A Simple Method to Simultaneously Detect and Identify Spikes from Raw Extracellular Recordings

Panagiotis C. Petrantonakis * and Panayiota Poirazi *

Computational Biology Laboratory, Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology-Hellas, Heraklion, Greece

The ability to track when and which neurons fire in the vicinity of an electrode, in an efficient and reliable manner can revolutionize the neuroscience field. The current bottleneck lies in spike sorting algorithms; existing methods for detecting and discriminating the activity of multiple neurons rely on inefficient, multi-step processing of extracellular recordings. In this work, we show that a single-step processing of raw (unfiltered) extracellular signals is sufficient for both the detection and identification of active neurons, thus greatly simplifying and optimizing the spike sorting approach. The efficiency and reliability of our method is demonstrated in both real and simulated data.

Keywords: multi-electrode recordings, spike sorting, extracellular signals

INTRODUCTION

Multi-electrode arrays comprise one of the most valuable tools for neuroscience research (Buzsáki, 2004). Neural prosthetics (Gilja et al., 2012), Brain Machine Interfaces (BMIs; Schwarz et al., 2014; Moxon and Foffani, 2015), and neurophysiology experiments (de Lavilléon et al., 2015) are some of the areas where multiple electrodes are routinely used to extract neuronal-activity related information. Analysis of that information toward the identification and discrimination of activity events (spikes) from different neurons, termed Spike Sorting (SS; Gibson et al., 2012), is a multi-step, computationally expensive process, thus limiting its application on wireless technologies that would enable experiments under more ecological conditions (Rey et al., 2015).

In this work we propose a novel, spike detection and identification algorithm that is applied on raw (unfiltered) extracellular recordings and effectively bypasses several steps of the common SS process. As such, it paves the way for new, wireless on-chip algorithmic implementations, considerably reducing the amount of data that needs to be transmitted through the wireless link. The commonly used state-of-the-art SS process (Einevoll et al., 2012; Figure 1) is divided in five basic steps: (1) the raw extracellular signal derived from a given electrode is band-pass filtered to remove the low frequency component (Local Field Potential) and reveal the spiking activity; (2) spikes are detected; (3) spikes are aligned; (4) specific features are extracted from spikes; (5) data are clustered based on extracted features. Each cluster is assumed to represent a different neuron. In this study, the raw, extracellular signal (derived from a single or multiple electrodes) is processed once and both spike times and discriminatory features are extracted simultaneously (Figure 1).

The proposed algorithm is based on the finding that neuronal spike waveforms are compressible in various basis sets (Thorbergsson et al., 2014). In other words, the spiking waveforms can be decomposed into a set of simpler waveforms, in terms of a coefficient vector, and a support set of few coefficients is sufficient to reliably reconstruct the initial waveform. More importantly, this support set of coefficients is specific for different spike waveforms, namely different neurons (Charbiwala et al., 2011) and obviously different from non-spiking waveforms. Thus, it is suggested that the
estimation of, e.g., the euclidean distance, between coefficient sets extracted from consecutive, sliding windows of an extracellular signal would reflect the existence of neuronal spiking. Even better, due to the specificity of the coefficient sets, the estimation of euclidean distance using a sliding window would also discriminate among waveforms originating from different neurons. For instance, the use of a simple basis set of discrete cosines and the decomposition of the signals in a sliding window framework using the Discrete Cosine Transform (DCT; Narasimha and Peterson, 1978) would result in a new signal with the aforementioned properties. Interestingly, since DCT is an orthogonal transformation (thus distance is preserved), the estimation of the new signal can be reduced to a very simple calculation: the euclidean distance between the consecutive windowed signals in their original form would be equivalent to the distance between their decompositions. Our approach is based on this observation.

