LETTER **Communicated by Gaute Einevoll**

A Detailed and Fast Model of Extracellular Recordings

Luis A. Camuñas-Mesa
lacm1@le.ac.uk
Centre for Systems Neuroscience, University of Leicester, Leicester LE1 7RH, U.K.

Rodrigo Quian Quiroga
rqqg1@le.ac.uk
Centre for Systems Neuroscience, University of Leicester, Leicester LE1 7RH, U.K.; Leibniz Institute for Neurobiology, University of Magdeburg, 39118 Magdeburg, Germany

We present a novel method to generate realistic simulations of extracellular recordings. The simulations were obtained by superimposing the activity of neurons placed randomly in a cube of brain tissue. Detailed models of individual neurons were used to reproduce the extracellular action potentials of close-by neurons. To reduce the computational load, the contributions of neurons further away were simulated using previously recorded spikes with their amplitude normalized by the distance to the recording electrode. For making the simulations more realistic, we also considered a model of a finite-size electrode by averaging the potential along the electrode surface and modeling the electrode-tissue interface with a capacitive filter. This model allowed studying the effect of the electrode diameter on the quality of the recordings and how it affects the number of identified neurons after spike sorting. Given that not all neurons are active at a time, we also generated simulations with different ratios of active neurons and estimated the ratio that matches the signal-to-noise values observed in real data. Finally, we used the model to simulate tetrode recordings.

1 Introduction

A classic experimental approach in neuroscience is to study the spiking activity of individual neurons using microelectrodes inserted in the brain. The first step in the analysis of the data recorded with these electrodes is to detect the spikes of nearby cells and classify their shapes into clusters corresponding to the firing of the different neurons, a procedure known as spike sorting (Lewicki, 1998; Quian Quiroga, 2007, 2012a; Gibson, Judy, & Markovi, 2012).

A supplemental appendix is available online at http://www.mitpressjournals.org/doi/suppl/10.1162/NECO_a_00433.

Spike sorting has been a challenging problem for decades, and several software packages implementing different detection and sorting strategies have been developed (Lewicki, 1998; Harris, Henze, Csicsvari, Hirase, & Buzsaki, 2000; Quian Quiroga, Nadasdy, & Ben-Shaul, 2004; Rutishauser, Schuman, & Mamelak, 2006; Kwon, Eldawlatly, & Oweiss, 2012). The end result of all these methods is the identification of the firing of individual neurons, which, depending on the method used, the quality of the data, and the expertise of the operator, could potentially differ dramatically from the actual real values (Pedreira, Martinez, Ison, & Quian Quiroga, 2012).

Although methods to evaluate the quality of spike sorting have been developed (i.e., different metrics quantifying the distance between the clusters of spike shapes; Schmitzer-Torbert, Jackson, Henze, Harris, & Redish, 2005; Hill, Mehta, & Kleinfeld, 2011), there is typically no objective measure of spike sorting performance with real data due to the lack of ground truth (Einevoll, Franke, Hagen, Pouzat, & Harris, 2011). Different strategies have been proposed to overcome this problem. One approach is to perform simultaneous extracellular and intracellular recordings using the intracellular recordings to validate the spike sorting performed on the extracellular data (Wehr, Pezaris, & Sahani, 1999; Harris et al., 2000). However, this type of recording is rare, and the number of clusters that can be validated is limited. This is because we have access to the ground truth for only one neuron per intracellular recording and it is experimentally challenging to record intracellularly from more than one or a couple of neurons at a time. Another approach is to use synthetic data, where the time and identity of each spike are known by construction. One way of doing this is to use already collected extracellular action potentials (EAPs), which are added to computer-generated noise (Lewicki, 1994). Simulations performed in this way are relatively simple and fast, but the challenge here is to reproduce in the simulations the features observed in real data, particularly those that are critical for spike sorting algorithms, such as the amplitude and power spectrum distribution of the noise and the amplitude distribution of the spikes (Martinez, Pedreira, Ison, & Quian Quiroga, 2009). A more realistic scenario is to sum up the contributions of individual neurons, using models of the EAPs generated by each neuron at a given location, based on their morphology and biophysical properties. Detailed models of EAPs of single neurons have been developed (Holt & Koch, 1999; Gold, Henze, Koch, & Buzsaki, 2006; Gold, Henze, & Koch, 2007; Smith & Mtetwa, 2007; Pettersen, Hagen, & Einevoll, 2008; Pettersen & Einevoll, 2008), and some preliminary synthetic data sets have been generated using this strategy (Menne, Folkers, Malina, Maex, & Hofmann, 2002; Mamlouk, Sharp, Menne, Hofmann, & Martinetz, 2005). However, due to the large computational power needed to run these models for each neuron, this approach has been limited to mimic the activity generated by relatively small networks.

