Journal of Neuroscience Methods 184 (2009) 285–293

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

# Realistic simulation of extracellular recordings

Juan Martinez, Carlos Pedreira, Matias J. Ison, Rodrigo Quian Quiroga*

*Department of Engineering, University of Leicester, Leicester, United Kingdom*

ARTICLE INFO

*Article history:*
Received 28 July 2009
Received in revised form 18 August 2009
Accepted 19 August 2009

*Keywords:*
Simulations
Spike detection
Spike sorting
Extracellular recordings

ABSTRACT

In this paper we present an efficient method to generate realistic simulations of extracellular recordings. The method uses a hybrid and computationally simple approach, where the features of the background noise arise naturally from its biophysical process of generation. The generated data resemble the characteristics of real recordings, as quantified by the amplitude and frequency distributions. Moreover, we reproduce real features that are specially challenging for the analysis of extracellular data, such as the presence of sparse firing neurons and multi-unit activity. We compare our simulations with real recordings from the human medial temporal lobe and exemplify their use for testing spike detection and sorting algorithms. These results show that this technique provides an optimal scenario for generating realistic simulations of extracellular recordings.



## 1. Introduction

To a large extent, neuroscientists have the goal of understanding how mental processes are encoded by neurons in the brain. A well established methodology to address this question is to study the extracellular activity of individual neurons, as recorded from microelectrodes in the brain. The analysis of single-cell responses is done by first detecting the spikes from the background noise (spike detection) and then by identifying the activity of different recorded neurons, grouping the recorded spikes into clusters based on the similarity of their shapes – a procedure known as spike sorting (Lewicki, 1994; Fee et al., 1996a; Harris et al., 2000; Letelier and Weber, 2000; Hulata et al., 2002; Pouzat et al., 2002; Quian Quiroga et al., 2004). The quality of the resulting data depends crucially on the particular detection and sorting technique used, whose performance is usually hard to establish with real data. Consequently, the use of synthetic data to compare the algorithm outcome with the original spike labels provides an optimal framework to quantify the performance of these methods.

In principle, two different approaches might be taken to generate synthetic extracellular recordings. On the one hand, one could generate simple simulations by adding spikes to Gaussian noise (Lewicki, 1994). This approach is very fast and easy to implement but it misses important features of the real recordings that make spike sorting especially challenging. In particular, some characteristics, like the presence of non-Gaussian distribution of clusters, multi-unit activity and spectral similarity between noise and spikes, might be critical for spike sorting. On the other hand, detailed models of extracellular waveforms (Hines and Carnevale, 1997; Holt and Koch, 1999; Gold et al., 2006) could be more accurate but too computationally intensive when simulating extracellular recordings generated by a large number of neurons.

Here we explored the possibility of a compromise between both alternatives, by developing relatively easy simulations that reproduce relevant real features of extracellular recordings for testing spike detection and sorting algorithms. In particular, our simulations will try to replicate the amplitude and spectral distributions of the noise, multi-unit and single-unit activity found in extracellular recordings. These simulations follow the approach presented by Quian Quiroga et al. (2004), where real spike shapes were used to create background noise and single-unit activity. Here, we introduced several improvements. First, we propose a biophysical strategy for generating the background noise. Second, we reproduce the multi-unit activity, thus making the synthetic data very similar to real recordings. Multi-unit activity is commonly found in extracellular recordings and it represents the activity of several distant neurons whose spikes can be detected but are not large enough to be clustered, because the difference in their shapes is masked by the background noise. This situation typically produces clusters with high variance and low signal-to-noise ratio. Third, we replicate actual values found in real recordings in terms of amplitude and power spectrum distributions. The resulting simulations also allowed the study of sparse firing neurons (Hahnloser et al., 2002; Perez-Orive et al., 2002; Quian Quiroga et al., 2005, 2007), which due to their very low baseline firing rates might be masked by neurons firing at higher frequencies, making their identification particularly difficult. Our method also allows an efficient generation of synthetic datasets, thus providing a simple strategy to generate large number of realistic experimental scenarios that are useful to

* Corresponding author. Tel.: +44 116 252 2314; fax: +44 116 252 2619.
*E-mail address:* rqqg1@le.ac.uk (R. Quian Quiroga).

0165-0270/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jneumeth.2009.08.017

quantify the performance and limits of different spike sorting algorithms.

## 2. Materials and methods

We generated synthetic extracellular recordings that modeled the contribution of the background noise, multi-unit and single-unit activity. The simulations were created using a database with 594 different averaged spike shapes, taken from real recordings in monkey neocortex and basal ganglia. In order to reproduce conditions of real recordings – in which misalignments may occur if the peak of the spikes do not occur during a recording sample, thus introducing misalignments – we used the following procedure: first, the data were first simulated at a sampling frequency of 96 kHz. Then, using interpolated waveforms of the original spike shapes, the spike times were simulated to fall continuously between samples (Quian Quiroga et al., 2004). Finally, the data were downsampled to 24 kHz.

Fig. 1 shows a summarized description of the method. Considering a recording electrode at the centre of the sphere, the extracellular space was divided into three major areas: (1) background noise; (2) multi-unit activity and (3) single-unit activity. The main idea of the method is to generate, in a computationally efficient way, the relevant features – such as amplitude distribution and power spectrum – that allow reproducing realistic extracellular recording scenarios for the three zones described in the figure. The neurons in area 1 generate the background noise, whose activity has been constructed based on the contributions of many individual point-source neurons. This strategy avoids complex models of single cells that do not have a relevant effect on the total signal at large distances from the recording electrode. The activity of neurons located at areas 2 and 3 has been simulated using spike waveforms from the above mentioned database and with amplitude distributions that follow real multi-unit and single-unit peak amplitude distributions, given that these values are accessible from real recordings. To consider the fact that different recording conditions can vary the levels of the background noise activity, we considered the amplitude values of the spikes of areas 2 and 3 relative to the detection threshold, which was in turn established from an estimation of the standard deviation of the noise (see Eq. (1)). Finally, the synthetic extracellular recordings were generated by superimposing the spikes corresponding to multi-unit and single-unit activity to the background noise, as detailed below. For this we used different single-unit amplitude levels, thus replicating the conditions of different signal-to-noise ratios.

