probability p_i that vector x_j will be correctly classified,

$$p_i = \sum_{j \in C_i} p_{ij} \tag{2}$$

where $C_i = \{j | c_i = c_j\}$ denote the set of points in the same class as i.

The objective that we maximize is the expected number of points correctly classified under this scheme,

$$f(A) = \sum_{i=1}^{n} p_i \tag{3}$$

which is the differentiable cost function. Then *A* can be obtained using an iterative solver, such as conjugate gradient descent. The NCA used a probabilistic way to obtain the differentiable cost function, in other words, the cost function can be calculated by the average probability that every point in sample set belongs to the same set.

2.4. Quality Metrics Based on Neighborhood Components Analysis

In extracellular recordings, spikes recorded on a microelectrode represent a mixture of spikes obtained from one or more neurons and some non-neural spikes, such as noise events from motion artifacts, electromyographic activity, electric field pickup and so on [14]. In this study, a quantitative metrics based on NCA was proposed for evaluating cluster quality of spikes. Note, the aim of NCA is not only to discover new features which could be of interest, but to also use the idea of NCA to estimate the probability that every event is correctly classified, and further to quantify cluster quality. The purpose of using NCA is always the data visualization and fast classification [14] but here we are discussing it from the cluster quality evaluation perspective.

Setting k spikes $s_1, \cdots, s_k$ from m neurons are recorded by a microelectrode and corresponding class labels $c_1, \cdots, c_k$ in $N^+ (1 \leq c_i \leq m)$. In other words, k spikes are sorted to m clusters $C_1, \cdots, C_m$. To reduce the variability of spike waveforms from the same neuron due to discrete sampling of analog traces, we firstly up-sampled the data using cubic spline interpolation, the factor was fixed at 2. Then, all events were aligned by the point of maximum amplitude and

normalized by standardized *z*-scores. All spikes after being prepressed were simply sorted two clusters, spike cluster and noise cluster. The spike cluster is a cluster classified as a single unit and the noise cluster is a cluster of events not classified with that unit in multi-units.

For each spike, a wavelet transform was calculated. A four-level multiresolution decomposition was implemented using Haar wavelets. The Haar wavelets were chosen due to their compact support and orthogonality, which allow the discriminative features of spikes to be expressed with a few wavelet coefficients and without a priori assumptions on the spike shapes. Thus, the Lilliefors modification of a Kolmogorov-Smirnov (KS) test [15] was used to select the optimal coefficients that best separate the different spike classes. The goal of this step is to reduce the differences among spikes from the same clusters and increase the differences among the different clusters, and improve estimation precision. In our method, the first 10 coefficients with the largest deviation from normality were used. For specific details please see reference [12].

For a given class, C_k contain N_k spikes, firstly computing the similarity between each event s_x in spike cluster, $s_x \in C_k$, to all other events s_y in noise cluster, $s_y \in C_1 \bigcup \cdots \bigcup C_{k-1} \bigcup C_{k+1} \bigcup \cdots \bigcup C_m$,

$$\text{similarity}(s_x, s_y) = \exp(-\|s_x - s_y\|^2 \lambda), \tag{4}$$

where $\|\cdot\|$ is the Euclidean distance between vectors s_x and s_y. The parameter λ is a gain contrast ($0 \le \lambda \le \infty$) that sets the gain of this stretch. For a given event s_x, when $\lambda \to 0$, all events are similar and similarity $(s_x, s_y) \to 1$. On the other hand, when $\lambda \to \infty$, all events are dissimilar and similarity $(s_x, s_y) \to 0$. Here, $\lambda = 0.9 \times d_0$, d_0 is the average Euclidian distance in spike cluster.

In order to turn the above similarity index into a probability-like quantity (positive values that sum to 1), the normalization was performed by the following formula,

$$P_x(y) = \frac{\text{similarity}(s_x, s_y)}{\sum_{z \ne x} \text{similarity}(s_x, s_z)}, p_x(x) = 0, \tag{5}$$

where $s_z \in C_1 \bigcup \cdots \bigcup C_m$. Note that here the similarity index was normalized by the sum of the similarity index between s_x and all other events from multiunits were recorded simultaneously.