As a compromise between detailed modeling and computational power, in this letter we present a method to simulate extracellular potentials

recorded by a finite-size electrode using a hybrid strategy. For this, we implemented detailed compartmental models to simulate the contribution of neurons near the recording electrode, while for farther-away neurons, we used previously recorded spike shapes, with an amplitude normalized by the distance to the electrode. The parameterization of this model allowed adjusting the simulations to mimic the variability observed on different recording conditions (e.g., in terms of signal-to-noise ratio, size of the electrode, neuronal density). Moreover, we used this approach to simulate tetrode recordings.

2 Materials and Methods

2.1 Basic Features of the Model. Figure 1 shows an illustration of our modeling approach. We considered a cube of brain tissue containing a large number of randomly placed neurons, with the recording electrode at its center (see Figure 1a). As described in previous studies (Markram et al., 2004), we considered a ratio of 80% pyramidal cells and 20% interneurons. The firing of each neuron followed a Poisson distribution, with a mean firing rate assigned according to the distribution of mean firing rates observed in real recordings from the medial temporal lobe (exponential distribution with a mean of 1.25 Hz for pyramidal cells and five times higher for interneurons) (Quian Quiroga, Reddy, Koch, & Fried, 2007; Ison et al, 2011). The real data correspond to a total of 624 recordings of about half an hour, performed in 13 experimental sessions with 48 channels each in two patients implanted with intracranial recordings for clinical reasons.

The activity of the neurons inside the cube was modeled using: i) a detailed biophysics model, for neurons near the recording electrode (see Figures 1c and 1d; Gold et al., 2007) and ii) previously recorded spike shapes with an amplitude normalized by the distance to the electrode for neurons farther away (see Figure 1e). Hence, the extracellular recording was obtained by adding the spiking activity generated by all the neurons at the electrode position, the closest giving the single- and multi-unit activity and the ones farther away contributing to the background noise. Given that the exposed area of the electrode (the electrode tip) is not a single point source, we also modeled the effects of finite-size electrodes by averaging across their surface and considering capacitance filtering effects. For this, we implemented a previous model of the electrode-tissue interface (see Figure 1b; Robinson, 1968) and also studied how the recordings and spike sorting outcomes changed with different electrode sizes. We deleted neurons whose cell body was placed inside a sphere with a diameter 20 μm larger than the one of the electrode centered at the electrode position, considering that it would have been damaged by the electrode. In the following sections, we describe in detail the generation of the background noise, the unit activity, and the electrode model.

