ORIGINAL RESEARCH

Offline Spike Sorting Using Approximate Entropy

Sajjad Farashi[1,2]

Abstract

Analysis of neuronal activities is essential in studying nervous system mechanisms. True interpretation of such mechanisms relies on detecting and sorting neuronal activities, which appear as action potentials or spikes in the recorded neural data. So far, several algorithms have been developed for spike sorting. In this paper, spike sorting was addressed using entropy measures. A method based on a modified version of approximate entropy was proposed for feature extraction, which captured the local variations in spike waveforms as well as global variation to create the feature space. Results showed that the entropy-based feature extraction method created more distinguishing features, which reduces spike sorting errors. The proposed method was capable of separate different spikes in small-scale structures, where the technique such as principal component analysis fails.

Keywords Neural data processing · Spike sorting · Entropy measures · Approximate entropy

Introduction

Neuronal activities in the nervous system control human behavior and functions. Neurons communicate by producing electrical signals called action potential (AP) or spike. To study the neuronal activities, it is necessary to record their produced electrical signals using inserted microelectrodes to the extracellular space [1]. Since the microelectrode tip is surrounded by several neurons, it records the activity of more than one neuron. For allocating each detected spike to its corresponding neuron, the spike sorting procedure is necessary.

The shape of action potential for each neuron is determined by biophysical properties of the neuron and its distance to the microelectrode tip. The shape of the action potential can be considered as a fingerprint for the cell; therefore, it is usually used to distinguish neighbor neurons [2]. So far, several spike sorting algorithms have been proposed. A wavelet-based spike classifier was introduced in [3] according to the time–frequency wavelet spectrum

analysis. The main idea of that method was based on the selection of the limited numbers of wavelet coefficients that distinguish waveforms. For such purpose, the wavelet coefficients with bimodal or multimodal distribution among all action potentials were selected manually [3]. Although the wavelet-based methods are potent in spike sorting procedures, such methods are susceptible to the selection of basis function. In fact, selection of the wavelet basis function is an important issue in the wavelet domain and needs a priori knowledge about spike shapes which are not accessible in the real experiments. Clustering with mixtures of multivariate t-distributions using log-likelihood maximization and expectation–maximization algorithm for parameter estimation was proposed in [4]. Compared with the traditional Gaussian model, this t-distribution mixture model decreased the effect of outlier spike waveforms in the clustering procedure [4]. The main weakness of methods based on neural networks is their necessity for learning procedure that needs a priori knowledge about the data, which usually is not accessible in the neural data processing. A combination of statistical analysis and neural networks is another widely used approach for sorting purposes [5]. Furthermore, self-organizing map (SOM) accompanied by independent component analysis (ICA) was another method that was proposed for clustering purposes [6].

Authors in [7] used the related trajectories of spike waveforms in the phase space as a tool for spike sorting. Mean

✉ Sajjad Farashi
  sajjad_farashi@yahoo.com

1 Autism Spectrum Disorders Research Center, Hamadan
   University of Medical Sciences, Hamadan, Iran

2 Neurophysiology Research Center, Hamadan University
   of Medical Sciences, Hamadan, Iran

of trajectories in the phase space were used for template construction, and the minimal distance criterion was utilized for spike classification. The selection of distance measure was important in this method and affected the results. In other studies, entropy measure was used for spike sorting in a way that the entropy cost function was used for optimal wavelet basis function selection and for selecting the limited number of wavelet coefficients [8, 9]. Pavlov et al. proposed a method of representative waveforms (rw) based on the averaging spikes related to points around PCA cloud centers. The wavelet coefficients in some decomposition levels were selected, which maximize the distance between rws in wavelet space [10]. In other words, for each spike, the coefficients related to some selected scale and translation levels were selected as spike features. To take into account the shape of the action potential, each coefficient was selected in one-half of the spike duration [10]. In another wavelet-based method, the authors used those discrete wavelet coefficients that their distribution among all spikes had a more significant deviation from a normal distribution. For measuring multimodality, the Kolmogorov–Smirnov test based on cumulative distribution function (CDF) was used [11]. Four-level decomposition was carried out and two wavelet coefficients in which their distribution among all spikes had a maximum distance from normality were chosen as new features for each spike [11]. For each dataset, the wavelet basis function that was more correlated with spike templates was used as basis function to get a sparser wavelet space. Autoregressive modeling of action potentials is another way for spike sorting [12]. Furthermore, the power spectrum density function of biological data might contain useful information for activity discrimination [13].

In the present work, an offline spike sorting methodology was proposed. For feature extraction purposes, a method based on approximate entropy (ApEn) was proposed. To correctly estimate the entropy for smooth and very short length waveforms like APs, the main ApEn algorithm proposed by Pincus [14] was modified slightly. The ApEn-based proposed method used the local variation of spike shapes as well as global variations. The results showed that ApEn based feature extraction obtained better performance than PCA and wavelet-based methods for spike sorting.

Material and Methods

Data Recording Procedure

A single tungsten microelectrode with impedance about 1MΩ was used to record neuronal activity from a cockroach restrained firmly on a plastic disk. This plastic disk was located in a faraday cage to reduce the effects of electromagnetic interference. After an initial amplification using a preamplifier, analog neural data were amplified by the main amplifier with a gain of 2000 and band-pass filtered in the range of 0.3–3 kHz. The analog data were digitized using a National Instrument ADC card (30 K samples/s). All procedures, including data acquisition and analyses, were controlled by a user-written program in the Labview environment (Version 8.6, National Instruments, USA). Recorded data were up-sampled (up-sampling factor = 2) to increase the number of data points and reduce the alignment error [15]. Using an automatic amplitude thresholding strategy, spikes were extracted. The labeling process for extracted spikes was performed by three experts. In the first stage, two experts (1 neurologist and 1 neuroscientist) labeled the extracted spikes; however, to resolve any conflicts the third person (a neuroscientist) finalized the labeling procedure. There was 93% agreement between the first two experts.