### 2.1. Generation of background noise

The core of our approach to generate the noise activity was to simulate the contribution of distant neurons, as given in real extracellular recordings. This provides a hybrid simulation strategy, in which the characteristics of the noise activity, such as its amplitude and frequency distribution, arise naturally from the realistic biophysical process of its generation and it is not imposed beforehand. Moreover, this approach reproduces real noise characteristics without modeling the computationally expensive details of the contribution of each neuron. For this, the far-apart neural activity was modeled by superimposing millions of spike shapes (one spike per sample of generated data) placed at uniformly distributed random times. The amplitude distribution of all the waveforms was obtained by first assuming that each neuron generates an action potential of peak amplitude 1. These simulated neurons represent different point-source charges distributed uniformly within a sphere of radius 1 (in adimensional units). Only far-away spikes, at a distance larger than 0.5, were considered for generating the background noise. The activity of close-by neurons was considered to contribute to the multi-unit and single-unit activity (areas 2 and 3 in Fig. 1), which was modeled following the typical amplitude ranges found in real data as described in the sections below. For a single spike shape $s_i$ its peak amplitude $a_i$ at the recording site

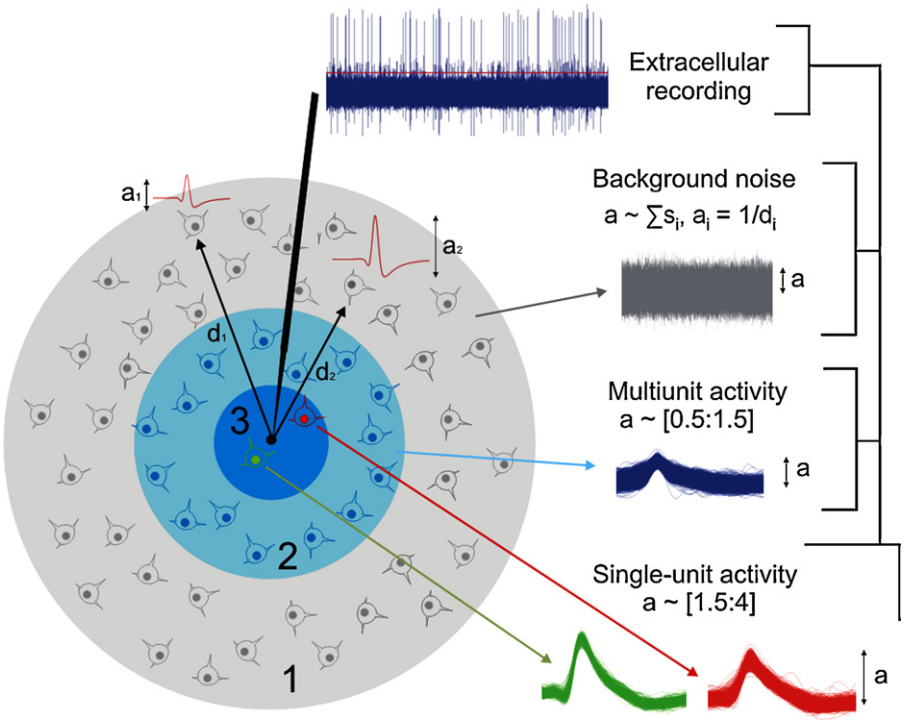


**Fig. 1.** Overview of the simulation approach. Neurons in area 1 generate the noise activity, which is given by the superimposition of their spikes $s_i$. Each peak amplitude $a_i$ is modulated by its distance $d_i$ to the recording electrode (in the center of the sphere). The activity of neurons in areas 2 and 3 (multi-unit and single-unit activity, respectively) is created by adding spikes on top of the background noise, following typical amplitude distributions $a_i$ measured for multi-unit and single-unit spikes.

(the centre of the sphere) was modeled to be equal to $1/d_i$ ($d_i$: distance of the cell $i$ from the centre), following Coulomb's law (Rall, 1962). As mentioned above, this assumption leads to a large simplification in our simulation of the extracellular recordings, since we do not consider specific morphologies and physiological characteristics of distant neurons. The effect of other possible sources of electrical activity that might contribute to the noise power spectrum, like electronic noise, axons, dendrites or synapse currents (Llinas, 1988; Farrant et al., 1994), was simulated by adding Gaussian noise. Note that these simulations were generated to test spike detection and sorting algorithms, for which the extracellular signal is typically highpass filtered with a cutoff frequency of 300 Hz or larger to visualize the spiking activity. Therefore, we did not consider or try to reproduce sources of low frequency activity – i.e. below 300 Hz – as the ones considered for simulations of local field potentials (Brunel and Wang, 2003; Bedard et al., 2006; Mazzoni et al., 2008).

### 2.2. Generation of multi-unit activity

Multi-unit activity – i.e. spikes that can be detected but cannot be clustered into different single units due to their relatively small amplitude – is largely produced by neurons located around 100–150 μm from the electrode tip (Buzsaki, 2004). The amplitude of the recorded multi-unit activity is close to the detection threshold (spikes from area 2 in Fig. 1). Note that in this case, the noise masks any difference between the spikes of different neurons, making the identification of single-units unfeasible. Spikes corresponding to multi-unit activity tend to produce large and typically non-Gaussian clusters. This is because clusters are created by the activity of more than one neuron and the amplitude distribution is truncated by the detection threshold.

In our simulations, multi-unit activity was created by mixing the activity of our whole database of 594 spike shapes using amplitudes uniformly distributed between 0.5 and 1.5 times the level of the detection threshold. Following previous criteria (Quian Quiroga et al., 2004) the detection threshold was set at 4 times the estimation of the standard deviation of the noise (see Eq. (1)). These amplitude ranges were chosen to follow the distribution of amplitude values for multi-unit activity found in real data (see Supplementary Material). For each unit contributing to the multi-unit activity the spike times were generated using a Poisson distribution with a mean firing rate of $20/N$ Hz (where $N$ is the number of different units used to create the multi-unit activity), which summed up gives an overall multi-unit firing rate of 20 Hz.

### 2.3. Generation of single-unit activity

The single-unit activity (area 3 in Fig. 1, spikes in green and red) was simulated by using spikes of different shapes added to the background noise. The peak spike amplitudes were assigned values between 1.5 and 4 times the detection threshold, which falls within the range of typical values found in real recordings (see Supplementary Material). For each single-unit, the variability of the spike amplitudes was given by the background noise. The spike train of each single-unit followed a Poisson process, with a mean firing rate randomly selected between 0.5 and 5 Hz. Spikes that fell within the 2 ms following another spike were removed, to introduce a refractory period and delete overlapping spikes.

### 2.4. Real data

The simulated extracellular data presented here came from a 30-min recording from the human medial temporal lobe. The subject was a pharmacologically intractable epileptic patient who was implanted with intracranial electrodes for clinical reasons. Besides the intracranial EEG contacts, each probe had a total of 9 micro-wires at its end, with 8 active recording channels and 1 reference to record single-neuron activity (Fried et al., 1997). The differential signal from the micro-wires was amplified, sampled at 32 kHz and filtered between 300 and 3000 Hz. The continuous data used in this study are available at www.le.ac.uk/neuroengineering. The main features of this recording were similar to other 192 recordings from the human medial temporal lobe, in terms of the power spectrum and amplitude characteristics, as shown in Fig. S1.