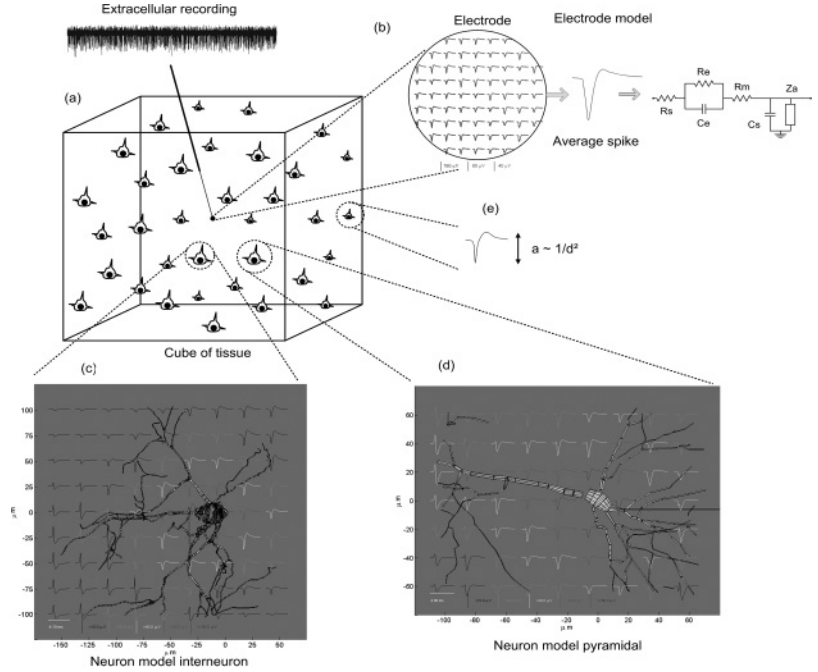

Figure 1: Illustration of the modeling approach. (a) Neurons inside the cube of brain tissue were modeled to obtain the spike shapes generated at the electrode position. The extracellular recording was created by adding the spiking activity of all the neurons, averaging over the electrode surface and filtering with the (b) Electrode model, where R_s is the spreading resistance, R_e is the leakage resistance, C_e is the capacitance of the interface, R_m is the metallic resistance, C_s is the shunt capacitance, and Z_a is the input impedance of the amplifier. A detailed model was used for neurons close to the electrode, either interneurons (c) or pyramidal (d), while previously recorded spike shapes were used for neurons farther away, with an amplitude normalized by the distance to the electrode (e).

2.2 Background Noise. The background noise was generated by adding the contribution of two sources: neural and thermal noise. The neural noise is produced by the spiking activity of distant neurons, and its level depends on the volume of brain tissue that is simulated. If the volume is too small, the noise level will be small compared to the one observed in real data; if the volume is too large, the processing time and the computational power need will increase without any relevant changes of the results. In other words, if the volume is too large, we would simulate neurons that do not contribute to the recording at all. Figure 2 shows the noise levels obtained (as quantified by their standard deviation) with simulations generated with different volumes of tissue, where we considered a density of 300,000 neurons/mm^3

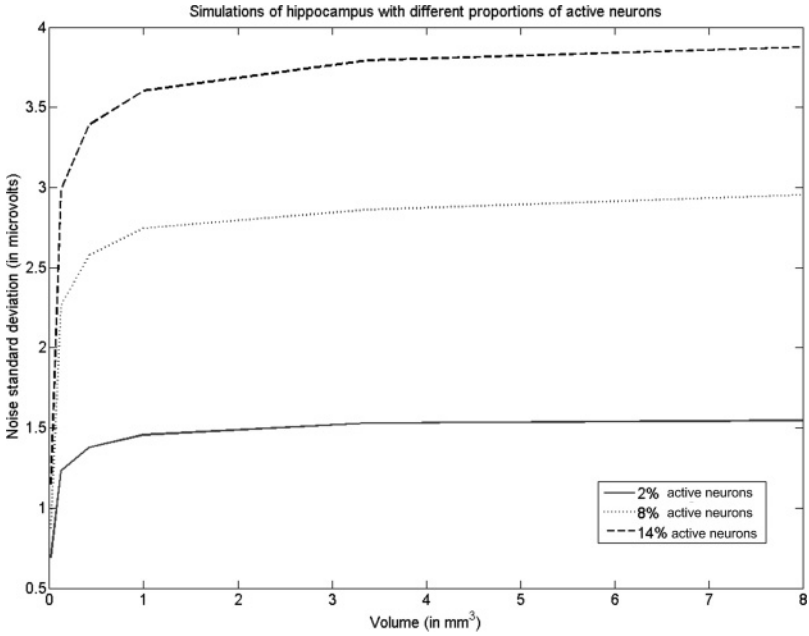

Figure 2: Effect of the volume of tissue on the noise level for different proportions of active neurons. In all cases, a volume of 1 mm^3 of tissue gives a noise level close to the saturation value.

as reported in the rodent hippocampus (Henze et al., 2000). Due to the fact that not all neurons are active at a given time, we also considered three proportions of active neurons, 2%, 8%, and 14%, which correspond to a low, medium, and high level of overall neural activity (Shoham, O'Connor, & Segev, 2006). For all ratios of active neurons, Figure 2 shows a saturation of the noise level at about 1 mm^3, where the standard deviation of noise had more than 92% of the value obtained for the largest volume. Therefore, we chose a volume of 1 mm^3 for all the simulations described below.