For each event in spike cluster, s_x, all the normalized similarity values $P_x(y)$ were summed for all the s_y in the C_k cluster,

$$P(x) = \sum_{s_y \in C_k} P_x(y), \tag{6}$$

where $P(x)$ is the probability that spike s_x belongs to this cluster. The calculation of $P(x)$ was illustrated in Fig. (**1**).

Finally, quality of the cluster is computed as,

$$\text{NCA} = \frac{1}{N_k} \sum_{s_x \in C_k} P(x) \tag{7}$$

and can be intuitively considered as the average probability that an event classified as a spike belongs to the cluster. The range of the NCA is from 0 to 1. If $\text{NCA} \to 1$, which means a better cluster quality, in other words, with a small distance among spike events, and a large distance between them and noise events; If $\text{NCA} \to 0$, it means a worse cluster quality, i.e. events from spike cluster are surrounded by the noise events.

Fig. (**1**) depicts an illustration of the calculation of $P(x)$. Correctly categorized events are shown in Fig. (**1A**), false positive and false negative errors are shown in Fig. (**1B** and **C**), respectively. If noise events are classified as spikes by sorting algorithms (false positive errors), that are nonetheless close to noise cluster. The $P(x)$ value would be reduced, due to its proximity to the other noise events, thus will reduce the overall NCA value (see Fig. (**1B**)). Likewise, when spikes are missed (false negative errors), they are close to spike cluster. As a result, events in the cluster that are close to such misses will be reduced $P(x)$, which in turn again reduce the overall value of NCA (see Fig. (**1C**)).

2.5. Quality Metrics Based on *Isolation Distance* and *L*ratio

This subsection described two measures used to compare with the NCA, namely *isolation distance* and *L*ratio, and a complete description of the two methods can be found in [4]. The *isolation distance* is a measure of how well-separated a cluster is from the rest of the data set and *L*ratio indicates the distribution of non-cluster spikes around a cluster.

For cluster c, containing n_c spikes, the *isolation distance* is defined as the squared Mahalanobis distance of the n_c-th closest non-c spike to the center of c. The squared Mahalanobis distance is calculated as,

$$D_{i,c}^2 = (x_i - \mu_c)^{\mathrm{T}} \Sigma_c^{-1} (x_i - \mu_c), \tag{8}$$

where x_i is the vector containing features for spike i, and μ_c is the mean feature vector for cluster c. Σ_c is the covariance matrix of spikes in cluster c. The *isolation distance* is not defined when n_c is greater than the number of non-cluster spikes. A higher value indicates that non-cluster spikes are located farther away. Note that the value of *isolation distance* is not normalized against cluster size, so that clusters with a large number of spikes will tend to have a higher *isolation distance*.

*L*ratio is calculated as follows for cluster c,

$$L_{\text{ratio}}(C) = \frac{1}{n_c} \sum_{i \notin c} (1 - \text{CDF}_{\chi^2_{\text{d.f.}}}(D_{i,c}^2)), \tag{9}$$