The main processing step of the proposed approach is the segmentation of the raw, single/multiple-electrode signal using a sliding, fixed-length window. A new signal, $D_w$, where $w$ denotes the length of the window, is produced by estimating the euclidean distances between the different signal segments (see Section Materials and Methods and Supplementary Figure 1). Within the resulting signal $D_w$, the time points of the local maxima indicate the time points of the spiking events while their amplitudes serve as unique markers (identifiers) for each specific spike waveform, namely each distinct active neuron (Figure 2). Strikingly, this is true for both the band-pass filtered and the raw, unfiltered version of the extracellular signal. In essence, different support sets of coefficients for different spike waveforms, lead to different amplitudes of local maxima for $D_w$. Thus with a single processing step, the proposed method transforms the unfiltered, extracellular signal into a new signal ($D_w$) that reveals both the occurrence and identity of spikes in a simple and highly efficient manner.

MATERIALS AND METHODS

$D_w$ and Feature Space Extraction

The $D_w$ signal consisted of the Euclidian distance values between the successive $X_k$ signal segments. For instance, $D_w(1) = \sqrt{\sum_{k=1}^{w} (X_k^1 - X_k^2)^2}$ where $X_k^1$ and $X_k^2$ are the first two successive segments of length $w$ using a predefined step. In this work we
used $w = 5$ and $w = 15$ as most significant decomposition coefficients varied within this range (Charbiwala et al., 2011). The step value was 1 sample. The $D_5$ and $D_{15}$ signals were used to extract the proposed feature space, which consisted of the polar coordinates of the amplitudes of their local maxima (Figure 2). Polar coordinates instead of Cartesian ones were used in order to avoid the linearity that is exhibited between $D_5$ and $D_{15}$ values (see Figure 2B). Expansion of the methodology to account for tetrode recordings was performed by estimating $D_w$ using four channels simultaneously. Specifically, the $X_k$ segments were concatenated before the euclidean distance calculation, resulting in a single $D_w$ signal that incorporated information from all four recordings (see Supplementary Figure 1B).

**Spike Detection**

In this work, spike-detection (i.e., the detection of maxima in the $D_w$ signal) and feature-extraction (i.e., amplitudes of maxima in the $D_w$ signal) are essentially a single step, as the maximum values of $D_w$ are detected mainly around the time points of a spike occurrence. The detection of maxima is based on $D_w$ amplitude thresholding, in particular using the $D_5$ signal. In all examples reported here (real or simulated), the threshold (Quiroga et al., 2004) was set to $T = 2\sigma_n$ where $\sigma_n = \text{median} \left\{ |D_w| \right\}$ is an estimate of the standard deviation of the background noise (Donoho and Johnstone, 1994). A lower threshold could also be used as the thresholding process is applied on the $D_w$ signal and not on the extracellular signal itself and, thus, it would not affect the subsequent clustering process. Moreover, to avoid multiple detections of the same spike, a minimum sample distance between two detected maxima of 15 samples (equal to the maximum window length used) was imposed. Finally, due to the sharper form of maxima peaks in $D_5$ compared to $D_{15}$, maxima detection was first performed in $D_5$. Afterwards, identification of the maxima amplitude in $D_{15}$...
Datasets

The examples analyzed were taken from two public databases, one consisting of real extracellular recordings and one consisting of simulated signals with integrated spike waveforms. The former is the hc-1 database (Henze et al., 2009) that consists of simultaneous intracellular and extracellular recordings in the CA1 hippocampal region of anesthetized rats. It contains both intracellular and extracellular recordings from the same neuron. Ten different signals, corresponding to 10 different cells, were analyzed from this database. For the single channel analysis the experimental sessions and single recording channels (see Supplementary Figure 2) were selected in order to account for various firing patterns (bursting etc.) and to have a reliable representation of the intracellularly recorded neuron in the extracellular recordings. Some of the recordings were from tetrodes and some from hexatetrodes settings with two shanks. In the four-channel analysis that we followed for the estimation of \( D_n \), in the case of double hexatetrodes, only four channels were used. These were selected based on their distance from the single electrode that was initially chosen for the single electrode analysis, in order to avoid problems that have to do with distances between electrodes (Rossant et al., 2015).