The thermal noise includes other possible sources of electrical activity that might contribute to the noise power spectrum, like electronic noise, axons, dendrites, and synaptic currents (Llinas, 1988; Farrant, Feldemeyer, Takahashi, & Cull-Candy, 1994; Martinez et al., 2009). To simulate this effect, we added gaussian noise with zero mean and a standard deviation of $1\mu V$. As in a previous work (Martinez et al., 2009), the sum of the thermal and neural noise replicated the frequency and amplitude distributions observed in real data.

2.3 Model of Extracellular Action Potentials. It has been shown that the neuropil can be modeled by an isotropic volume conductor and that

the electric potential Φ in the extracellular space is governed by Laplace's equation (Plonsey, 1969),

$$\nabla^2 \Phi = 0. \tag{2.1}$$

For a single-point source of amplitude I in an unbounded isotropic volume conductor, the solution of Laplace's equation is analogous to the classical physics problem of point charges in free space (Coulomb's law):

$$\Phi = \rho I/(4\pi r), \tag{2.2}$$

where ρ is the extracellular resistivity and r is the distance from the source (Gold et al., 2006). In real neurons, membrane currents are distributed over elongated cylindrical neurites whose length largely exceeds the width. The line source approximation (LSA) (Holt & Koch, 1999) makes the simplification of locating the transmembrane net current for each neurite on a line across its center. Then, by assuming a line distribution of current, the resulting potential in a single neurite k with length Δs is given by

$$\Phi_k(r,h) = \frac{\rho}{4\pi\,\Delta s}\int_{-\Delta s}^{0} \frac{Ids}{\sqrt{r^2 + \left(h-s\right)^2}}, \tag{2.3}$$

where r is the radial distance from the line, h is the longitudinal distance from the end of the line, and s is the longitudinal dimension (Gold et al., 2006). The potential generated by the neuron is then built up from a compartmental model consisting of a group of cylindrical segments. Thus, the extracellular action potential for neuron j when firing spike i is $\widetilde{\varphi}_i^j = \sum_k \Phi_k$.

We used compartmental neuron models to simulate the shape of the extracellular action potentials at the electrode position (Hines & Carnevale, 1997). These models describe the cellular morphology (soma, axons, and dendritic tree) and include the variety of ionic channels and corresponding conductance densities (Gold et al., 2006). In particular, we used the same four sets of parameters—determining the passive membrane behavior, current kinetics, and conductance density for the various channels—described in Gold et al. (2007), combined with five different neuron morphologies, which gave 20 types of neurons (see the online supplementary material). Of the five morphologies used, four corresponded to pyramidal cells and one to an interneuron (Destexhe, Contreras, Steriade, Sejnowski, & Huguenard, 1996; Contreras, Destexhe, & Steriade, 1997; Rudolph, Pelletier, Pare, & Destexhe, 2005). Simulated synaptic inputs were applied to depolarize the cells until action potentials were triggered. The propagation medium was considered to be purely ohmic, with negligible capacitive effects (Logothetis, Kayser, & Oeltermann, 2007). Figures 1c and 1d show examples

of the simulation of the action potentials generated by a pyramidal cell and an interneuron at different positions within a plane. For each single neuron, one spike shape was simulated and repeated at each spike time.

2.3.1 A Computational Simplification. The model described can accurately simulate the spike shapes of a neuron at different locations. However, the computational load when taking into account the many neurons that contribute to an extracellular recording is way too large. If we consider a 1 mm^3 cube of hippocampus with 8% active neurons, this gives 24,000 neurons to be simulated. In a standard Intel Core i5 PC with a clock frequency of 3.33 GHz, the mean processing time to generate each spike shape is around 3 seconds, meaning that around 20 hours would be needed to calculate the spike shapes for all the individual neurons inside the cube and about 57 days to simulate the recording with a finite-size electrode as the one in Figure 1b. We also have to add the time needed to build up the extracellular recording by adding up the contribution of each of these neurons, but the time taken by this process is much lower than the one to calculate the spike shapes of all neurons at a given location.