where $D_{i,c}^2$ is the squared Mahalanobis distance between non-c spike i and the center of c;
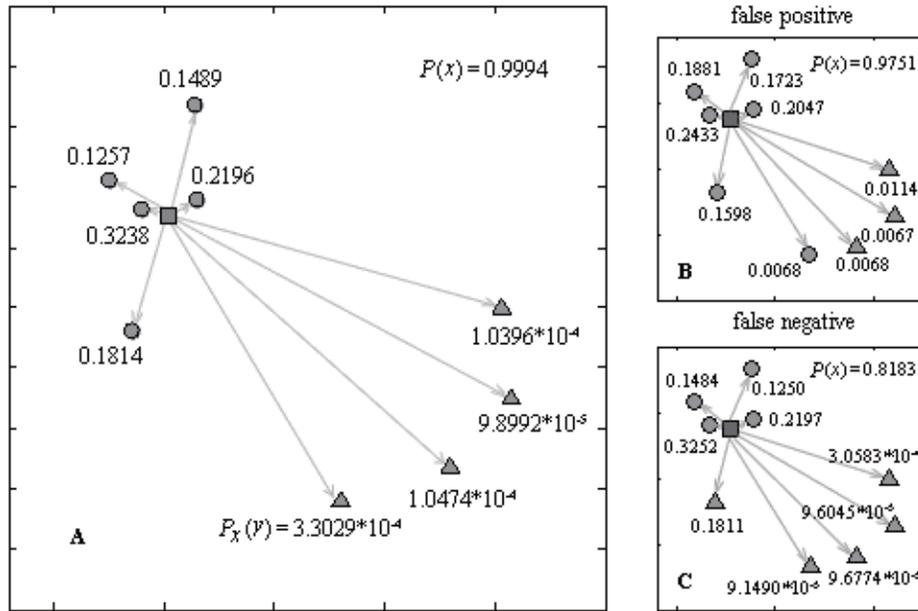
Fig. (1). Illustration of $P(x)$ **calculation**. This figure is a schematic representation. The x-y coordinates represent the m dimensions of a waveform from the spike and noise cluster. The gray circles represent spike events in the spike cluster, whereas the gray triangles represent points in the noise cluster. The gray square represents a given point in the spike cluster. (**A**) Schematic representation of correctly categorized events. For a given spike s_x in the spike cluster (gray squares), the numbers next to each of the other spikes, s_y, are $P_x(y)$. The arrows denote the Euclidean distance. $P(x)$ for the given point (gray squares), is the sum of all $P_x(y)$ values for all other spike events (gray circles). Note that for events far from s_x, $P_x(y)$ is infinitesimal, and hence they have only a small influence on the $P(x)$. (**B**) Schematic representation of false positive errors. The $P(x)$ values would be decreased when noise events are detected as the spikes, *e.g.* the gray circle in the down middle corner close to the noise events. (**C**) Schematic representation of false negative errors. The $P(x)$ values would also be significantly decreased when spike events are detected as the noises, *e.g.* the gray triangles in the left middle corner close by the spike events.

$\text{CDF}_{x^2_{\text{d.f.}}}$ is the Chi-squared cumulative distribution function describing the distribution of spikes in cluster c. The number of degrees of freedom is equal to the number of features used in the cluster space. A low L_{ratio} indicates that there is a relatively empty space between the cluster and other spikes in the data set. Therefore, a lower L_{ratio} and a higher *isolation distance* together indicate better cluster quality.

In addition to *isolation distance* and L_{ratio}, the SNR was calculated for cluster c. The calculating formula of the SNR is as follows [3]:

$$\text{SNR} = \frac{1}{n_c} \sum_{i=1}^{n_c} \frac{\max(s_i) - \min(s_i)}{2 \times \text{std}(\varepsilon_i)} \qquad (10)$$

where s_i is a waveform vector of spike i; $\text{std}(\varepsilon_i)$ is the standard deviation of the noise ε_i, $\varepsilon_i = s_i - \overline{s}$ is the average spike waveform in cluster c. The SNR is a traditional method for evaluating quality of a cluster, a low SNR indicates worse cluster quality.

3. RESULTS

3.1. Validation of Algorithm by Simulation Data

To evaluate the performance of the proposed algorithm, a signal with the noise level of 0.1 was selected, which contains 3 well-sorted spike clusters and 521 spikes. The events which were included in a circle of radius r around center of a given cluster composed the spike cluster and the rest of the events composed the noise cluster. The false negative and false positive errors could be simulated as the radius increase or decrease. The results are shown in Fig. (**2**). The principal component (PC) projections, spike waveforms and two-dimensional histogram of waveforms are shown in Fig. (**2A**, **B** and **C**), respectively. In Fig. (**2A**), the red and black points represent events in the spike cluster and the noise cluster respectively. Note that the PC features only for representing conveniently representing the signals, other features, such as peak amplitude, energy, wavelet coefficients and so on, can also be selected.