## Feature Extraction Procedure

Feature Extraction Method

In this paper, the modified ApEn-based method was proposed for feature extraction. Since neurons produce spikes with stereotyped shape, the spike waveform is considered a useful tool to discriminate spikes [2, 16, 17]. ApEn is a measure of complexity or uncertainty in a time series [14]. The most important features of ApEn are resistance to short transient interferences, robustness against noise and consistency with short length data. These features make ApEn an interesting tool for spike processing because spikes are very short-length data that usually are affected by short and strong transient noises induced by electronic devices. In this paper, ApEn was used as a measure of variability in spike time-series which was affected by spike shape. More fluctuations in the spike waveform increased the value of ApEn. A fast algorithm for ApEn calculation was proposed in [14] that used the variation between segmented patches of a time-series. In this paper, that algorithm was revised to be consistent with a very short-length time series like spikes.

The proposed ApEn-based algorithm for feature extraction is displayed in Fig. 1 and was implemented as follows:

1. Suppose N-sample spike. Select ith L-sample ($L \leq N$) segment of spike by starting point i and jth segment with the same length and starting point j where $j = \{1, \ldots, N-L+1\}$. Compare the two segments as step 2 explains.
2. The absolute point-wise difference between ith segment and jth segment of a spike is computed. The absolute differences obtain the elements of vector $\boldsymbol{D}$, the abbreviation for difference vector.

Fig. 1 Block diagram of the ApEn-based feature extraction method. Based on the proposed method the variation between different segments of each spike was returned in a matrix called *C*. when such a matrix was created for all spikes, two matrix elements were selected where the distribution of entries among all matrices was more deviated from a normal distribution. In this regard, (m, n) and (m', n') elements among all matrices had multi-modal distribution (indicated by arrows) while (m^*, n^*) had mono-modal distribution. This indicated the entries had no ability to distinguish spikes originated from different neurons. The selected features for each spike were multiplied by the calculated ApEn of that spike to create the final feature
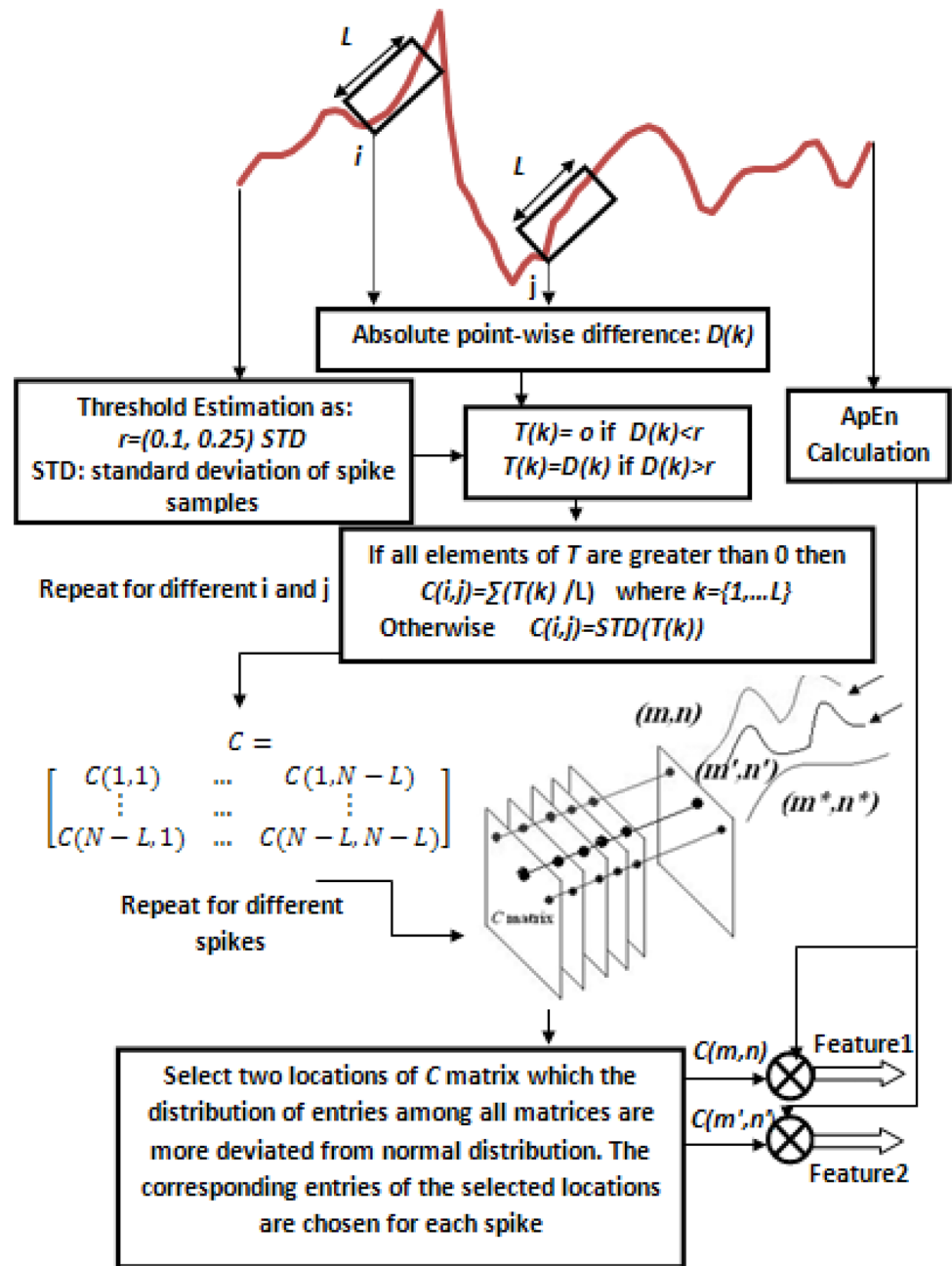
3. A new vector called T is created from D based on (1) as follows:

$$T(k) = \begin{cases} D(k), D(k) > r \\ 0, D(k) \leq r \end{cases}, \tag{1}$$

where in (1), $(k = 1, \ldots, L)$ and r is a predefined positive threshold for decreasing the noise effect in the calculation of ApEn. Greater r causes the noise-induced variations between segments to be neglected. Such thresholding enables ApEn to be robust against noise. T is an abbreviation for the thresholded vector. As suggested by

Pincus [14], r can be taken as $(0.1, 0.25) \times \mathrm{SD}_X$ where SD_X is the standard deviation of the original spike waveform (X).