### 2.5. Spike detection and sorting

The real and synthetic extracellular recordings were analyzed using 'Wave_Clus', an unsupervised spike detection and sorting algorithm (Quian Quiroga et al., 2004). First, the data were bandpass filtered within the range of 300–3000 Hz, using a non-casual filter to avoid filtering distortions of the spike shapes (Quian Quiroga, 2009). Spike detection was done using an amplitude threshold set at 4 times an estimation of the standard deviation of the noise, using the following expression:

$$Thr = 4\sigma_n; \qquad \sigma_n = \text{median}\left\{\frac{|x|}{0.6745}\right\} \tag{1}$$

where $x$ is the bandpass filtered signal. For each detected spike, 64 datapoints were stored and aligned with their maximum at datapoint 20 for further processing. After spike storage, features of the spike shapes were extracted by using the wavelet transform. Wavelet coefficients represent the features of the spike shapes at different scales (i.e. frequency ranges) and times. The coefficients that best separated the signal were selected by using a Kolmogorov–Smirnof test of Normality (as in Quian Quiroga et al., 2004), choosing the coefficients with the least Normal distribution (i.e. those more likely to represent more than 1 cluster of spike shapes). In the last step, 'Wave_Clus' does an unsupervised classification to group spikes into classes using superparamagnetic clustering (Blatt et al., 1996; Quian Quiroga et al., 2004). The code for spike detection and sorting is available at www.le.ac.uk/neuroengineering.

## 3. Results

We generated 5 sets of simulations of 2 min each with a sampling rate of 24 kHz. Each simulation contained the activity of one multi-unit and two single-units. The noise level $\sigma_n$ was normalized to give a detection threshold of 28 μV, as typically set for real data (see Fig. S1 in Supplementary Material). The peak amplitudes of the spikes generating the multi-unit clusters were uniformly distributed between 0.5 and 1.5 times the detection threshold for all the simulations. Each single-unit in examples 1–5 generated a spike train following a Poisson distribution. The values of peak amplitude and mean firing rates are described in Table 1. The different amplitudes (i.e. signal-to-noise ratios) and firing rates simulate different recording conditions to test spike sorting algorithms in varying regimes. The processing time to generate each example was around

**Table 1**
Characteristics of the 5 simulated data sets. The values of the single-unit amplitudes are all relative to the detection threshold\α and $r^2$ correspond to the slope of decay of the power spectrum and the correlation coefficient of the linear fit.

| Example number | Amplitude | Firing rate (Hz) | α | $r^2$ |
|---|---|---|---|---|
| 1 | 4 | 1 | 1 | 0.991 |
| 2 | 4 | 5 | 1.18 | 0.994 |
| 3 | 2 | 5 | 0.9 | 0.984 |
| 4 | 2 | 5 | 0.93 | 0.984 |
| 5 | 3 | 0.5 | 0.97 | 0.988 |

9 min in an Intel Core 2 PC with a clock frequency of 2.4 GHz. All the examples of synthetic data used in this paper are available from download at www.le.ac.uk/neuroengineering.

### 3.1. Real data

Fig. 2a shows 30 s of the continuous real data, which contains spikes with amplitudes between 30 and 130 μV. The estimation of the standard deviation of the noise ($\sigma_n$ in Eq. (1)) was in this case equal to 6.5 μV and the threshold for spike detection was of 26 μV (i.e. 4 times this value). For the same data, in Fig. 2b we show the amplitude distribution of the signal, which followed a Gaussian distribution. Note that in the raw (bandpass filtered) data as well as in the amplitude distribution of Fig. 2b there is no clear distinction between the noise and the multi-unit activity (see inset), a typical challenging feature of real extracellular recordings. The black trace in Fig. 2c shows the power density spectrum (PDS) of the data in Fig. 2a (filtered between 300 and 3000 Hz). For comparison, the curve in gray displays the non-filtered PDS. Both PDS traces were smoothed with a 3-datapoints moving average. The inset in Fig. 2c depicts the log–log PDS within the red vertical lines. The plot in the inset of Fig. 2c suggests that the PDS falls off following a power law distribution, with a $1/f^{\alpha}$, $\alpha = 1.1$ scaling dependence ($r^2 = 0.98$). Analyses performed on a large database of 192 channels of real recordings on the human medial temporal lobe, revealed similar PDS with $\alpha$ equal to $0.98 \pm 0.21$ (mean ± S.D.) and a mean $r^2$ coefficients equal to $0.992 \pm 0.007$ (see Supplementary Material).

From the full recording (of which the first 30 s are shown in Fig. 2), three different clusters were automatically identified using 'Wave_Clus' (see Section 2), as illustrated in Fig. 3. The first cluster was classified as a multi-unit and the other two as single-units. Note the high variance and low amplitude observed in the multi-unit (left) as compared to the single-unit clusters, a typical feature of extracellular recordings.

### 3.2. Simulation of the background noise activity

In this section we will describe the simulation of the background noise following the characteristics of the real recordings shown in the previous section in terms of the amplitude and frequency spectra. The noise activity was obtained by superimposing a large number of spikes, simulating the contribution of far-away neurons. For this, neurons were uniformly distributed inside a sphere and their spike amplitudes were inversely proportional to the distance to the centre of the sphere, which corresponds to the position of the electrode (see Section 2).

#### 3.2.1. Matching of the noise amplitude distribution

For the noise activity we did not consider point-source charges from relatively close distances $d_i$ because the resulting amplitude of such spikes are in the range of the multi-unit and single-unit activity. To have a rough approximation of the cutoff distance contributing to the noise, multi-unit, and single-unit activity, we evaluated the amplitude distribution of the simulated noise for three different cutoff distances ($d_i = 0.01$; 0.1 and 0.5, Fig. 4a–c, respectively) and compared these simulations with a real continuous recording that did not have any visible single- and multi-unit activity (Fig. 4d). As seen in Fig. 4a, a low cutoff distance gives a noise distribution with too many large amplitude spikes that clearly differs from the one of real recording. In fact, in this case there were $n = 285$ spikes crossing the detection threshold (see Eq. (1)), many more than the ones crossing the threshold for the real data ($n = 1$). On the other hand, considering only far apart action potentials (Fig. 4b and c) we reproduced a more realistic representation of the continuous data and amplitude distribution for the synthetic