Fig. (**2D**) shows the performance of the NCA and ACC methods. The ACC was defined as [16] ACC = TP/(TP + FN + FP), where TP (True Positive) indicates the number of spikes which were classified as the spike cluster; FN (False Negative) indicates the number of spikes which were classified as the noise cluster, FP (False Positive) indicates the number of noises which were classified as the spike cluster. The ACC provides a quantitative metrics of the false negative and false positive errors. The value closer to 1, the smaller the errors and the higher the cluster quality. As shown in Fig. (**2D**), as the cluster was moved out of the distribution of spikes (see the black circle in (**A**)), both measures improved. When the radius $r < 3.8$, the value of NCA is consistent with the ACC. It is clear that the NCA can
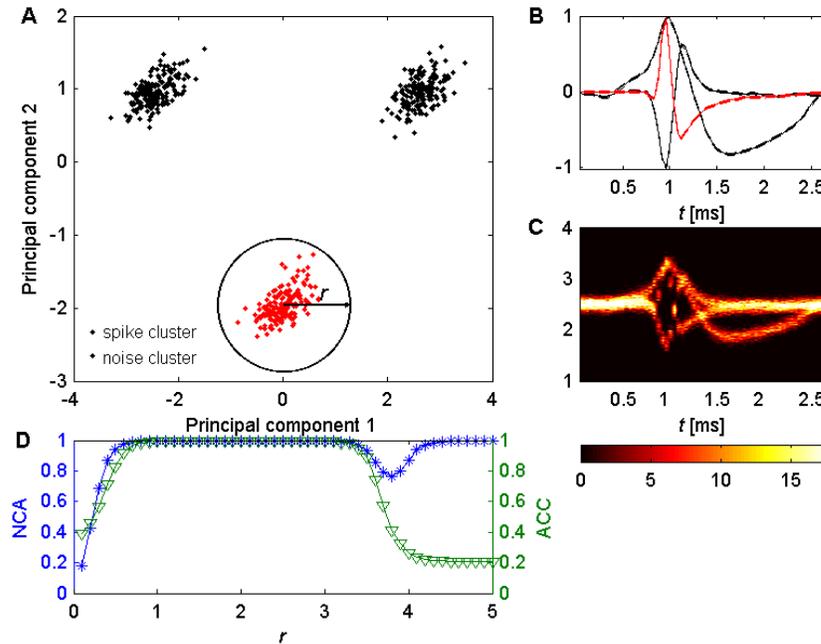
Fig. (2). Comparison of the cluster quality between the NCA and the ACC. (**A**) Scatter plot of spike waveforms from all events using the first two principal components. The red points denote the spike cluster and the black points denote the noise cluster (**B**). The average waveform for each cluster in (**A**). Thickness of the waveform represents the standard error (**C**). Two-dimensional histogram of waveforms as in (**A**). (**D**) Performance of the NCA and the ACC as a function of the radius r (as shown black circle in (**A**)).

effectively measure the quality of spike cluster for cases in which the number of cluster spikes is not far greater than the number of noise events.

To further evaluate the performance of the NCA method, it was compared with L_{ratio} (LRO), *isolation distance* (ISD), and SNR. Fig. (**3**) shows a comparison of the cluster quality among the NCA, the SNR, the ACC, the ISD and the LRO. For an easy comparison among them, the ISD, the LRO and the SNR were normalized by their maximum respectively. As shown in Fig. (**3**), with the increase of radius r, the values of NCA, ISD and ACC increased at first and then

decreased, the LRO, by contrast, descend at first and then increased, and the SNR increased all along. All methods attain the stability value when the ACC reaches its maximum value except the SNR. This demonstrates that these methods can be successfully used for measuring the cluster quality. Compared to other methods, however, the NCA is the nearest to the ACC. Although the performance of the LRO and the ISD is similar to the NCA, they don't have a global scale to differentiate between good and poor separation. The SNR on spike cluster also increased with the radius r, but did not provide an unambiguous measure of cluster quality.
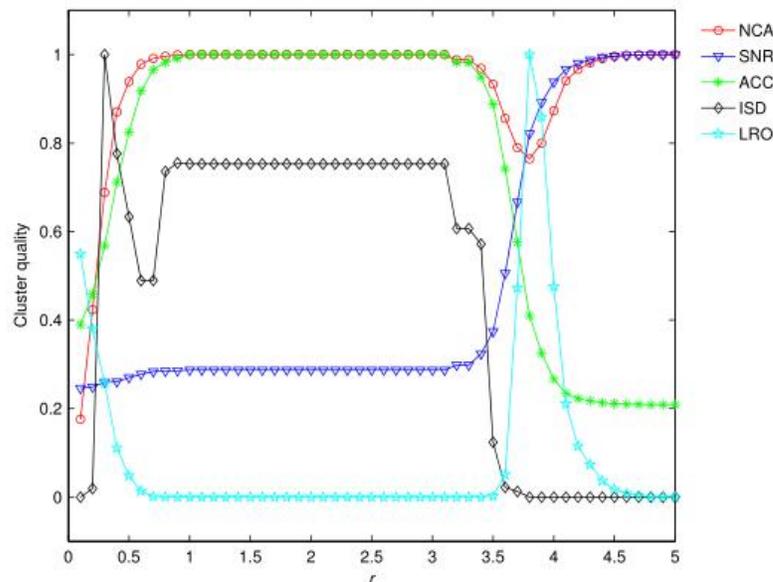

Fig. (3). Performance of the NCA, the SNR, the ACC, the ISD and the LRO method as a function of the radius r.