4. If all T elements are non-zero, the mean value of vector T is considered in the (i,j)th entry of a matrix called C; otherwise, the standard deviation of values is considered (see Fig. 1). Note that the mean value shows the average dissimilarity between two patches of a spike. Also, as the threshold r is usually a small value, in the case of existence of zero elements in the vector T, the standard deviation will be a small value and hence indicates that

there is no significant variation between ith and jth segments. In this regard, each entry of C matrix shows the level of variation between two segments of a spike. C is a symmetric matrix with zero main diagonal which is produced for each spike. As extracellular recorded signals are inevitably corrupted by noise and far-field action potentials which induce relatively small variations in the spike waveform, considering mean value reduces the effect of such small variations in entropy calculation.

5. Averaging ith row of C, which gives c_i exhibits the relative variation of ith segment of spike relative to whole spike duration. More variation of a segment leads to greater value for c_i. For example, if time-series is a sequence of equable values, the relative variation between segments is zero which leads to zero value for c_i. For each spike the level of variation is computed based on (2) as follows:

$$\varnothing_L = \frac{1}{N - L + 1} \sum_{i=1}^{N-L+1} C_i. \tag{2}$$

It should be noted that in the original ApEn algorithm proposed by Pincus [14], if all elements of $T(k)$ be greater than r then C entry is replaced by 1; otherwise, entry is replaced by 0 and this creates a binary matrix. In this way, the c_i is the total number of signal patches that are close to one intended patch, but in the modified ApEn algorithm, c_i is an estimation of the variation of each segment in the time-series. For a short length and smooth waveform like a spike, this calculates the variation of waveform more precisely. If the length of segments, L, which is called the dimension of calculation, is increased to $L + 1$, and the above steps are repeated, another measure is obtained (φ_{L+1}) which shows the level of variation between signal segments with length of $L + 1$.

6. Finally, ApEn is calculated as follows (3):

$$\text{ApEn} = \ln\varphi_L / \varphi_{L+1} \tag{3}$$

Note that for complex signals, which contain a higher level of variation, changing the dimension causes the value of φ to be changed dramatically. In comparison, for the lower level of variations, this change is negligible and consequently, ApEn tends to zero.

As ApEn is a global measure of time series variation, it is probable that two different waveforms with different local variations have an equal value of ApEn. For differentiating spikes with differences in small-scale structures, some C matrix entries with local discrimination capability are considered in feature extraction, as is explained in step 7.

7. In the above steps, for each spike, a distinct C matrix is produced. Distribution of (m,n) entries among all C matrices are calculated using histogram-based probability density function (pdf) estimation and finally a limited numbers of entries which their estimated cumulative distribution function ($F(x)$) are more deviated from Gaussian distribution with the same mean and variance ($G(x)$) are selected. Deviation from normality is quantified by $\max|F(x) - G(x)|$ [11]. In Fig. 1 the distribution of entries in three locations among all C matrices are displayed. If there are different spike templates in the dataset, most-discriminative entries among all C matrices are those that their distribution is multimodal. In this regard, in Fig. 1 the distribution of entries in (m^*, n^*) location among all matrices is mono-modal and, therefore, corresponding entries cannot discriminate different spikes. However, the distribution of entries in (m, n) location is multimodal. This indicates stronger discrimination.

8. For constructing a 2D feature space, two entries, as explained in step 7, are selected for each spike and multiplied by ApEn of that spike. Such multiplication creates two features for each spike (see Fig. 1). Note that both selected $C(m, n)$s and ApEn are dependent on spike shape which the former considers the local variability and the latter considers the global variability of spike. The global and local variations of spike shape can be used for discriminating spikes that originated from different neurons.

Selection of Parameter *L* for Feature Extraction

The most important parameter for the proposed ApEn-based feature extraction method is dimension L. ApEn is proportional to the ratio of close signal segments which remain close by increasing L. For time-series which contains high level of variations, as proposed in [14], it is better to choose L to be 2 or 3 to capture newly presented patterns of variation; however, for highly smooth and short length data like spikes, it is better to assign a larger value for L. It should be noted that the patterns of spikes generally consist of a rising segment which finally reaches the spike dominant peak and a segment followed by the peak. The main variation in such a waveform is around the peak location. To capture such variations during increasing L to $L + 1$ in ApEn calculation, it is proposed to set L as the spike peak duration. Due to the special morphology of a spike which consists of four phases (falling, rising, hyperpolarization and resting state), the spike peak duration can be considered as one-quarter of spike duration. Since the refractory period of APs is rarely greater than 1 ms [18], spike length can be calculated based on the sampling rate

of analog-to-digital conversion. For our datasets (dataset 1 and dataset 2) which were used for comparison purposes, this value on average was 16 samples. Such selection was evaluated in "Sensitivity of the proposed method to *L* and threshold applied to Pearson's correlation".

Sorting Procedure

Mapping each spike into a *2D* feature space explained in "Feature extraction method", creates some clouds in the feature space. The optimal situation is that each cluster belongs to a distinct neuron. These clouds can be overlapped or well separated. The powerful features are those that produce well-separated and more compact clouds in the feature space. Averaging spikes related to points in the small neighborhood of each cluster center created a template for that cluster as a representative waveform. The centers were found manually. It should be noted that the feature space was constructed for building representative templates. In this paper, Pearson's correlation as a simple distance metric was used to classify spikes. The correlation coefficient defined by (4) was calculated between each spike and all constructed templates. In statistics, Pearson's product–moment correlation coefficient is a measure of the correlation (linear dependence) between two waveforms like *X* and *Y*, giving a value between -1 and $+1$. It is widely used in the sciences as a measure of the strength of linear dependence between waveforms [19]. Suppose *X* as spike and *Y* as the representative template of one cluster, with average of signal samples $\overline{X}$ and $\overline{Y}$, respectively, and the standard deviation of their samples, S_X and S_Y, respectively. Pearson's correlation between spike and each template is computed by (4) as follows:

$$r = \sum_{i=1}^{N}(X_i - \overline{X})(Y_i - \overline{Y})/(n-1)S_X S_Y, \tag{4}$$

where *N* is the number of spike samples or representative template samples.