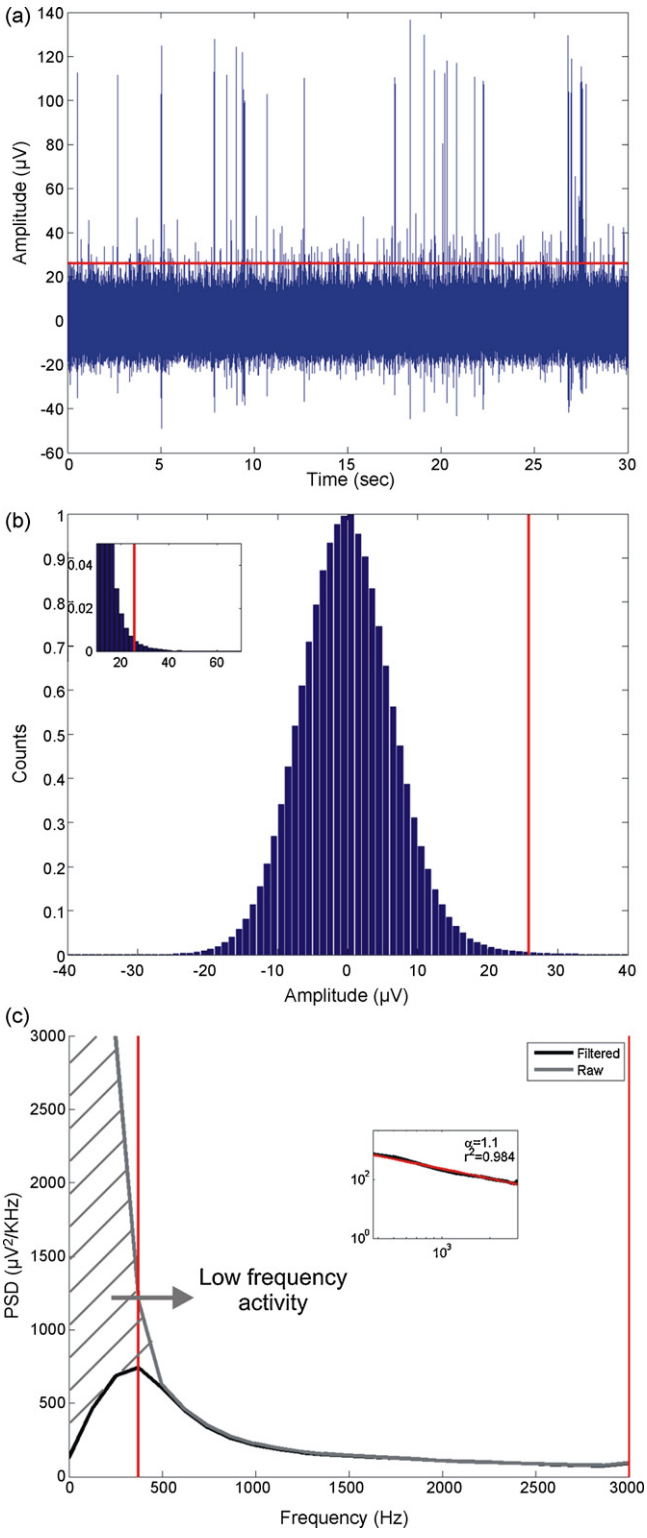


**Fig. 2.** (a) 30 s of a continuous recording from the human medial temporal lobe. The voltage signal was bandpass filtered between 300 and 3000 Hz. (b) Amplitude distribution of the voltage signal. The vertical red line represents the threshold for spike detection. The inset shows a zoom of the amplitude distribution around the detection threshold. (c) Power density spectrum (PDS) of the data in (a) with (black) and without (gray) bandpass filtering between 300 and 3000 Hz. The inset depicts the log–log PDS between the two red vertical lines. The red line in the inset indicates a linear fit to the data, which gave a slope $\alpha = 1.1$ and a correlation coefficient $r^2 = 0.984$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

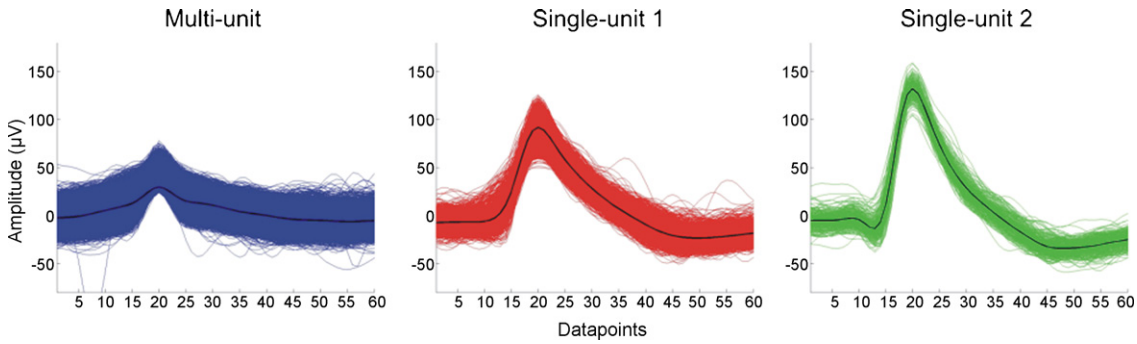


**Fig. 3.** Clusters obtained after an automatic spike sorting of the data of Fig. 2, using the whole duration of the recording (approximately 30 min). The cluster on the left corresponds to a multi-unit (in blue) and the other clusters correspond to two single-units (in red and green). For each cluster the thick black line marks the mean spike shape. Note the higher variance for the multi-unit compared to the single-units. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

background noise, with few events above the detection threshold (16 and 4, respectively). Given these results, we chose a cutoff distance of 0.5 for simulating the noise activity. The processing time for creating these simulations approximately 5 min using a standard PC with a frequency of 2.4 GHz.

### 3.2.2. Matching of the noise power spectra

As mentioned before, a simple alternative to simulate the background activity would be to use Gaussian noise, following the amplitude distribution discussed in the previous section. But as shown in Fig. 5a, after bandpass filtering between 300 and 3000 Hz (as done for the real recordings), the power spectrum looks flat and clearly different to the spectrum found in real recordings (see Fig. 2c). On the contrary, the procedure described in the previous sections, based on a realistic simulation of the biophysical process contributing to the noise activity, gives a frequency distribution very similar to the one obtained in real recordings, as shown in Fig. 2b. However, the value of the slope of a linear fit of the log–log plot of the power spectra, gave a value of $\alpha = 1.8$, with $r^2 = 0.94$. The lower amounts of high frequency components in the power spectrum compared with real data (Fig. 2c) might be due to other possible sources such as electronic noise or fast synaptic currents (Farrant et al., 1994), which could in principle be generated by an extra source of Gaussian noise. In order to generate a more realistic synthetic noise that includes higher frequency components to reproduce the power law decay found in real data, we decided to combine the two approaches described below, adding to the noise of Fig. 5b, in a proportion of 40% of Gaussian noise as the one of Fig. 5a. Fig. 5c shows the resulting noise characteristics, which leaded to an increase in the high frequency activity, giving a PDS decay of $1/f^{\alpha}$, with $\alpha \sim 1$. Note that the voltage signal and amplitude distributions remain the same in all cases.

## 3.3. Addition of the single-unit and multi-unit activity

Next, we added the single- and multi-unit spikes to the noise activity generated as described in the previous sections. Note that in this case, rather than assuming a biophysical generation process, as for the noise, we can use the amplitude and firing rate distribution of the real single-and multi-unit data. Table 1 describes the main features of 5 simulated datasets that reproduce different recording conditions to test spike sorting algorithms. The last two columns represent the scaling exponent\$\alpha$ and the correlation factor $r^2$ of the bandpass filtered data for each example, which are within the range of the values obtained from real extracellular recordings (see Supplementary Material).

Fig. 6 shows an example of one simulated data set with single-unit peak amplitude of to 4 and a firing rate of 1 Hz. Note in the continuous recording showed in Fig. 6a the presence of small multi-unit spikes at the level of the detection threshold, as also observed in the real recordings (see Fig. 2a). Moreover, in Fig. 6b there is a smooth transition between the background noise and the multi-unit activity around the detection threshold (see inset), as observed in the real recordings (see inset in Fig. 2b). Fig. 6c shows the PDS of this simulation, with a frequency dependence similar to the one of the real recording (Fig. 2c). The values of $\alpha$ (1) and $r^2$ (0.991) depicted in the figure (see inset) show that the simulation is able to reproduce major frequency features with a power law distribution, as with the real extracellular recording situation (see Fig. 2c and Supplementary Material).