The latter database consists of four different datasets that refer to different levels of difficulty in terms of spike detection and identification (Quiroga et al., 2004), hereafter referred as datasets 1, 2 (easy) and 3, 4 (difficult). Moreover, the four different datasets were constructed using different levels of background noise (Quiroga et al., 2004). The noise level was determined from its standard deviation relative to the amplitude of the integrated spike waveforms. Four levels of noise, associated with four different standard deviation values, i.e., 0.05 (noise level 1 etc.), 0.1, 0.15, 0.2 were used. For dataset 1, four additional levels of noise were applied with deviations equal to 0.25, 0.3, 0.35, and 0.4 (see Supplementary Figure 3), since spike identification was relatively easy (more information about the construction of the simulated data used can be found in Quiroga et al., 2004). It should be noted that simulated data lack, in principle, the low frequency component, i.e., cannot be considered as raw (unfiltered) simulated signal. They are primarily used for further evaluation of the proposed approach for multiple ground truth neuronal activity.

Classification

For the classification of spikes corresponding to the intracellularly recorded neuron (hc-1 database) against all other “noisy” spikes (spikes that do not correspond to that particular neuron) the \( k \)-Nearest Neighbor (k-NN) classifier (Cover and Hart, 1967) was used. For each one of the ten different cells (see Supplementary Figure 2) a 100-iteration procedure was implemented. Every time, 50% of the data were randomly selected for training and the rest were used for testing. The mean type I and type II error values were used for the extraction of Figure 3 (see also Supplementary Table 1 for mean classification performance for each one of the different cells). For the \( k \)-NN algorithm, the number of nearest neighbors was set to \( k = 4 \).

\[ L_{ratio} \text{ and Isolation Distance (ID) Estimation} \]

In order to evaluate the capacity of the proposed features to separate the spikes that belong to different neurons we used the \( L_{ratio} \) and the Isolation Distance (ID) measures (Schmitzer-Torbert et al., 2005). Both measures make use of the Mahalanobis Distance (MD; Mahalanobis, 1936). The MD of spike \( i \) from cluster \( C \) is estimated using the formula: \( MD_{i,C} = (x_i - \mu_C)^T \Sigma_{C}^{-1} (x_i - \mu_C) \), where \( x_i \) is the feature vector for spike

\[ \text{Type I (False Positives, FP) and Type II (False Negatives, FN) error rates against } L_{ratio} \text{ and Isolation Distance (ID).} \]

\[ \text{Boxplots 1, 2, 3, and 4 depict the IDs of the clusters of neurons 1, 2, 3 and falsely detected spikes (see Figure 2D).} \]

\[ \text{Boxplots 5, 6, 7, and 8 are the IDs of the clusters of neurons 1, 2, 3, and falsely detected spikes from the corresponding cluster itself (within-cluster IDs).} \]

\[ \text{All comparisons (ANOVA) of four pairs of within- and between-cluster IDs, e.g., pair 1 and 5, reveal high statistical significance (} p < 0.001). \]
i, μC is the mean of the feature vector values in cluster C, and ΣC is the corresponding covariance matrix. The first measure, i.e., \( L_{\text{ratio}} \) for cluster C is calculated by \( L_{\text{ratio}}(C) = \frac{L(C)}{n_C} \) where \( n_C \) is the number of spikes in cluster C. The quantity \( L \) for C is calculated as \( L(C) = \sum_{i \neq C} 1 - CDF_{\chi^2}(i) \) for \( i \) not in C. This is the cumulative distribution function of the \( \chi^2 \) distribution with degrees of freedom that equal to the length of the feature vector. A low value of \( L \) indicates that cluster C is well separated from the rest of the data and vice versa.

The ID value of a cluster C that contains \( n_C \) spikes is defined as the MD value of the \( n_C \) closest noise spike (a spike that does belong to cluster C). ID estimates how distant the spikes of cluster C are from the other spikes. In Figure 3A the \( L_{\text{ratio}} \) and ID values were estimated for the spikes corresponding to the intracellularly recorded cell (cluster C). For Figure 3B the ID values were calculated both between clusters and within clusters. In particular, for the boxplots 1, 2, 3, and 4 (see Figure 3B), ID was estimated for spikes that belong to cells 1, 2, 3, and spikes that were falsely detected (black dots), respectively. On the other hand, the spikes for boxplots 5, 6, 7, and 8, ID was estimated within the same cluster. For instance, in boxplot 5, the spikes that belong to cell 1 were randomly split into two groups, which consisted of 40 and 60% of the spikes of that cluster, respectively. Subsequently, ID was estimated using the first group (40%) as the reference cluster C. This procedure was repeated 10 times and the mean ID was finally used for boxplot 5. The same process was applied for boxplot 6, 7, and 8, i.e., for cells 2, 3, and 3 falsely detected spikes, respectively.