Finally, the spike was allocated to a cluster, which was more correlated to its corresponding representative template. Due to the false alarm error that usually occurs in the majority of spike detection algorithms, a predefined threshold was applied on the correlation coefficient to remove non-spike events, which had a low correlation with constructed representative templates.

The block diagram of the proposed method is depicted in Fig. 1.

Spike Sorting Performance Evaluation

In this paper, the performance of the sorting procedure was quantified by the percentage of misclassified spikes and the percentage of unclassified spikes for each algorithm. For performance evaluation, also these two types of errors were combined to obtain a unique error-index [7] (Eq. 5) as follows:

$$\text{error index} =$$

$$\sqrt{\sum_{i=1}^{M}(\text{unclassified error}(\%))^2 + \sum_{i=1}^{M}(\text{misclassified error}(\%))^2}, \tag{5}$$

where in (5), *M* is the number of clusters. This error-index returns the aggregation of errors in all classes. Unclassified spikes were referred to spikes which their normalized correlation with all constructed templates was smaller than a pre-defined threshold and hence not assigned to any clusters. The misclassified error referred to spikes which belong to one neuron but are allocated to another one.

In the field of neural spike sorting, there are two common sets of spike datasets. The spikes in a dataset can have different morphology or can be different in small-scale structures. In this paper, two different spike datasets were used for comparison. Dataset 1 consisted of two templates, which were different in small-scale structures and dataset 2 consisted of three spike templates which were different in shape. Such datasets were extracted from the real recorded data. Templates were displayed in Fig. 2.

Fig. 2 Spike templates for **a** dataset 1 which were different in small-scale structures and **b** templates for dataset 2 which were different in shapes. Two datasets were extracted from real recorded neural data
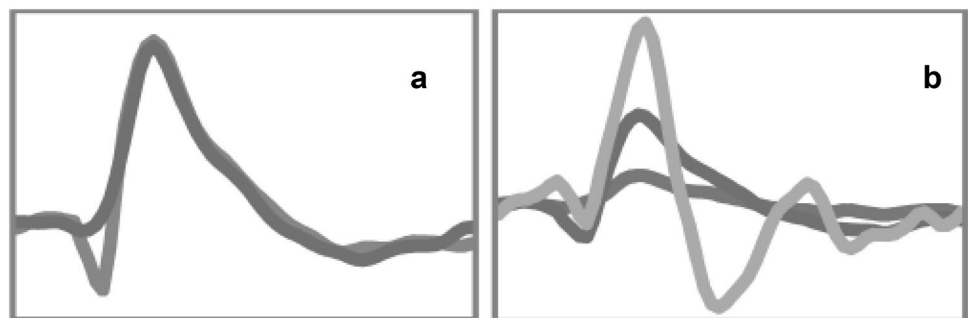
Results and Discussion

Apen-Based Feature Extraction Performance

The proposed ApEn-based method for feature extraction was compared with other methods, including PCA, the combination of PCA and wavelet which was named wavelet shape accounting classifier (WSAC) [10] and the combination of wavelet-based method and Kolmogorov–Smirnov (KS) which here was called WKS [11]. In the proposed ApEn-based algorithm, parameter r was set to 0.25SD_{SD} where SD_{SD} was the standard deviation of spike samples and L was selected as one-quarter of spike length.

After mapping each spike dataset (dataset 1 and dataset 2) to the feature space by each method, the centers of clouds in the feature space were found manually. By averaging spikes related to points in the small neighborhood of the cloud centers, templates for the sorting procedure were created. Each template was considered as the cluster marker. For sorting, Pearson's correlation was applied as the distance metric where the correlation between each spike and constructed templates was computed. Finally, the spike was allocated to the cluster which was more correlated with its marker (template). Due to the inevitable false alarm errors in most spike detection algorithms, it was possible to find waveforms in dataset that were not spike events but detected falsely. For eliminating such events in the sorting step, a predefined threshold was applied to Pearson's correlation coefficient, where if the correlation of spike with all templates was lower than the threshold, that spike was considered as the non-spike event and not assigned to classes. In this paper, the threshold was considered to be 0.5 for all methods (see supporting material, appendix B for the reason of selecting threshold level of 0.5).

Figure 3 shows the mapping of dataset 1 to the feature space using different methods. Clearly, PCA failed to separate spikes with differences in small-scale structures. This is the major weakness of PCA in spike sorting. For solving such difficulties, authors in [10] proposed WSAC method even though visually it seemed inappropriate for dataset 1. It could be seen from Fig. 3 that WKS obtained the well-separated clusters as ApEn-based method.

Table 1 shows that although WKS visually obtained clusters with greater distance than the ApEn-based method, it

Table 1 Comparison between ApEn-based method and other methodologies for clustering of spikes different in small-scale structures

Method	Unclassified error (%)	Misclassified error (%)
ApEn	2.6	2.2
WKS	3.1	2.5
PCA	3.6	7.6
WSAC	5.6	10.9