## 3.4. Testing of the spike sorting algorithm

The results of the unsupervised spike sorting of the synthetic data of Fig. 6 using 'Wave_Clus' are shown in Fig. 7. As in the example of real data (Fig. 3), the algorithm correctly identified all units generated in the simulation (one multi-unit and two single-units). The multi-unit cluster (left plot) has a wider spike shape due to the contribution of several units, as typically found in real recordings (see leftmost spike in Fig. 3). Furthermore, for this cluster it is possible to see a non-uniform variance along the spike, which leads to non-Gaussian clusters found in real recordings. The presence of such non-Gaussian clusters are very challenging for many spike sorting algorithms that assume any particular spread of the clusters. The middle and right plots depict the single-unit clusters, in which the amplitude variability is solely given by the background noise activity.

Table 2 summarize the performance of the unsupervised spike sorting on 5 examples described in Table 1. The spike sorting performance (misses and false alarms) as well as the amplitude, firing rate and number of spikes, corresponds to the single-unit activity. The number of detected spikes for examples 1 and 5 were lower due

**Table 2**
Performance of the unsupervised spike sorting algorithm. All values are referred to single-unit spikes. Values in the spike sorting section of the table include cluster classification errors.

| Example number | Number of detected spikes | Spike detection | | Spike sorting | |
|---|---|---|---|---|---|
| | | Misses | False alarms | Misses | False alarms |
| 1 | 236 | 0 | 0 | 0 | 0 |
| 2 | 1058 | 0 | 0 | 0 | 0 |
| 3 | 1163 | 5 | 80 | 31 | 106 |
| 4 | 1127 | 10 | 87 | 10 | 87 |
| 5 | 118 | 0 | 0 | 0 | 0 |

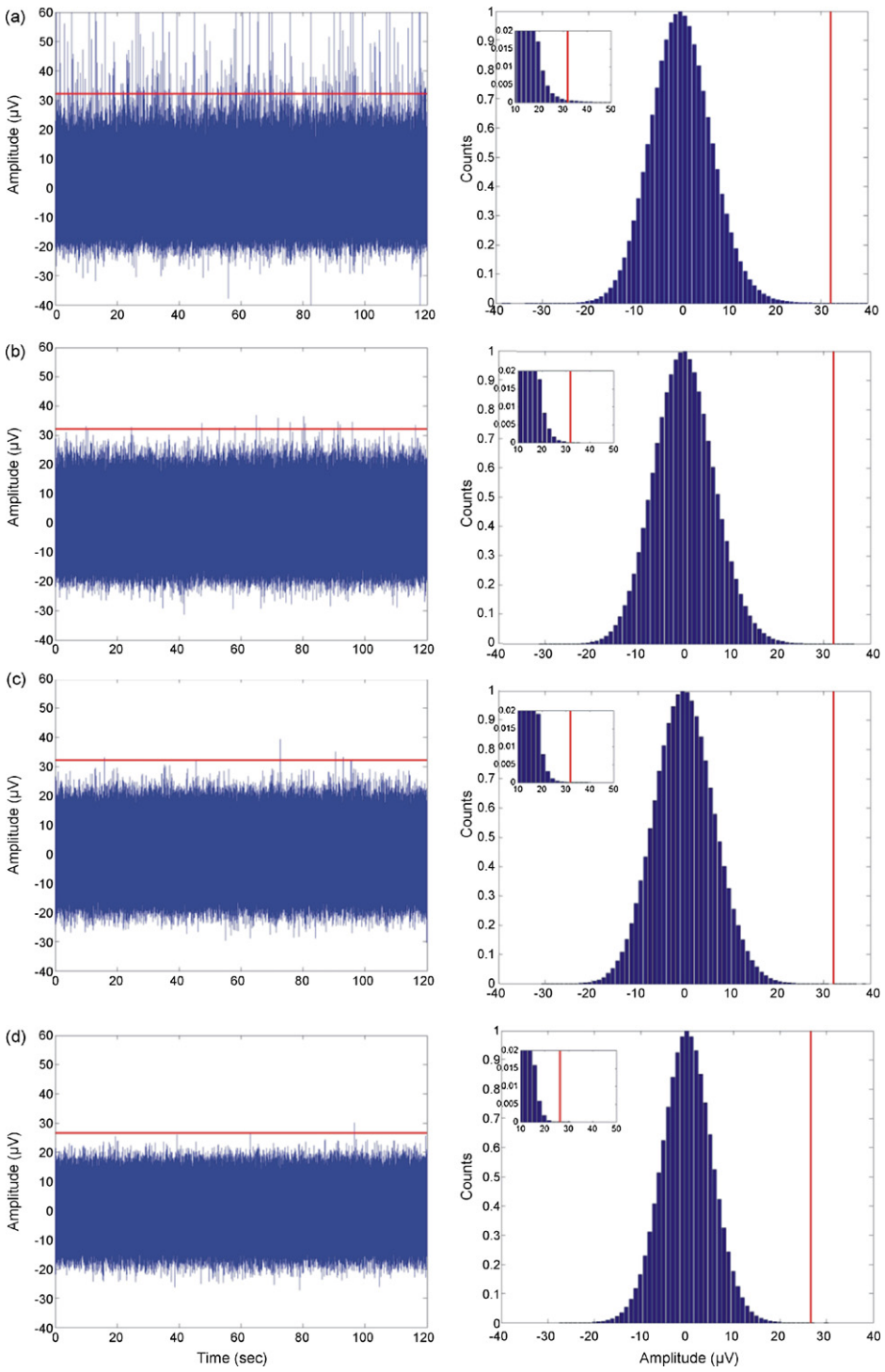


**Fig. 4.** (a–c) Voltage signal (left) and amplitude distribution (right) of 3 synthetic noise signals with different values of $d_i$ (from top to bottom, $d_i$ = 0.01, 0.1 and 0.5, respectively; see text for details). The red lines represents the detection threshold. (d) Noise from a real extracellular recording without spiking activity. For low values of $d_i$, (a) the amplitude distribution of the synthetic noise differs from the one observed in the real data. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

to the low single-unit firing rate for these simulations (see Table 2). In these two examples and in example 2 all spikes were correctly detected and classified. For simulations using single units of lower amplitudes, as in the cases of examples 3 and 4, the single-unit spikes tend to overlap with the multi-unit cluster, leading to some false alarms and misses. Note that in the case of the example 3, where the spike shapes of the two single-units were similar, the number of misses and false alarms is larger.