## 3.2. Validation of Algorithm by Real Data

Although the NCA performed well on simulated data, it was also important to demonstrate its performance using real data. Figs. (4−6) show three signals taken from micro-electrode array collected in rat V1 area. Fig. (4) shows data from the same electrode recording in which a well-separated cluster of spikes was observed. Histograms of the projection of each cluster onto Fisher's linear discriminant between the spike cluster and the noise cluster are also shown (the linear discriminant, H, between the features of spike event s and noise event n were given by $H = (\bar{\bar{\Sigma}}_s + \bar{\bar{\Sigma}}_n)^{-1}(\mu_s - \mu_n)$ where μ_s and μ_n are the mean vectors $\bar{\bar{\Sigma}}_s$ and $\bar{\bar{\Sigma}}_n$ are the covariance matrices of the features, respectively. The projection of an event V onto the linear discriminant was given by $H \times V$). The values of LRO, ISD, SNR and NCA

for spike cluster were indicated through the black text on the right top corner of figure. Dots are the measured distribution and the solid line is the fitted curve using Gaussian function. For both clusters, a very little overlap of the projection of the spike and noise cluster was observed, and values of the NCA indicated that the clusters were well separated.

As can be seen in the PC projections and the histograms of the Fisher's projection, the spike cluster in Figs. (5 and 6) are not as well separated as the examples shown in Fig. (4). The worst separation was observed in Fig. (6), while Fig. (5) had an