Fig. 3 Dataset 1 (different in small-scale structures) was mapped into the feature space by other methods. The methods were ApEn-based method (Entropy), Principal Component Analysis (PCA), wavelet shape accounting classifier (WSAC) [10] and WKS (combination of wavelet and Kolmogorov–Smirnov criterion) [11]
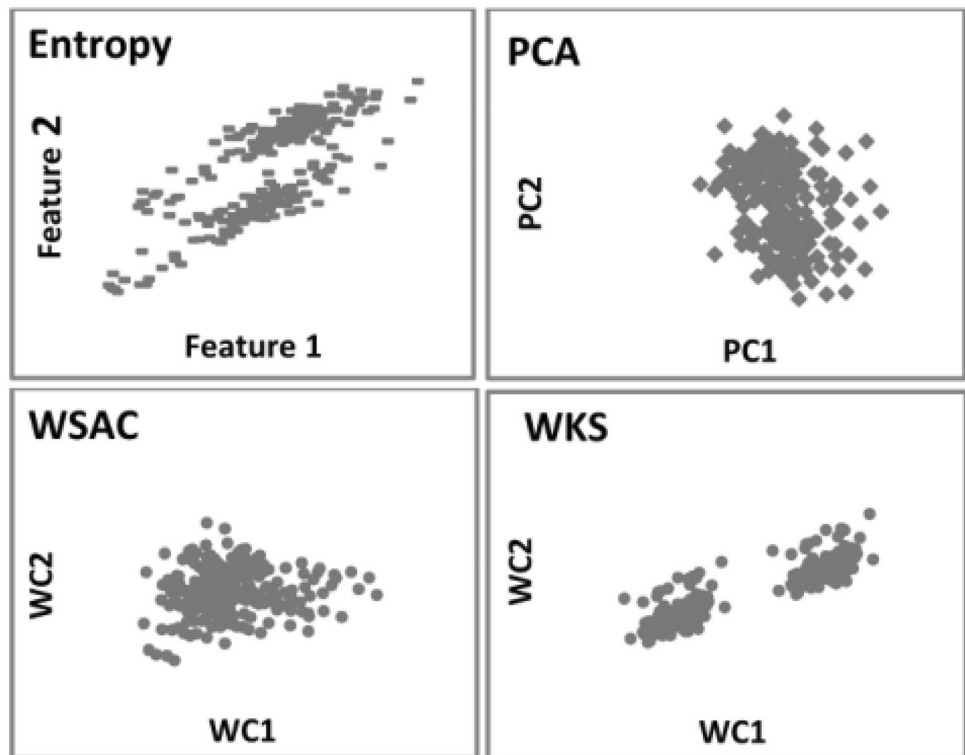
returned more misclassified and unclassified spikes. The reason was that in WKS each spike was transformed into the wavelet space and each coefficient represented some energy content of that spike. Any unwanted fluctuation in the spike waveform changed the energy distribution. This could move the related points in the feature space from one cluster to another. This is the case for especially datasets 1, in which there were differences in small-scale structures, and in this manner, the clustering error increased. In the case of the ApEn-based method, parameter r in ApEn calculation reduced the effect of such unwanted fluctuations.

For dataset 2 where there were 1580 spikes of three different spike templates, the percentage of unclassified and misclassified spikes using different feature extraction algorithms followed by Pearson's correlation is reported in Table 2. Unlike dataset 1, for dataset 2 it was no longer easy to evaluate the performance of methods based on reported unclassified and misclassified percentages because even though ApEn had lower misclassified errors but it was faced with higher unclassified error than PCA. Therefore, another criterion which was a combination of misclassified and unclassified errors was used for comparison purposes, which was called error-index (Eq. 5).

Table 2 showed the lower unclassified percentage for PCA because in PCA, due to mean-centering, the spike-related

points in the feature space were closer to the cluster centers. This caused more reliable representative templates that obtained a lower number of unclassified spikes. Table 2 showed that the ApEn-based algorithm for feature extraction obtained the lower number of misclassified spikes in comparison with PCA and WKS because instead of using some limited number of wavelet coefficients (in WKS) or a limited number of scores (in PCA), global and local variations of spikes were used for feature extraction. Also, due to parameter r in ApEn calculation, denoising was carried out which made ApEn-based feature extraction robust against correlated or uncorrelated noises. Figure 4 indicated the sorting results for dataset 2 where empty circles showed unclassified spikes. Due to the large error for WSAC method (SEE Table 2), this method was excluded for further analyses.

Furthermore, the sensitivity of ApEn-based feature extraction to noise was accessed. For this aim, another spike dataset from [11] was used which contained 1000 spikes of two different templates. Figure 5 shows the result of the clustering for ApEn-based, PCA and WKS feature extraction methods. In each implementation, Gaussian noise with different strength and variance was added to dataset; then features were extracted and representative waveforms were constructed by averaging related spikes around cluster centers, and finally, the classification was carried out by allocating each spike to the more similar representative template. Gaussian noise was added because it is usually presented in neural data due to the noise induced by electronic devices or thermal perturbation. The first column in Fig. 5 was related to noisy templates; other columns were related to clustering results for WKS, PCA and ApEn, respectively. Note that the noisy templates were chosen from the dataset directly and differed from representative waveforms. The rows from top to down were related to different noise strengths. For WKS, symlet 6 mother wavelet was chosen because it was more

Table 2 Comparison between feature extraction methods for sorting dataset 2 with different Spike templates