## 4. Discussion

The development of optimal spike detection and sorting algorithms play a crucial role for the analysis of extracellular recordings (Buzsaki, 2004; Quian Quiroga et al., 2005; Quian Quiroga and Panzeri, 2009). In recent years many spike sorting techniques have been proposed (Lewicki, 1994; Fee et al., 1996a; Harris et al., 2000; Letelier and Weber, 2000; Hulata et al., 2002; Pouzat et

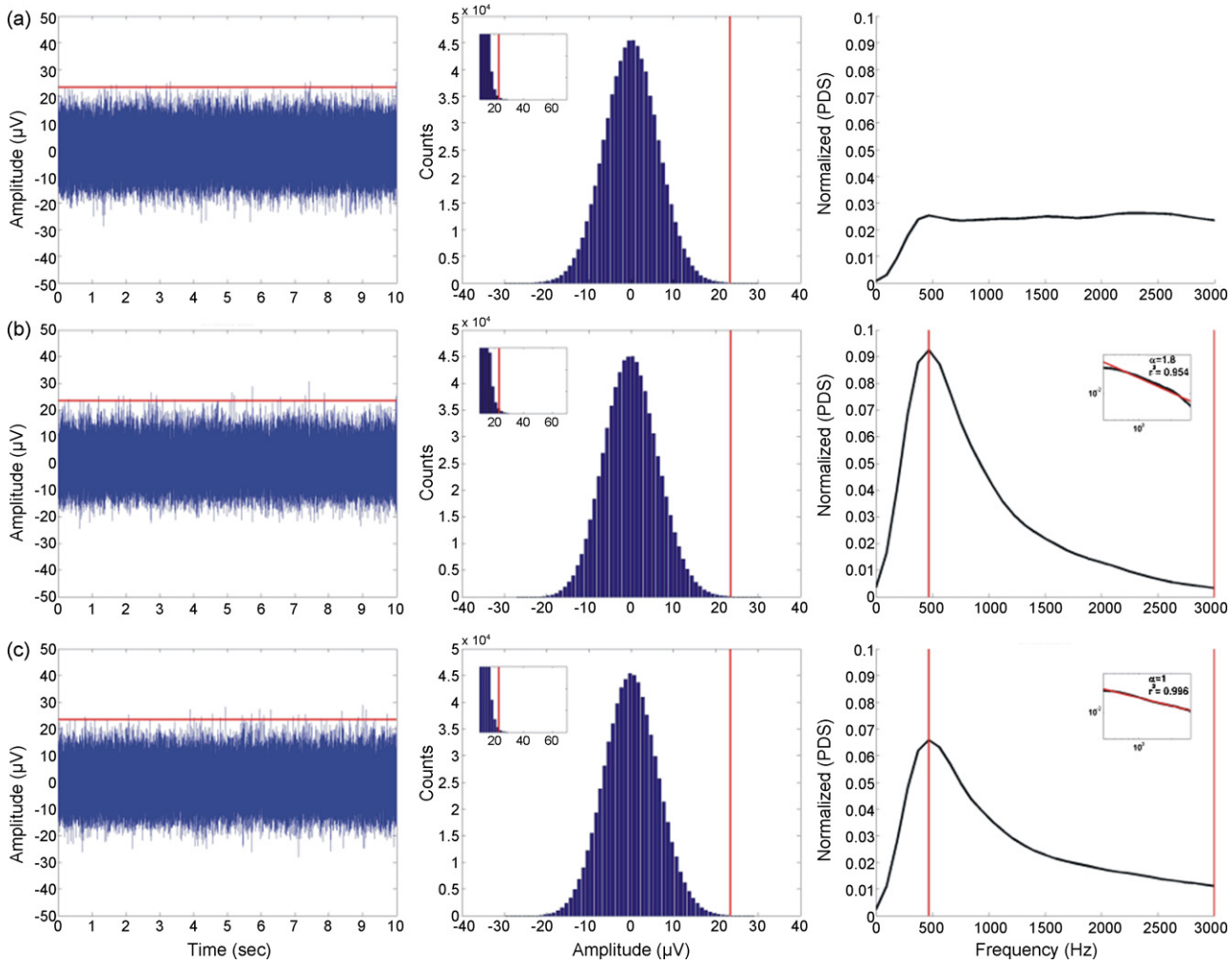


**Fig. 5.** Continuous data (left), amplitude distribution (centre) and power density spectrum (PDS, right) of 3 synthetic noise realizations of 120 s each (only the first 10 s are displayed). (a) Gaussian noise. (b) Noise generated by the superimposition of spikes (see text for details). (c) Combination of (a) and (b). Note that the amplitude and frequency characteristics of the noise in (b and c) are very similar to the one of the real recording in Fig. 2c.

al., 2002; Quian Quiroga et al., 2004), but it is still very difficult to compare these algorithms due to the lack of shared datasets that allow to unequivocally quantify their performance. On the one hand, it is in general not possible to perform such quantifications with real datasets because we do not have access to 'ground truth', i.e. the original labels marking which spike belongs to which neuron (but see Wehr et al., 1999; Harris et al., 2000; Henze et al., 2000 for notable exceptions). On the other hand, simulations of extracellular potentials allow an exact quantification of performance but, in general, they do not reproduce very challenging properties of real datasets, such as a power law spectrum of the noise and the presence of multi-unit activity and non-Gaussian clusters. In this respect, here we presented a hybrid and computationally simple approach to generate synthetic datasets reproducing these features of real recordings that are particularly challenging for any spike detection and sorting algorithm.

### 4.1. Hybrid simulation approach

A significant component of this work is the use of a hybrid twofold strategy to characterize most salient features of real data relevant for an efficient performance quantification of different algorithms. The first component of our approach was to simulate the background noise by adding the contribution of many different far-away action potential point-sources uniformly distributed around the electrode, where the amplitude of each of these action potentials was determined by its distance from the electrode following Coulomb's law. This simple and physiologically plausible model reproduced the main characteristics of the noise activity in real recordings, avoiding the need of simulating computationally expensive details of each of the far-away neurons. Note that this strategy would not be plausible for close-by neurons – which generate the multi-unit and single-unit activity – since in this case the particular physiological characteristics of the neurons and their relative positions and orientations with respect to the recording site become relevant (Rall, 1962; Hines and Carnevale, 1997; Holt and Koch, 1999; Gold et al., 2006; Harris et al., 2000). The second component of our approach was to generate the single-unit and multi-unit activity by simply using template spike shapes and the peak amplitude distributions available for a large number of experimental observations.

This simple twofold strategy to generate the noise and the spiking activity allowed the performance of relatively fast simulations – about 9 min for a simulation of a 2-min recording using an Intel Core 2 PC with a clock frequency of 2.4 GHz with Matlab – which reproduced the main features of extracellular recordings, as detailed below.