In order to reduce the processing time, we considered the fact that the exact shapes of the spikes from far-away neurons do not need to be modeled in detail, given that they contribute only to the background noise. Therefore, we defined two zones: zone A, with close-by neurons ($d \leqslant d_{lim}$) that were simulated using the compartmental model described above, and zone B, for farther-away neurons ($d > d_{lim}$) that were randomly assigned a spike shape taken from a database of 594 previously recorded neurons, with an amplitude normalized by the distance to the electrode.

To estimate the distance d_{lim} separating both zones, we simulated each of the 20 neuron models used (5 morphologies $\times$ 4 sets of parameters) and calculated the spike shapes at increasing distances along the six different Cartesian directions from the soma ($\pm x$, $\pm y$, and $\pm z$), assuming a point source electrode. Therefore, we estimated how the spikes changed as a function of the distance to the soma, averaging across the 20 models and 6 directions. In particular, we calculated the average amplitude of the simulated spikes (see Figure 3a) and the variability of the spike shapes (estimated as the mean Euclidean distance between the spike shapes at contiguous positions; see Figure 3b). From these results, we chose a value $d_{lim} = 150$ μm, given that at this distance, the spike amplitude reached an asymptotic value of 20 μV (well into the noise level) and the spike shape variability was negligible. For this value of d_{lim}, only the neurons located in a sphere representing 1.4% of the total volume should be simulated in detail, which reduces the processing time by about 80 times. Note that the value $d_{lim} = 150$ μm is well into the asymptotic regime; indeed, a value of 100 μm would also have been appropriate, and it would have given a 240-fold reduction compared to modeling each neuron.

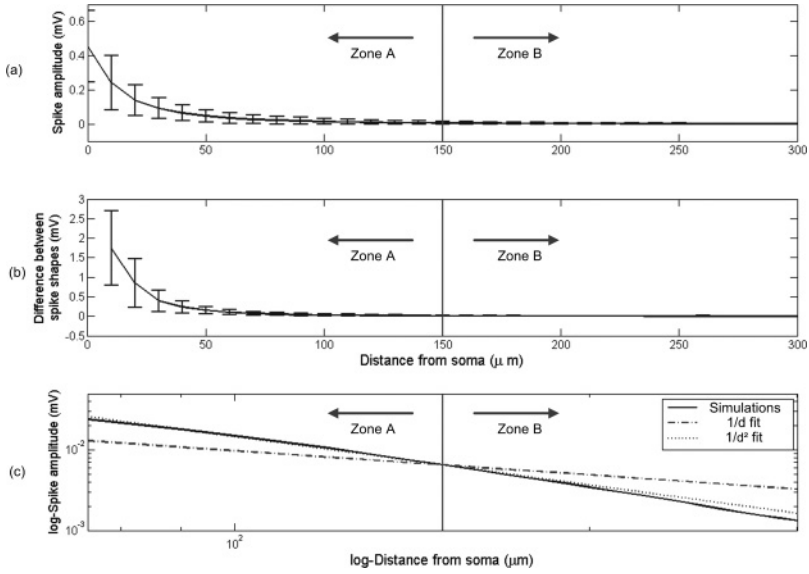

Figure 3: Estimation of the distance d_{lim} separating zones A and B. Simulations were done using the 20 neuron models (5 morphologies $\times$ 4 sets of parameters; see text for details) by calculating the spike shapes at increasing distances from the soma along the six Cartesian directions. (a) Average amplitude of the simulated spikes as a function of the distance to the soma. (b) Variability of the spike shapes estimated as the mean Euclidean distance between the spike shapes at contiguous positions. From these results, we choose a value $d_{lim} = 150$ μm. (c) Comparison of the mean amplitude of the simulated spike shapes in zone B (large distance) with two approximations, $1/d$ and $1/d^2$, in log-log scale. This graph shows that the best approximation is given by $1/d^2$.

After setting the value of d_{lim}, the amplitudes of the spike shapes in zone B had to be scaled considering the dipole approximation, which, following Coulomb's law, gives a $1/d^2$ decay with distance (Pettersen & Einevoll, 2008; Linden et al., 2011). Figure 3c shows the comparison between the simulated amplitudes and two different approximations at large distances: taking the neuron as a monopole or a dipole ($1/d$ and $1/d^2$ decay with distance, respectively). As expected, the dipole model gave a better fit to the data.