Method	Unclassified error (%)	Misclassified error (%)	Error index (%)
ApEn	2.1	0.9	2.284
PCA	1.4	1.9	2.36
WKS	4.5	3.9	5.954
WSAC	7	4.5	8.32
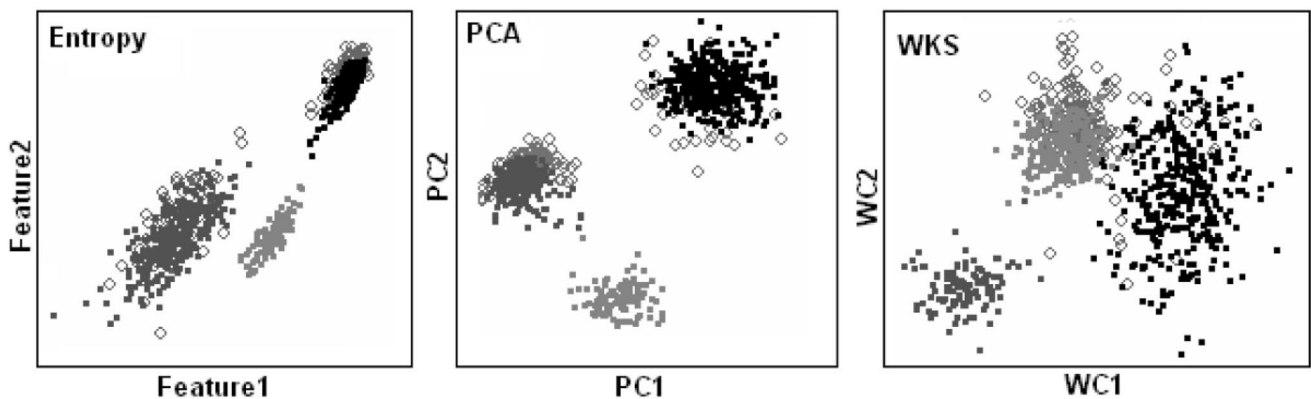
Fig. 4 Sorting result for dataset 2 which consisted of three different spike templates. Empty circles indicate unclassified spikes. The methods were ApEn-based (Entropy), Principal Component Analysis (PCA), and WKS (combination of wavelet and Kolmogorov–Smirnov criterion)
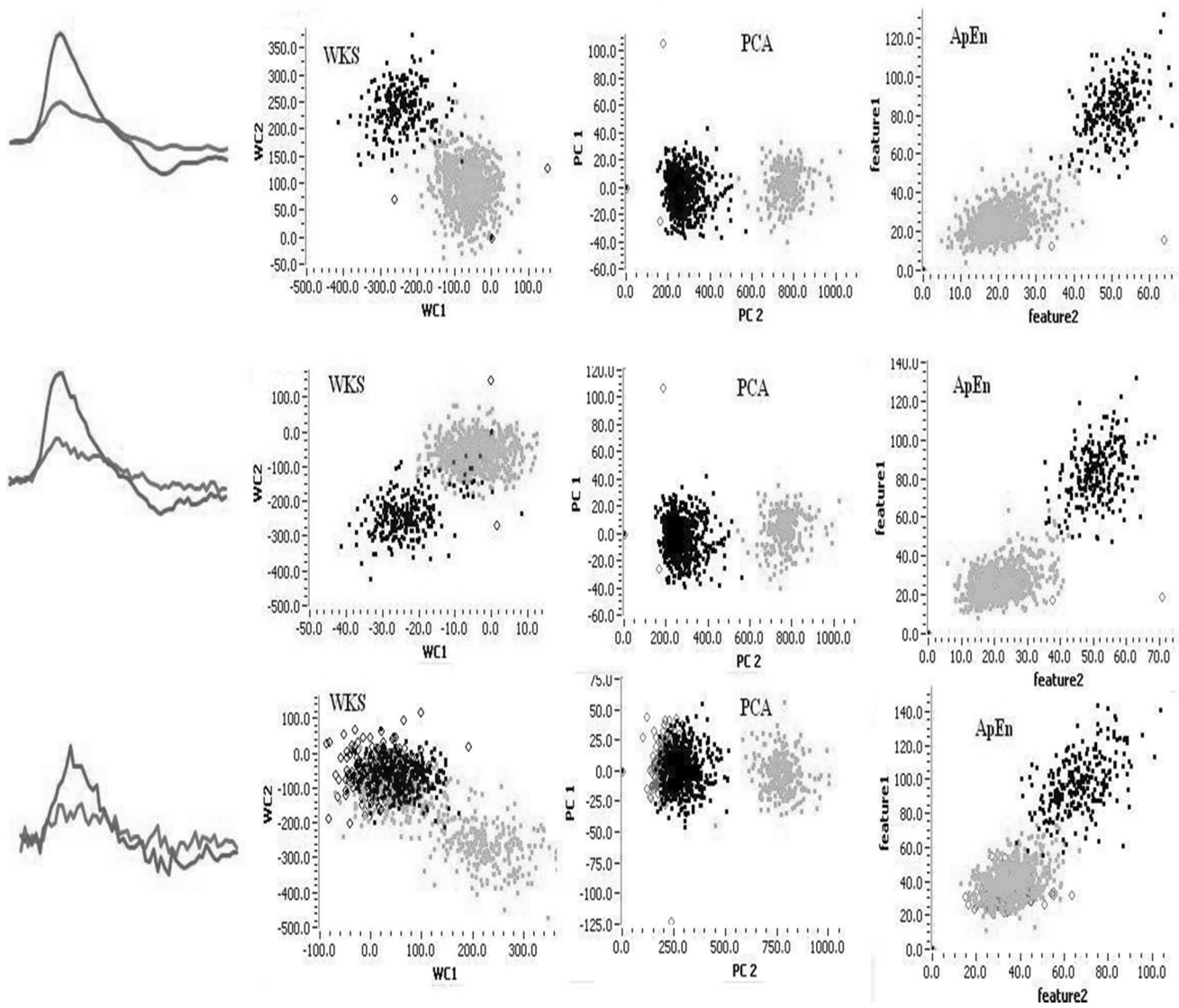
Fig. 5 Robustness of spike sorting to noise. The dataset consisted of two template waveforms. Different levels of uncorrelated noise were added to the dataset and clustering was performed for ApEn-based, PCA and WKS methods. The first column was related to the noisy waveform templates with different strengths of noise from top to down. Other columns were related to the clustering of spikes by various methods

similar to the waveforms of the selected dataset. Results in Fig. 5 showed that WKS had the lowest robustness against noise as the number of misclassified spikes increased by adding stronger noise which were 5, 27 and 202 spikes, respectively. The location of the cloud centers in WKS feature space changed as noise was added because the wavelet coefficients were affected by noise strength. These results showed that PCA had the highest robustness against noise and the number of misclassified spikes by increasing noise strength was 0, 0 and 1, respectively. The number of misclassified spikes in ApEn-based feature extraction was 1, 6 and 23 as noise strength was increased. In this regard, ApEn-based feature extraction made clustering more

robust against noise than WKS, but PCA outperformed ApEn-based method. The reason was that such uncorrelated noise was projected on components other than first two components, where these components were not used in PCA-based feature extraction and only the first two principal components were considered in this paper. In the case of ApEn-based feature, extraction parameter r in ApEn calculation caused noise effect to be reduced; therefore, its robustness was higher than WKS method. Since the added noise was uncorrelated, noise reduced the similarity between constructed templates and spikes which increased the number of unclassified spikes.