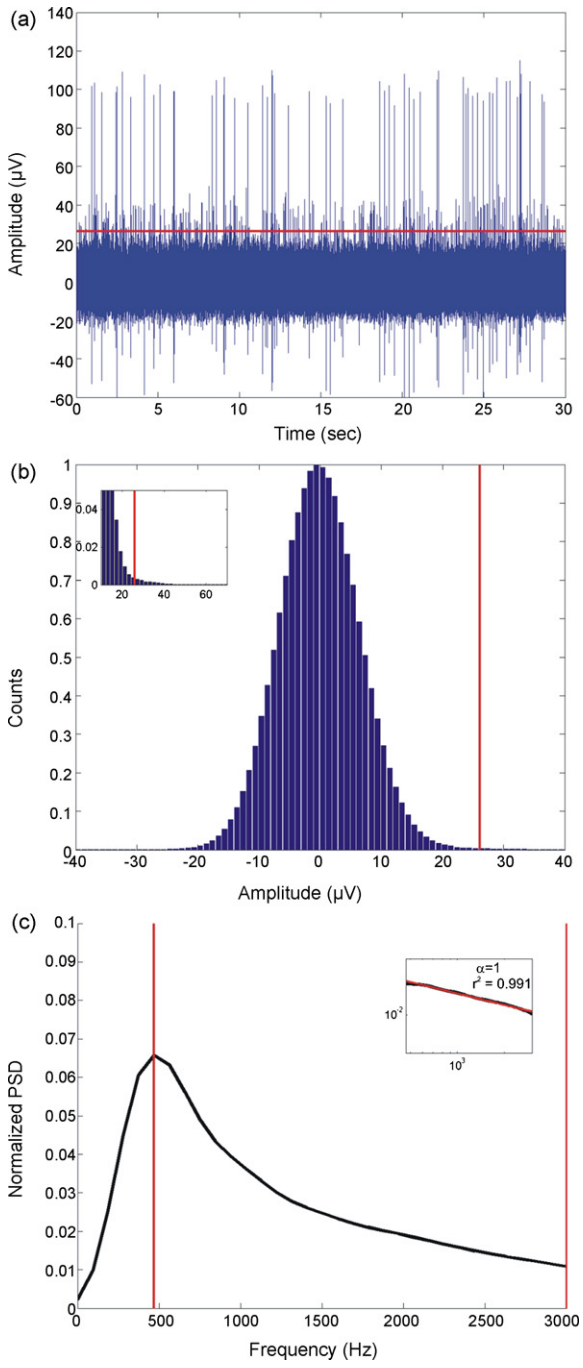


**Fig. 6.** Example of a 2-min simulation of an extracellular recording including background noise, multi-unit activity and two single-units (example 1 in Table 1). (a) First 30 s of continuous data filtered between 300 and 3000 Hz. (b) Amplitude distribution of the voltage signal. The red lines represents the amplitude detection thresholds. The inset shows a zoom around the detection threshold. (c) Power density spectrum (PDS) of the continuous data from 300 to 3000 Hz. The inset depicts the log–log PDS between the two red vertical lines. The red line in the inset shows a linear fit to the data, which gave a slope $\alpha = 1$ and a correlation coefficient $r^2 = 0.991$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

### 4.2. Multi-unit activity

The presence of multi-unit activity in extracellular recordings makes the processes of spike detection and sorting particularly challenging, an issue that has not been dealt with in previous simulation studies (Lewicki, 1994; Quian Quiroga et al., 2004; Smith and Mtetwa, 2007). This is mainly for two reasons. First, the low signal-to-noise ratio and high amplitude variance of multi-unit clusters complicates the detection of spikes because there is no clear separation between background noise and spikes. Second, multi-unit clusters are usually non-Gaussian, which compromises the use of spike sorting algorithms that assume Gaussian clusters or introduce a pre-whitening of the data (Pouzat et al., 2002; Lewicki, 1998). Third, multi-unit clusters can be relatively large compared to those of single-units, given that they are composed by the activity of several units (in spite of the fact that a large proportion of spikes of these units may be not detected). This can complicate the accurate sorting of single-units and even their identification, since they can be easily merged within the large multi-unit clusters. This issue becomes more problematic for the identification of sparse firing neurons (Quian Quiroga et al., 2005, 2007).

### 4.3. Background noise

The characteristics of the background noise recorded from *in vivo* extracellular recordings play a fundamental role in the identification and classification of single- and multi-unit activity. In this respect our noise simulations, generated by superimposing a large number of spikes from far-away neurons, reproduced the amplitude and power spectra distribution of the real recordings. In this study we did not systematically study up to which distance neurons contribute to the single- and multi-unit activity and beyond that to the background noise. Rather, we used a heuristic minimum distance for the noise contribution that could reproduce the amplitude distribution observed in real recordings and used typical amplitude values observed in real data for the multi- and single-units. For the noise generation, decreasing this heuristic distance gave a large number of spikes crossing the detection threshold, which were not present in the real data (see Fig. 4). A more precise and systematic investigation of the distance ranges giving rise to the noise, multi-and single-unit activity is ripe of future investigation, but it would have also to consider other factors, such as the impedance profile of the electrodes and more detailed characteristics of the neurons nearby the recording site. These results could be then compared to those reported with real recordings, in which acutely implanted electrodes are retracted observing at what distance the spikes of a single neuron gets mixed with those of other neurons – in between 50 and 100 μm – and at what distance the spikes of the neurons are no longer detected – at about more than 150 μm (Gerstein and Clark, 1964; Henze et al., 2000; Buzsaki, 2004).

It has to be noted that with this approach we did not try to simulate the high frequency components generated by single-cell currents (Llinas, 1988; Farrant et al., 1994) or by relatively lower electronic noise from the recording system (Fee et al., 1996b). To replicate the power law decay in the frequency spectrum observed in real data, the high frequency noise was introduced by superimposing Gaussian noise to the background activity generated by the contribution of the spikes from far-away neurons. Moreover it was not necessary to model the contribution of LFPs, since their activity is outside the frequency range considered for spike sorting, which is in fact typically filtered to perform the spike detection.

Alternatively to our approach, it is in principle possible to obtain noise signals from real recordings by eliminating the spiking activity (Pouzat et al., 2002). Then, identified spikes corresponding to different neurons can be added to the noise to evaluate the performance of spike sorting algorithms. However, the difficulty of such an approach is to eliminate the spiking activity because this depends on the setting of a detection threshold which will certainly leave some of the (low amplitude) spikes and delete part of the noise activity due to random crossings of the threshold. To overcome these problems, one could still use real recordings without

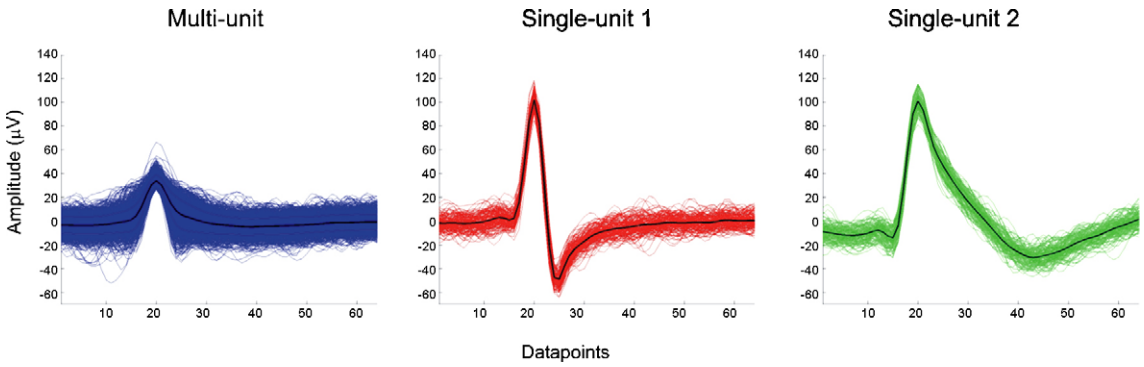


**Fig. 7.** Result of the automatic clustering of the recording shown in the previous figure. The algorithm correctly identified one multi-unit (left) and two single-units. Note the similarity with to the real recording shown in Fig. 2.

spiking activity, as we used in Fig. 4d to compare the amplitude distributions of our noise simulations. However, the availability of such data is not guaranteed, especially if one requires a large number of simulations.