intermediate separation. The values of NCA agreed with this subjective categorization of cluster quality. The clusters shown in Fig. (4) had the best quantitative separation for both measures, Fig. (5) had an intermediate value, and Fig. (6) had the worst values of NCA. These examples support the use of these cluster quality measures as an objective method for evaluating cluster quality.
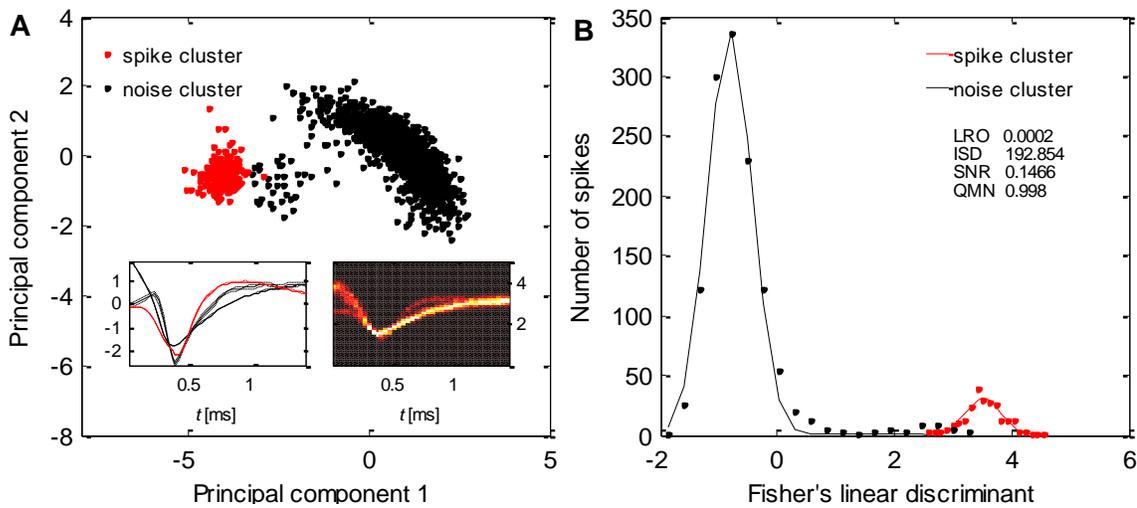
Fig. (4). Real data with good separation. (**A**) Scatter plot of spike waveforms from all events using the first two principal components. The spike events and the noise events are indicated by red and black, respectively. (**B**) Histograms of the projection of each cluster onto Fisher's linear discriminant between the spike cluster and the noise cluster. Dots indicate measured distribution of the projection of each cluster and solid line indicates fit curve using Gaussian function.
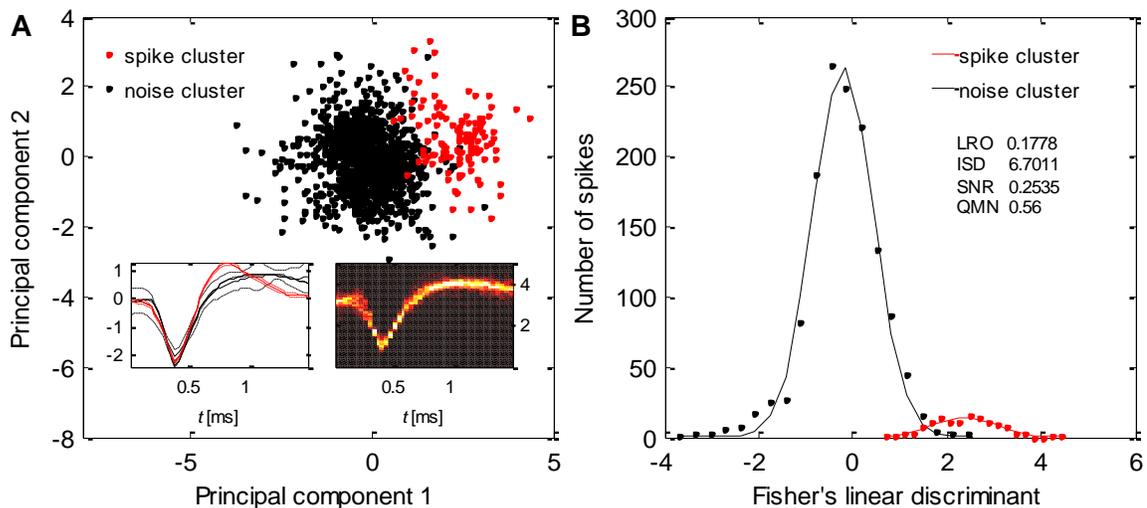
Fig. (5). Real data with intermediate separation. (**A**) Scatter plot of spike waveforms from all clusters using the first two principal components. (**B**) Histograms of the projection of each cluster onto Fisher's linear discriminant between the spike cluster and the noise cluster.
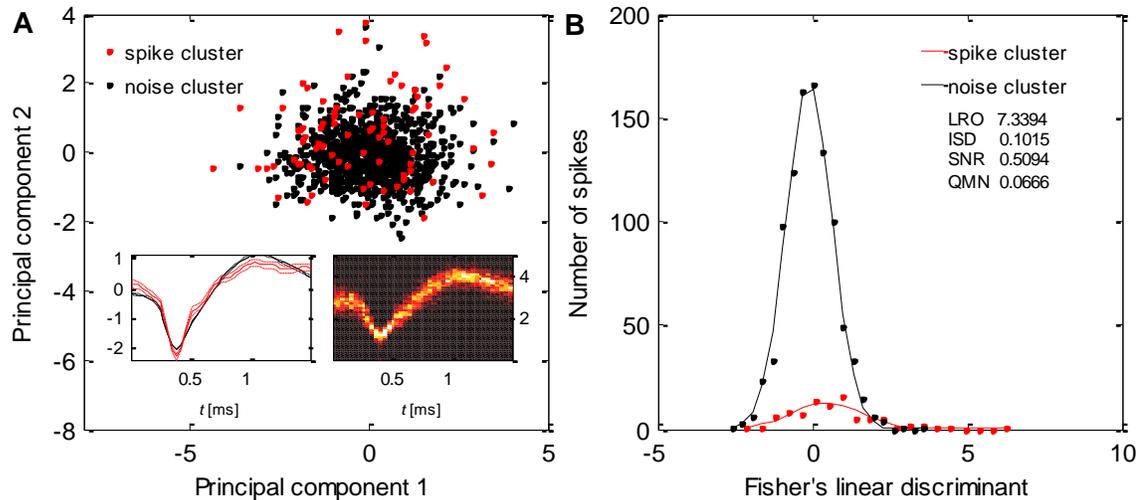
Fig. (6). Real data with poor separation. (**A**) Scatter plot of spike waveforms from all clusters using the first two principal components. (**B**) Histograms of the projection of each cluster onto Fisher's linear discriminant between the spike cluster and the noise cluster.