A Note on ApEn-Based Feature Extraction

As mentioned in the material and methods, for each spike the extracted features were obtained by multiplication of two selected elements of C matrix and the corresponding ApEn of that spike. The question is that if multiplying the ApEn (as a measure of overall variability of spike waveform) has a considerable effect on the discriminative power of extracted features? To address this question, different spike datasets with the templates that are depicted in Fig. 2 were mapped into the feature space with and without considering multiplication of ApEn in the feature generation step; then the inter-cluster and intra-cluster distances [20] were calculated. The inter-cluster distance was computed based on the distance between cluster centers which were selected manually. Higher inter-cluster distance indicates that the clusters were more separated. The intra-cluster distance was computed as the average distances between cluster members and the corresponding cluster's center which was a measure of cluster compactness. Smaller average intra-cluster distance is related to the more compact clusters. In this regard, Davies–Bouldin index (DB) [20] was used to assess the quality of the clustering algorithm which is defined as (6) follows:

$$DB = \frac{1}{n} \sum_{i=1}^{M} \max_{i \neq j} \left(\frac{\sigma_i + \sigma_j}{d(O_i, O_j)} \right), \tag{6}$$

where M is the number of clusters, O_i is the centroid of cluster i and $d(O_i, O_j)$ is the distance between centroids O_i C$_i$ and O_j C$_j$. Also σ_i is the average distance of all elements in the cluster i to centroid O_i. In comparison between two algorithms that produce clusters, the one with the smallest DB index is more favorable. The DB index was calculated for both datasets 1 and 2. It is also calculated for conditions of including or excluding the multiplication of ApEn. The results were reported in Table 3. These results showed that by considering ApEn of each spike in the feature generation procedure, the value of DB index was found lower than the case of neglecting ApEn in feature generation. Such a result was expected because considering the ApEn of each spike as a measure of global spike variation in the feature generation accompanied by local variation in the spike waveform made generated features more sensitive to the spike shape which was a specific characteristic of each neuron.

Table 3 The effect of considering ApEn in the feature selection. The numbers in this table were related to Davies–Bouldin index

| Dataset	Considering ApEn	Without ApEn
Dataset 1	1.9315	2.2042
Dataset 2	3.1458	3.2442

Sensitivity of the Proposed Method to L and Threshold Applied to Pearson's Correlation

To capture the level of variation in each spike waveform, L-sample segments of each spike were compared with each other. Here, the sensitivity of the sorting procedure to different L values was tested for datasets 1 and 2. The sensitivity was defined as TP/(TP + FN), where TP was true positive(true classified) and FN (false negative). In order to evaluate the sensitivity of the algorithm to L and find its optimal value, bootstrap selection procedure was used and one-quarter of spikes in each dataset were selected for the test. The results for both datasets are depicted in Fig. 6, where the sensitivity of spike sorting to different values of L for both datasets is shown. For dataset 1, when L was large, as templates were similar and the average value of differences between spike segments were replaced in entries of C matrix, the detail differences were eliminated by averaging and, therefore, the entries of the same elements in C matrices were very close and this increased FN. When L was near to the duration that the main difference between templates that occurred (L between 7 and 18), the most sensitivity (the least FN) was achieved because in this case, the difference between selected entries among all C matrices increased which resulted in more separated clusters in the feature space. For dataset 2, where spike templates were different, the sensitivity had lower dependence on L because of the larger differences between different spikes. In this case, higher sensitivity was achieved when L was chosen near the spike peak duration.

Furthermore, the sensitivity of the proposed method to the threshold applied to the Pearson's correlation during sorting procedure is shown in Fig S2 (see supplementary material). For this analysis, receiver operating characteristic curve (ROC) was used. The sensitivity of other methodologies that were used for comparison (i.e. PCA, WKS and WSAC) was also evaluated. In ROC analysis, true positive
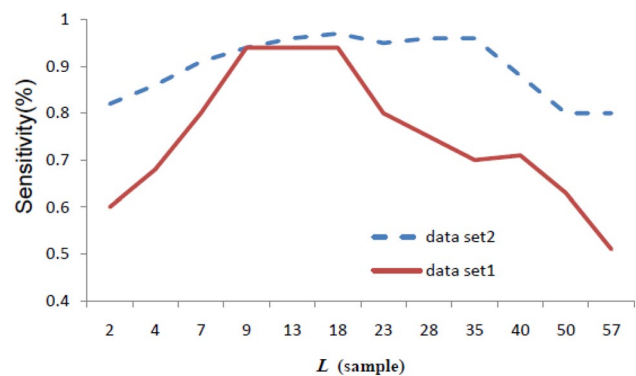
Fig. 6 Sensitivity of the proposed method to parameter L
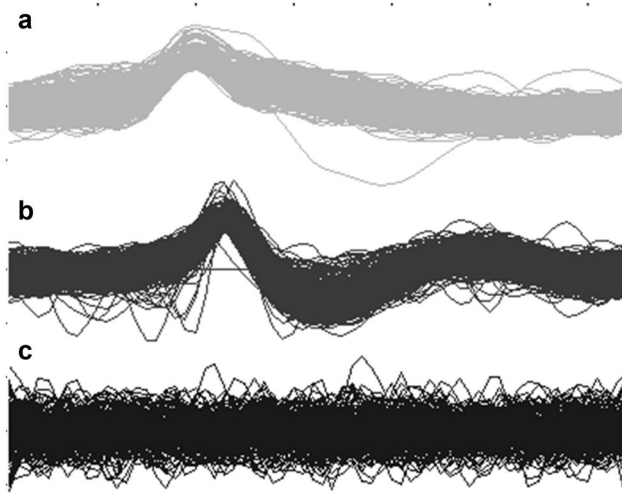
Fig. 7 Results of the proposed entropy-based detection and feature extraction algorithm for a real neural recording. **A** Spike train 1(1250 spikes), **B** spike train 2 (984 spikes), **C** background noise as unclassified waveforms (52 waveforms) which had low correlation with representative templates

Table 4 The result of clustering based on ApEn feature extraction and other methods for spikes detected from real neural data

Error	Entropy	PCA	WKS	WSAC
%FN	8.1	7.4	8.2	14.6
%FP	1.6	3.2	4.2	8.4
Sum of errors	9.7	10.6	12.4	23

rate (TPR) vs. false positive rate (FPR) for different values of threshold levels applied to the Pearson's correlation levels was calculated. With higher threshold levels, spike was allocated to the more similar pattern; therefore, true positive value will be enhanced while the number of falsely classified spikes (FP) was reduced. Since L is the most important parameter in the proposed methodology, to check if the sensitivity analysis obtained a generalized result, another dataset was used and the above-mentioned sensitivity analysis was performed again. The result for this analysis is shown in Fig S2 (see supplementary material).