The noise model presented here did not assume any particular distribution of data as when using white or colored Gaussian noise in which clusters have hyper-spheres or ellipse boundaries (Chandra and Optican, 1997). It has to be noted that several experimental conditions, such as electrode drifts, bursting activity, overlapping spikes, multi-unit activity or misalignments during detection can produce non-Gaussian clusters (Quian Quiroga, 2007), which compromise the use of algorithms assuming clusters with a multidimensional Gaussian shape (Lewicki, 1994; Pouzat et al., 2002).

In conclusion, we presented a method to generate simulations that reproduced the most salient characteristics of extracellular recordings like, power spectrum and amplitude distribution, which can be useful for testing spike detection and sorting algorithms. Moreover, these simulations might also be used to quantify filter distortions in the spike shapes (Quian Quiroga, 2009), to simulate recording conditions with different electrode designs (Robinson, 1968) or as a platform to test effects of spike detection and sorting for inferring information about, for example, time patterns and synchronized activity in a neural population (Quian Quiroga and Panzeri, 2009). Our hybrid approach to generate the synthetic data was able to provide several simplifications making this algorithm simple and fast.

## Acknowledgements

This work was supported by EPSRC (CARMEN: EP/E002331/1; and EP/D052254/1), MRC and the Royal Society.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jneumeth.2009.08.017.

## References

Bedard C, Kroger H, Destexhe A. Model of low-pass filtering of local field potentials in brain tissue. Phys Rev E 2006;73:051911.
Blatt M, Wiseman S, Domany E. Super-paramagnetic clustering of data. Phys Rev Lett 1996;76:3251–4.
Brunel N, Wang XJ. What determines the frequency of fast network oscillations with irregular neural discharges? J Neurophysiol 2003;90:415–30.
Buzsaki G. Large-scale recording of neuronal ensembles. Nat Neurosci 2004;7:446–51.
Chandra R, Optican LM. Detection, classification, and superposition resolution of action potentials in multi-unit single-channel recordings by an on-line real-time neural network. IEEE Trans Biomed Eng 1997;44:403–12.
Farrant M, Feldmeyer D, Takahashi T, Cull-Candy SG. NMDA-receptor channel diversity in the developing cerebellum. Nat Lond 1994;368:335–9.
Fee MS, Mitra PP, Kleinfeld D. Automatic sorting of multiple unit neuronal signals in the presence of anisotropic and non-Gaussian variability. J Neurosci Methods 1996a;69:175–88.
Fee MS, Mitra PP, Kleinfeld D. Variability of extracellular spike waveforms of cortical neurons. J Neurophysiol 1996b;76:3823–33.
Fried I, MacDonald KA, Wilson CL. Single neuron activity in human hippocampus and amygdala during recognition of faces and objects. Neuron 1997;18:753–65.
Gerstein GL, Clark WA. Simultaneous studies of firing patterns in several neurons. Science 1964;143:1325–7.
Gold C, Henze DA, Koch C, Buzsaki G. On the origin of the extracellular action potential waveform: a modeling study. J Neurophysiol 2006;95:3113–28.
Hahnloser RHR, Kozhevnikov AA, Fee MS. An ultra-sparse code underlies the generation of neural sequences in a songbird. Nature 2002;419:65–70.
Harris KD, Henze DA, Csicsvari J, Hirase H, Buzsaki G. Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements. J Neurophysiol 2000;84:401–14.
Henze DA, Borhegyi Z, Csicsvari J, Mamiya A, Harris KD, Buzsáki G. Intracellular features predicted by extracellular recordings in the hippocampus in vivo. J Neurophysiol 2000;84:390–400.
Hines ML, Carnevale NT. The neuron simulation environment. Neural Comp 1997;9:1179–209.
Holt G, Koch C. Electrical interactions via extracellular potential near cell bodies. J Comput Neurosci 1999;6:169–84.
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.
Letelier JC, Weber PP. Spike sorting based on discrete wavelet transform coefficients. J Neurosci Methods 2000;101:93–106.
Lewicki M. A review of methods for spike sorting: the detection and classification of neural action potentials. Network: Comput Neural Syst 1998;9:R53–78.
Lewicki M. Bayesian modeling and classification of neural signals. Neural Comput 1994;6:1005–30.
Llinas RR. The intrinsic electrophysiological properties of mammalian neurons: insights into central nervous system function. Sci Wash 1988;242:1654–64.
Mazzoni A, Panzeri S, Logothetis NK, Brunel N. Encoding of naturalistic stimuli by local field potential spectra in networks of excitatory and inhibitory neurons. PloS Comput Biol 2008;4:1–20.
Perez-Orive J, Mazor O, Turner GC, Cassenaer S, Wilson RI, Laurent G. Oscillations and sparsening of odor representations in the mushroom body. Science 2002;297:359–65.
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.
Quian Quiroga R, Nadasdy Z, Ben-Shaul Y. Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. Neural Comput 2004;16:1661–87.
Quian Quiroga R, Reddy L, Kreiman G, Koch C, Fried I. Invariant visual representation by single neurons in the human brain. Nature 2005;435:1102–7.
Quian Quiroga R. Spike sorting. Scholarpedia 2007;2(12):3583.
Quian Quiroga R, Reddy L, Koch C, Fried I. Decoding visual inputs from multiple neurons in the human temporal lobe. J Neurophysiol 2007;98:1997–2007.
Quian Quiroga R. What is the real shape of extracellular spikes? J Neurosci Methods 2009;177:194–8.
Quian Quiroga R, Panzeri S. Extracting information from neuronal populations: information theory and decoding approaches. Nat Rev Neurosci 2009;10:173–85.
Rall W. Electrophysiology of a dendritic neuron model. Biophys J 1962;2:145–67.
Robinson DA. The electrical properties of metal microelectrodes. IEEE Trans Biomed Eng 1968;56:1065–71.
Smith LS, Mtetwa N. A tool for synthesizing spike trains with realistic interference. J Neurosci Methods 2007;159:170–80.
Wehr M, Pezaris J, Sahani M. Simultaneous paired intracellular and tetrode recordings for evaluating the performance of spike sorting algorithms. Neurocomputing 1999;26:1061–8.