Summarizing the previous considerations, the contribution to the extracellular potential of neuron j when firing spike i is given by:

$$\phi_i^j = \begin{cases} \sum_k \Phi_k, & d \leqslant d_{lim} \\ \dfrac{1}{d^2}\bar{\phi}, & d > d_{lim} \end{cases}, \tag{2.4}$$

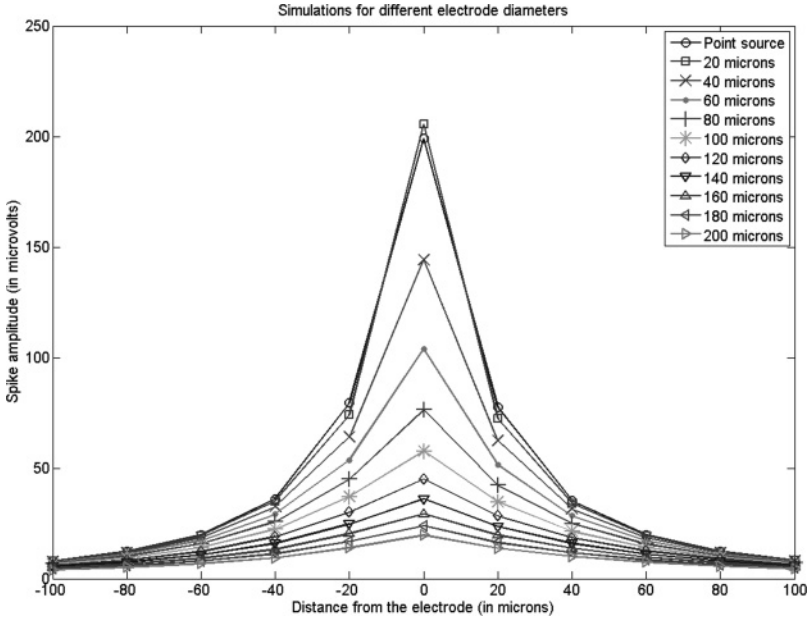

Figure 4: Simulation of the amplitude of the spike shapes of a single neuron as a function of the distance to the soma for different electrode sizes. The electrodes with smaller diameters give the larger spikes at the soma. For each diameter, the spike amplitude decays with the distance.

where k are the different neurites for neuron j and $\bar{\phi}$ is a randomly selected average spike shape from the database of 594 previously recorded neurons (Quian Quiroga et al., 2004; Martinez et al., 2009).

2.4 Electrode Model. To make our simulations more realistic, we considered a finite-size electrode model. For this, we assumed a circular area of the electrode tip (the exposed area of the electrode that records the extracellular potential) and averaged the spike shapes across a grid of points separated by 5 μm along its surface (see Figure 1b). Figure 4 shows the amplitude of the spike shapes of a single (simulated) neuron as a function of the distance to the soma for different electrode sizes. The electrodes with smaller diameters give the larger spikes at the soma, and, as expected, for each diameter, the spike amplitude decays with the distance.

As in previous work (Robinson, 1968), we also modeled the interface between the electrode and the extracellular space with a capacitive filter. The equivalent circuit of this filter, shown in Figure 1b, includes the following components: spreading resistance of the saline bath (R_s), leakage resistance (R_e), capacitance of the electrode-tissue interface (C_e), metallic

resistance (R_m), shunt capacitance to ground (C_s), and input impedance of the amplifier (Z_a). Basically, the effective electrode impedance is the sum of impedances due to the resistance of the electrolyte, the resistance of the electrode metal, and the resistance and capacitance at the double layer that forms the electrode-electrolyte interface at the electrode tip (see Robinson, 1968, for details). The effective amplifier input impedance is the total impedance to ground seen from the electrode, and it includes a path through the first amplifier stage and shunting routes to ground outside the amplifier (Nelson, Pouget, Nilsen, Patten, & Schall, 2008).