**RESULTS**

The proposed analysis was applied both on real (Henze et al., 2009) and simulated (Quiroga et al., 2004) data (see Section Materials and Methods, Figure 2, and Supplementary Figures 2, 3) from single and multiple electrode (tetrode) recordings. Spike detection success rates for single electrode data reached 94.82 ± 9.08 and 92.22 ± 2.64 for real and simulated data, respectively (see Supplementary Tables 1, 2). In both cases, spike detection parameters were identical (see Section Materials and Methods) and the feature vectors consisted of the polar coordinates of the amplitudes of the local maxima in \( D_5 \) and \( D_{15} \) signals (Figure 2 and Section Materials and Methods). Spike detection rates of the proposed method are very similar to those of a complex, multi-step state-of-the-art method (Quiroga et al., 2004) applied on the same simulated data, which demonstrated mean detection rates up to 91.08 ± 6.5 (see Supplementary Table 2). Another important advantage of our method lies in the thresholding applied for spike detection which in our case is done in \( D_w \) (see Section Materials and Methods), as opposed to existing methodologies where the thresholding is applied on the filtered extracellular signal. Since detected maxima are subsequently used for clustering, false spike detection due to low thresholds would lead to the formation of additional clusters with very small \( D_w \) values that could ultimately be discarded (e.g., see black dots cluster in Figure 2D), thus not influencing the performance of the method and avoiding the process of finding optimum thresholds for spike detection (Rizk and Wolf, 2009).

In addition to detection accuracy, the discrimination power of the extracted features (polar coordinates of the amplitudes of the local maxima in \( D_5 \) and \( D_{15} \)) was assessed based on the \( L_{\text{ratio}} \) and Isolation Distance (ID) measures (Schmitzer-Torbert et al., 2005). Both measures reveal the degree of isolation of a particular data group (in this case the intracellularly recorded neuron, Figure 2C, red dots) from the rest of the data in a feature space (see Section Materials and Methods and Supplementary Figure 4). Their relation with the type I and type II error rates reflects the capacity of the proposed features to isolate spikes from a particular neuron (Schmitzer-Torbert et al., 2005). \( L_{\text{ratio}} \) and ID were related linearly on the log scale to type I and type II error rates (see Figure 3A and Section Materials and Methods), confirming the reliability of the proposed features. Type I and II errors were estimated using a k-NN classifier (see Section Materials and Methods).

Moreover, we examined whether the proposed feature space assigns spike waveforms of different neurons to discrete and different clusters (see Figure 2D, red, blue, and green dots). Thus, we estimated the ID of each one of the three ground truth clusters of the simulated datasets (Quiroga et al., 2004), firstly, between clusters and then within the same cluster (see Section Materials and Methods and Figure 3B). In all cases tested, ID values within the same cluster were significantly smaller than the ID values estimated between clusters, confirming the initial assumption that waveforms from different neurons have different support sets of coefficients, which is, in turn, reflected in the proposed feature space (\( D_w \) maxima).

In the above-mentioned examples, the \( D_w \) signal was estimated using recordings from a single electrode. Despite this radical simplification, the method was shown to achieve reliable discrimination of the intracellularly recorded neuron in terms of the ID and \( L_{\text{ratio}} \) measures. The proposed methodology however can easily expand to account for multi-electrode experiments. Toward this goal, we reformulated the \( D_w \) estimation process to incorporate tetrode signals. Reformulation amounts to a simple concatenation of signal segments corresponding to individual electrodes (see Section Materials and Methods and Supplementary Figure 1B). We then estimated the ID and \( L_{\text{ratio}} \) measures of the isolation of the cluster under consideration (intracellular recorded neuron) using the new \( D_w \) signal. While median ID and \( L_{\text{ratio}} \) values were slightly higher and lower (showing better cluster isolation compared to the single electrode case), respectively, (Figure 4, single: median ID = 2.77, median \( L_{\text{ratio}} \) = 4.5; tetrode: median ID = 3.08, median \( L_{\text{ratio}} \) = 1.4) compared to the single electrode analysis, the mean difference was not significant (ANOVA). Importantly, though, the distribution of ID and \( L_{\text{ratio}} \) values was narrower (shorter whiskers of boxplots in Figure 4) exhibiting a relevant consistency within the 10 different datasets of real recordings used. Thus, the expansion of the proposed methodology to tetrode signals both slightly enhanced the isolation of the cluster under consideration (intracellularly recorded neuron) and the consistency of the ID and \( L_{\text{ratio}} \) values among the datasets.
**Computational Complexity**