Result for Another Real Recorded Data

In the previous sections, the ApEn-based sorting method was tested for real action potentials extracted from the cockroach recording and action potential from reference [11]. In this section, the proposed sorting algorithm was applied to another real spike dataset, which was recorded from a cockroach. Figure 7 depicts sorted spikes using

ApEn feature extraction accompanied by Pearson's correlation where spikes were sorted into two distinct classes (A and B). The waveforms labeled by C in Fig. 7 were falsely detected spikes which had a low value of correlation with constructed templates. The templates were constructed by averaging spikes related to points in the small neighborhood of cluster centers. The results of clustering for extracted spikes by different feature extraction methods were summarized in Table 4. As there was no a priori information about the number of clusters or spike templates, the real data and result of clustering were investigated by some expert persons to quantify the performance of feature extraction methods. The experts found the presence of two different templates and approximately 2200 spike waveforms in the recorded data. Experts consisted of three persons (2 neuroscientists and 1 neurologist). Results in Table 4 were based on the average values reported by these experts. These results showed that ApEn feature extraction obtained a lower percentage of false-negative and false-positive errors for real data. This was because in real data recording, the activity of neurons in far-field and noise sources like electrode displacement or motions superimposed on the waveform of the spike of the intended neurons. This changed the waveform slightly which affected the energy distribution in wavelet coefficients or PC scores. Although, in ApEn method, r parameter made the calculation insensitive to the noise which this decreased classification errors (see Table 4).

Conclusion

In this paper, a method based on entropy measure was proposed for offline spike sorting. For this purpose, the variation of action potential shape was considered and a method based on approximate entropy (ApEn) was proposed for measuring the variability of spike waveforms. For addressing the variability in the smooth and short-length spike waveform, the ApEn proposed in [14] was modified to accommodate the spike event. Focusing on spike variations in ApEn-based feature extraction caused spikes with differences in small-scale structures to be separated as well as spikes with different templates. The majority of feature extraction methods like PCA failed to separate spikes with detail differences. Against the wavelet-based method, the ApEn-based method was more robust against noise. Results showed that selected features based on ApEn, reduced the overall misclassified and unclassified spikes and consequently the overall classification error.

Declarations

Conflict of interest Author has nothing to declare.

Ethics approval The study was approved by the Ethical Council of Research of the Faculty of Medicine, Tehran University of Medical Sciences, Tehran, Iran (Ref. 289218720).

Consent to participate Not applicable.

Consent for publication Not applicable.

References

1. Eliades SJ, Wang X. Chronic multi-electrode neural recording in free-roaming monkeys. J Neurosci Methods. 2008;172:201–14.
2. Vargas-Irwin C, Donoghue JP. Automated spike sorting using density grid contour clustering and subtractive waveform decomposition. J Neurosci Methods. 2007;164:1–18.
3. Letelier JC, Weber PP. Spike sorting based on discrete wavelet transform coefficients. J Neurosci Methods. 2000;101:93–106.
4. Shoham S, Fellows MR, Normann RA. Robust, automatic spike sorting using mixtures of multivariate t-distributions. J Neurosci Methods. 2003;127:111–22.
5. Kyung Hwan K, Sung June K. Neural spike sorting under nearly 0-dB signal-to-noise ratio using nonlinear energy operator and artificial neural-network classifier. IEEE Trans Biomed Eng. 2000;47:1406–11.
6. Hermle T, Schwarz C, Bogdan M. Employing ICA and SOM for spike sorting of multielectrode recordings from CNS. J Physiol Paris. 2004;98:349–56.
7. Aksenova TI, Chibirova OK, Dryga OA, Tetko IV, Benabid AL, Villa AE. An unsupervised automatic method for sorting neuronal spike waveforms in awake and freely moving animals. Methods. 2003;30:178–87.
8. Hulata E, Segev R, Ben-Jacob E. A method for spike sorting and detection based on wavelet packets and Shannon's mutual information. J Neurosci Methods. 2002;117:1–12.
9. Farashi S, Abolhassani MD, Salimpour Y, Alirezaie J. Combination of PCA and undecimated wavelet transform for neural data processing. Ann Int Conf IEEE Eng Med Biol. 2010;2010:6666–9.
10. Pavlov A, Makarov V, Makarova I, Panetsos F. Sorting of neural spikes: when wavelet based methods outperform principal component analysis. Nat Comput. 2007;6:269–81.
11. Quiroga RQ, Nadasdy Z, Ben-Shaul Y. Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. Neural Comput. 2004;16:1661–87.
12. Farashi S. Spike sorting method using exponential autoregressive modeling of action potentials. World Acad Sci Eng Technol Int J Med Health Biomed Bioeng Pharm Eng. 2015;8:864–70.
13. Golnar-Nik P, Farashi S, Safari M-S. The application of EEG power for the prediction and interpretation of consumer decision-making: a neuromarketing study. Physiol Behav. 2019;207:90–8.
14. Pincus SM. Approximate entropy as a measure of system complexity. Proc Natl Acad Sci USA. 1991;88:2297–301.
15. Jung HK, Choi JH, Kim T. Solving alignment problems in neural spike sorting using frequency domain PCA. Neurocomputing. 2006;69:975–8.
16. Pouzat C, Mazor O, Laurent G. Using noise signature to optimize spike-sorting and to assess neuronal classification quality. J Neurosci Methods. 2002;122:43–57.
17. Franke F, Natora M, Boucsein C, Munk MH, Obermayer K. An online spike detection and spike classification algorithm capable of instantaneous resolution of overlapping spikes. J Comput Neurosci. 2010;29:127–48.
18. Gray PR. Conditional probability analyses of the spike activity of single neurons. Biophys J. 1967;7:759–77.
19. Benesty J, Chen J, Huang Y. On the importance of the Pearson correlation coefficient in noise reduction. IEEE Trans Audio Speech Lang Process. 2008;16:757–65.
20. Agustín-Blas LE, Salcedo-Sanz S, Jiménez-Fernández S, Carro-Calvo L, Del Ser J, Portilla-Figueras JA. A new grouping genetic algorithm for clustering problems. Expert Syst Appl. 2012;39:9695–703.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.