4. DISCUSSIONS

In this study, we presented a quantitative measure of cluster quality for evaluating the separation of a cluster from other spikes recorded on the same electrode. It is based on the idea of NCA to estimate the probability that every event is correctly classified and further estimate the cluster quality. In this method, the spikes were firstly preprocessed, as being up-sampled, aligned and normalized; Then, they were sorted as the spike cluster and the noise cluster; Thirdly, the wavelet transform and the KS test were used to extract the features of spike waveforms; Lastly, the quality of spike cluster was estimated by the NCA. The best advantage of our method is its range from 0 to 1, which enables easy comparison of units recorded at different times and even by different research groups.

Closer to our proposed method, the isolation score algorithm proposed by Joshua and co-workers [8] applied the idea of NCA to measure the overlap between the noise (non-spike) and spike clusters, and used a nearest-neighbors algorithm to estimate the proportion of false positive and false negative classifications. However, there are two drawbacks of this approach: first, this method was applied directly to the spike waveform which reduces the difference of spike waveform from different neurons, and the performance was much more variable and less reliable; second, it applies the time stamps of spike trains and the entire analog signals as the initial input, largely increases the computation burden. In particular, we showed that the method proposed here overcomes these problems, giving a more robust estimation and a better performance.

In our method, the wavelet-based features were used as the input, which did not require the entire analog signals. The advantage of using the wavelet transform as a feature extractor is that very localized shape differences of the different units can be discerned, and the interference of noise can be greatly reduced. The information about the shape of spikes is distributed in several wavelet coefficients, whereas with PCA or other features [4, 17], most of the information about the spike shapes is captured only by very few variables, such as the first three principal components, which

are not necessarily optimal to represent signals [12]. Since wavelets are a linear transform, therefore all the wavelet coefficients yield nearly the same results as the entire spike shape, the dimensionality reduction achieved by combining wavelets with the KS test may have a broad range of interest [12].

The performance of the proposed method was examined with simulated signals and it was then applied to real data from V1 of rats. With the synthetic signals, the performance evaluation was better than the one obtained using *isolation distance* and L_{ratio}. The NCA is consistent with the ACC which reflects the real proportion of false positive and false negative errors (see Fig. (**2**) and Fig. (**3**)), so this shows that it correlates with both the false positive and false negative errors. For the real data, the NCA can effectively discriminate well- and poorly-separated clusters, and its value is related to the correct identification of spikes (see Figs. (**4-6**)). Although *isolation distance* and L_{ratio} are similar to the method in this way, they do not have a global scale to differentiate between well and poorly isolated units. Following Joshua *et al.* [8], we suggested to exclude units with the NCA value below 0.8 in studies whose conclusions may be influenced by the quality of the recorded units.

Although, the spike detection and sorting technology have been rapid development at present, the quality of single-unit after clustering is not uniform due to the interference of noise with larger amplitude or other neuron spikes. In this paper, we used the idea of NCA to provide a quantitative method for evaluating cluster quality. Unfortunately, the proposed method is invalid for cases in which the number of spike events is far greater than the number of noise events (see Fig. (**2**) and Fig. (**3**)). In future, we will conduct research on this question. In most of the cases however, we found that the number of cluster spikes is lower than the number of the rest.

CONCLUSION

In this paper, a quantitative measure to evaluate spike cluster quality was presented and compared with *isolation distance*, L_{ratio} and so on. The proposed method is based on

the idea of neighborhood components analysis to quantify cluster quality. In this method, we used the wavelet-based features to represent the events for reducing the effects of noise. Its performance was tested with simulated data and real data, and the results proved that it provides better metrics for spike as there are no assumptions about the shape of the distribution of spikes and insensitivity to the number of spikes. Quantitative metrics of cluster quality is a necessary step in interpreting studies based on extracellular recording, so it is suggested that wider use of quantitative measures of cluster quality would likely improve the reproducibility of results across laboratories.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

REFERENCES

[1]   R.Q. Quiroga, and S. Panzeri, "Extracting information from neuronal populations: information theory and decoding approaches," *Nat. Rev. Neurosci.*, vol. 10, no. 3, pp. 173-185, 2009.