To demonstrate the efficiency of the proposed method, we examined its computational complexity compared to commonly used state-of-the-art algorithms for spike detection and feature extraction (Gibson et al., 2008). With respect to spike detection (detection of maxima in $D_w$), the computational complexity of our method is equal to that of the most widely used thresholding method for filtered extracellular signals, i.e., 1.058 additions and multiplications per waveform sample (Gibson et al., 2008) which is also the one with the lowest complexity. Note that we by-pass the filtering step as our method is applied on raw signals. With respect to feature extraction, an explicit comparison is not possible as the proposed methodology bypasses several processing steps that would add to the complexity of the ultimate spike sorting algorithm. Nonetheless, the obvious simplistic and straightforward calculation of the $D_w$ makes the proposed approach a very efficient framework in terms of complexity and comprises a significant step closer to wireless on-chip algorithmic implementations.

**DISCUSSION**

In sum, we introduced a novel, simple, and efficient approach for simultaneous detection and identification of spikes from raw, unfiltered extracellular recordings. In addition to redesigning the widely used spike sorting process via bypassing several processing steps, the proposed method also overcomes important optimization issues such as the identification of thresholds for spike detection, the scaling to larger sets of recordings and the parallelization of the method. Both scaling and parallelization are straightforward as the $D_w$ signal can be estimated for both single and multiple electrodes using a very similar algorithmic process (see Section Materials and Methods).

In this work we did not propose an unsupervised clustering algorithm for the extracted features. Nevertheless, the cluster isolation analysis revealed the capacity of proposed features to accurately distinguish the activity generated by different neurons.

These findings suggest that our methodology and extracted features can be coupled to any clustering algorithm in order to complete the spike sorting procedure. Importantly, for datasets with low noise levels (e.g., simulated datasets, Supplementary Figure 3) cluster discrimination can be achieved using just one of the two polar coordinates, thus further enhancing the efficiency of our method. This is possible because in such low noise cases, even $D_w$ maxima (before polar coordinates estimation) lie in different amplitude levels (identifiers) for different neurons (Supplementary Figure 5), thus, allowing for the implementation for thresholding based clustering algorithms that exhibit significantly low computational complexity and would further support the ultimate goal for on-chip spike sorting implementations.

Future work will be toward that direction, i.e., to investigate the possibility of exploiting the proposed method toward small, wireless, and low-energy consuming recording devices that would either transmit only the time point and the amplitude of the detected maxima in the $D_w$ signal or even perform on-chip, holistic spike sorting realization. Finally, the expansion of the proposed approach to multi-electrode settings (thousands of electrodes) will also be investigated, further diffusing its applicability in contemporary recording settings.

**ACKNOWLEDGMENTS**

This work was supported by the ERC Starting Grant dEMORY (ERC-2012-StG-311435) and the BIOSYS research project, Action KRIPI, project No MIS-448301 (2013SE01380036) that was funded by the General Secretariat for Research and Technology, Ministry of Education, Greece.

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fnins.2015.00452
REFERENCES

Buzsáki, G. (2004). Large-scale recording of neuronal ensembles. *Nat. Neurosci.* 7, 446–451. doi: 10.1038/nn1233

Charbiwala, Z., Karkare, V., Gibson, S., Marković, D., and Srivastava, M. B. (2011). “Compressive sensing of neural action potentials using a learned union of supports,” in *Proceedings of the 2011 International Conference on Body Sensor Networks* BSN 2011 (Dallas, TX), 53–58.