In our simulations, we used the following values for the filter parameters (see Robinson, 1968): $Z_a = 10\,M\Omega$; $C_s = 100\,pF$; $R_m = \rho_m \frac{L}{\pi r^2}\Omega$; $C_e = 0.2 \times \pi r^2 pF$; $R_s = \frac{\rho_s}{4\pi r}\Omega$; $R_e = \frac{1.33 \cdot 10^4}{\pi r^2}\Omega$, where $\rho_m = 10^{-5}\Omega \times$ cm is the specific resistivity of the electrode metal, L is the length in cm of the portion of the electrode, r is the radius in cm of the electrode tip, and$\rho_s = 72.5\Omega \times$ cm is the specific resistivity of the saline bath.

3 Results

In a previous study (Martinez et al., 2009) we used a similar approach to the one described here to simulate the background noise (but in that study the spike shapes corresponding to the single- and multi-unit activity were extracted from a database of previously recorded spikes rather than being simulated in detail, as here).

In line with our previous results (Martinez et al., 2009), the overall modeling strategy used here reproduced the main characteristics of extracellular recordings. This is exemplified in Figure 5, which shows a simulation with the model described above (left) and a real intracranial recording from the human medial temporal lobe (right) (Quian Quiroga et al., 2007), together with the clusters identified after spike sorting and the normalized power density spectrum (PDS). In this and the following examples, spike detection and sorting was performed using the software package Wave_clus (Quian Quiroga et al., 2004). The synthetic recording was generated using a medium proportion of active neurons (8%) and an electrode diameter of 40 μm. As illustrated in the example, the continuous recording, the sorted spike shapes, and the PDS look very similar for the simulation and the real data.

3.1 Effect of the Electrode Size. Figure 6a illustrates the effect of the size of the electrode, where the three different recordings were generated with the same distribution of neurons for a small (10 μm), medium (40 μm), and large (300 μm) electrode diameter. Small electrodes have typically very high impedances and high signal-to-noise ratios, and they tend to detect the activity of neurons very close to their tip (left); in contrast, large electrodes detect a large number of neurons, but their spikes, averaged over the whole electrode surface area, hardly surpass the noise level (right). The center plot