[2]   S.P. Kim, J.D. Simeral, L.R. Hochberg, J.P. Donoghue, and M.J. Black, "Neural control of computer cursor velocity by decoding motor cortical spiking activity in humans with tetraplegia," *J. Neural. Eng.*, vol. 5, no. 4, pp. 455-476, 2008.

[3]   S. Suner, M.R. Fellows, C. Vargas-Irwin, G.K. Nakata, and J.P.Donoghue, "Reliability of signals from a chronically implanted, silicon-based electrode array in non-human primate primary motor cortex," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 13, no. 4, pp. 524-541, 2005.

[4]   N. Schmitzer-Torbert, J. Jackson, D. Henze, K. Harris, and A.D. Redish, "Quantitative measures of cluster quality for use in extracellular recordings," *Neuroscience*, vol. 131, no. 1, pp. 1-11, 2005.

[5]   M.S. Lewicki, "A review of methods for spike sorting: the detection and classification of neural action potentials," *Network: Comput. Neural Syst.*, vol. 9, no. 4, pp. R53-R78, 1998.

[6]   C. Pouzat, O. Mazor, and G. Laurent, "Using noise signature to optimize spike sorting and to assess neuronal classification quality," *J. Neurosci. Meth.*, vol. 122, no. 1, pp. 43-57, 2002.

[7]   K.D. Harris, H. Hirase, X. Leinekugel, D.A. Henze, and G. Buzs´aki, "Temporal interaction between single spikes and complex spike bursts in hippocampal pyramidal cells," *Neuron*, vol. 32, no. 1, pp. 141-149, 2001.

[8]   M. Joshua, S. Elias, O. Levine, and H. Bergman, "Quantifying the isolations quality of extracellularly recorded action potentials," *J. Neurosci. Meth.*, vol. 163, no. 2, pp. 267-282, 2007.

[9]   J. Liu and S.C. Chen, "Discriminant common vectors versus neighbourhood components analysis and laplacianfaces: a comparative study in small sample size problem," *Image Vision Comput.*, vol. 24, pp. 249-262, 2006.

[10]   J. Goldberger, S. Roweis, G. Hinton, and R. Salakhutdinov, "Neighbourhood Components Analysis," in *Advances in Neural Information Processing Systems (NIPS'2004)*, Vancouver, Canada: Dec. 2004, pp. 513-520.

[11]   L. Shi, X.Y. Li, and H. Wan, "A predictive model of anesthesia depth based on SVM in the primary visual cortex," *Open Biomed. Eng. J.*, vol. 7, pp. 71-80, 2013.

[12]   R.Q. Quiroga, Z. Nadasdy, and Y. Ben-Shaul, "Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering," *Neural Comput.*, vol. 16, no. 8, pp. 1661-1687, 2004.

[13]   J. Martinez, C. Pedreira, M.J. Ison, and R. Quian Quiroga, "Realistic simulation of extracellular recordings," *J. Neurosci. Meth.*, vol. 184, pp. 285-293, 2009.

[14]   K.J. Paralikar, C.R. Rao, and R.S. Clement, "New approaches to eliminating common-noise artifacts in recordings from intracortical microelectrode arrays: Inter-electrode correlation and virtual referencing," *J. Neurosci. Meth.*, vol. 181, no. 1, pp. 27-35, 2009.

[15]   W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes* in *C*. Cambridge University Press: Cambridge; 1992.

[16]   V.P. Oikonomou, A.T. Tzallas, and D.I. Fotiadis, "A kalman filter based methodology for EEG spike enhancement," *Comput. Methods Programs Biomed.*, vol. 85, no. 2, pp. 101-108, 2007.

[17]   A.B. Wiltschko, G.J. Gage, and J.D. Berke, "Wavelet filtering before spike detection preserves waveform shape and enhances single-unit discrimination," *J. Neurosci. Meth.*, vol. 173, no. 1, pp. 34-40, 2008.