Cover, T., and Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory* 13, 3–22. doi: 10.1109/TIT.1967.1053964

de Lavilléon, G., Lacroix, M. M., Rondi-Reig, L., and Benchenane, K. (2015). Explicit memory creation during sleep demonstrates a causal role of place cells in navigation. *Nat. Neurosci.* 18, 493–495. doi: 10.1038/nn.3970

Donoho, D. L., and Johnstone, I. M. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika* 81, 425–455. doi: 10.1093/biomet/81.3.425

Einevoll, G. T., Franke, F., Hagen, E., Pouzet, C., and Harris, K. D. (2012). Towards reliable spike-train recordings from thousands of neurons with multielectrodes. *Curr. Opin. Neurobiol.* 22, 11–17. doi: 10.1016/j.conb.2011.10.001

Gibson, S., Judy, J. W., and Markovic, D. (2008). Comparison of spike-sorting algorithms for future hardware implementation. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2008, 5015–5020. doi: 10.1109/IEMBS.2008.4650340

Gibson, S., Judy, J. W., and Markovic, D. (2012). Spike sorting: the first step in decoding the brain. *IEEE Signal Process. Mag.* 29, 124–143. doi: 10.1109/MSP.2011.941880

Gilja, V., Nuyujukian, P., Chestek, C. A., Cunningham, J. P., Yu, B. M., Fan, J. M., et al. (2012). A high-performance neural prosthesis enabled by control algorithm design. *Nat. Neurosci.* 15, 1752–1757. doi: 10.1038/nn.3265

Henze, D. A., Harris, K. D., Borbely, Z., Csicsvari, J., Mamiya, A., Hirase, H., et al. (2009). Simultaneous intracellular and extracellular recordings from hippocampus region CA1 of anesthetized rats. doi: 10.6080/K02Z13FP

Mahalanobis, P. C. (1936). On the generalised distance in statistics. *Proc. Natl. Inst. Sci. India* 2, 49–55.

Moxon, K., and Foffani, G. (2015). Brain-machine interfaces beyond neuroprosthetics. *Neuron* 86, 53–67. doi: 10.1016/j.neuron.2015.03.036

Narasimha, M. J., and Peterson, A. M. (1978). On the computation of the discrete cosine transform. *Commun. IEEE Trans.* 26, 934–936. doi: 10.1109/TCOM.1978.1094144

Quiroga, R. Q., Nadasy, Z., and Ben-Shaul, Y. (2004). Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. *Neural Comput.* 16, 1661–1687. doi: 10.1162/08997660474201631

Rey, H. G., Pedreira, C., and Quiroga, R. Q. (2015). Past, present and future of spike sorting techniques. *Brain Res. Bull.* 1–12. doi: 10.1016/j.brainresbull.2015.04.007

Rizk, M., and Wolf, P. D. (2009). Optimizing the automatic selection of spike detection thresholds using a multiple of the noise level. *Med. Biol. Eng. Comput.* 47, 955–966. doi: 10.1007/s11517-009-0451-2

Rossant, C., Kadir, N. S., Goodman D. F. M., Schulman, J., Belluscio, M., Buzsaki, G., et al. (2015). Spike sorting for large, dense electrode arrays. bioRxiv. doi: 10.1101/015198

Schmitzer-Torbert, N., Jackson, J., Henze, D., Harris, K., and Redish, A. D. (2005). Quantitative measures of cluster quality for use in extracellular recordings. *Neuroscience* 131, 1–11. doi: 10.1016/j.neuroscience.2004.09.066

Schwarz, D., Lebedev, M., Hanson, T., Dimitrov, D., Lehew, G., Meloy, J., et al. (2014). Chronic, wireless recordings of large-scale brain activity in freely moving rhesus monkeys. *Nat. Methods* 11, 670–676. doi: 10.1038/nmeth.2936

Thorbergsson, P. T., Garwicz, M., Schouenborg, J., and Johansson, A. J. (2014). Strategies for high-performance resource-efficient compression of neural spike recordings. *PloS ONE* 9:e93779. doi: 10.1371/journal.pone.0093779

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Petrantonakis and Poirazi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.