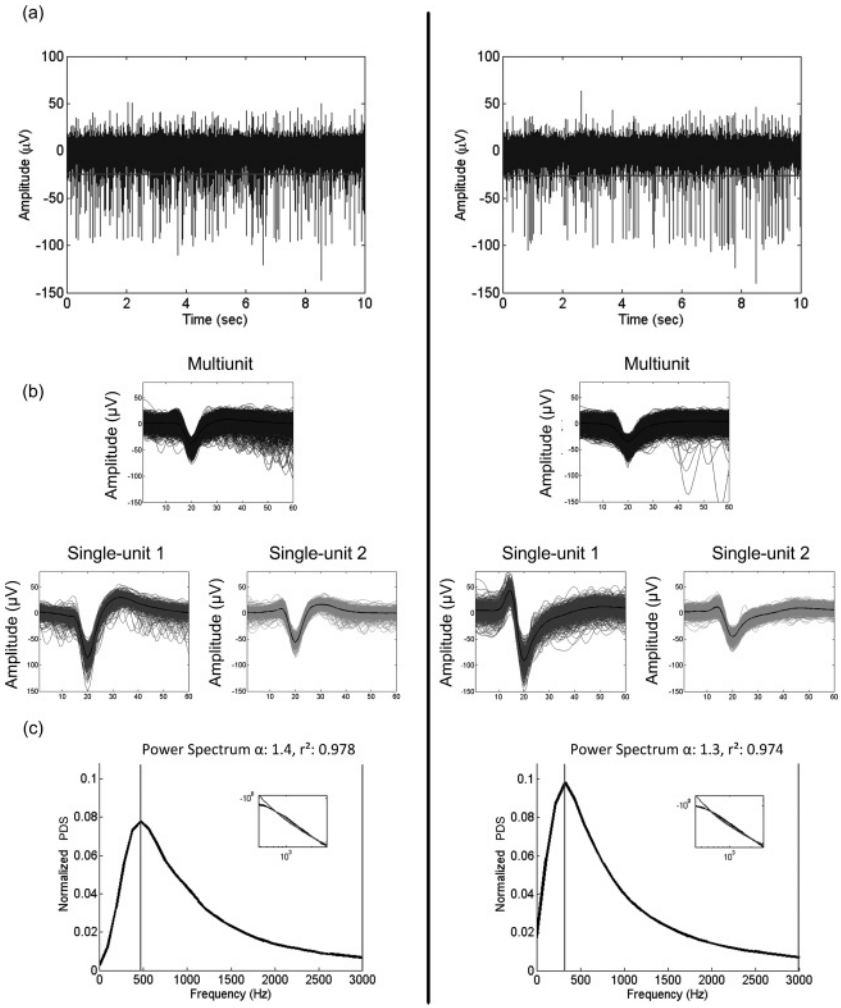

Figure 5: Simulated data set with 8% of active neurons and an electrode diameter of 40 μm (left) and a real recording from human MTL (right). (a) Ten seconds of the continuous signals, where the horizontal lines represent the detection threshold. (b) Clusters identified after an automatic spike sorting of the data, using the whole duration recordings. In both cases, a multiunit and two single units were detected. (c) Normalized power density spectrum (PDS) of the data. The inset depicts the log-log PDS between the two vertical lines, indicating the linear fit to the data, where α represents the slope and r^2 the correlation coefficient.

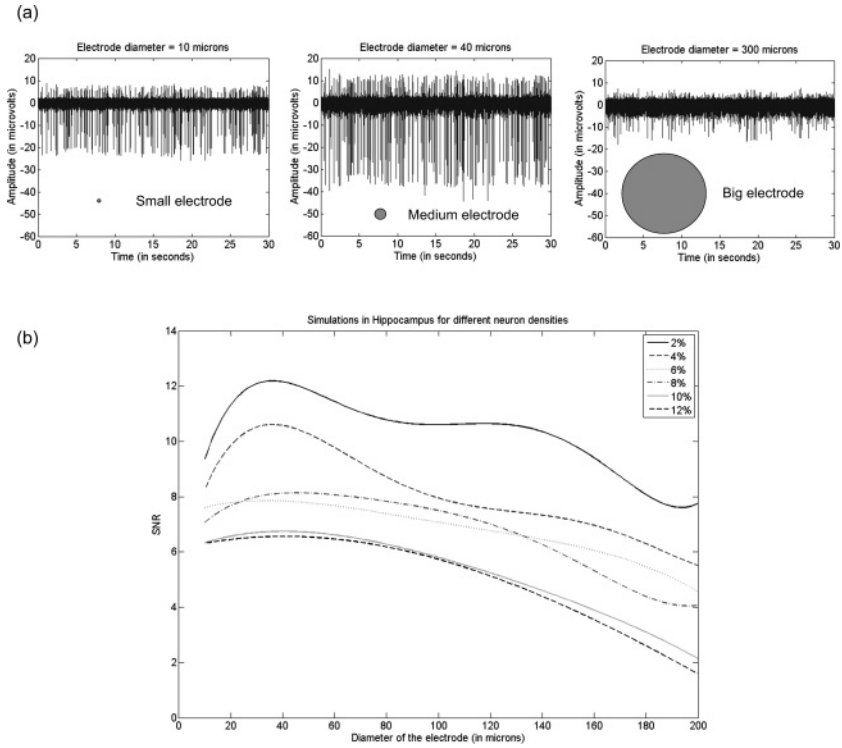

Figure 6: (a) Recording with different electrode sizes for a fixed distribution of neurons. While small electrodes can see only close-by neurons, big electrodes can hardly see individual neurons, as the spike shapes are integrated along their surface. Intermediate electrode sizes give an optimal trade-off between both cases. (b) Signal-to-noise ratio (mean amplitude of spikes over standard deviation of noise) of the recordings as a function of the diameter of the electrode for different proportions of firing neurons.

shows a middle-size electrode, which offers a compromise between these two cases. Note that given the fact that the neurons were placed at random locations, the larger the electrode, the larger the probability of getting very near neurons with high-amplitude spikes. For this reason, the spikes for the medium electrode are larger than for the small electrode, while for the largest electrode, the spikes are washed out by the averaging across its surface. In real life, the spikes seen with small electrodes are typically very large because the electrodes are manually placed close to the neuron to be recorded.

As described in section 2.4, this effect is produced by two features: the spatial averaging of the EAPs across the electrode surface and the filtering

by the equivalent electrical circuit. On the one hand, the spatial averaging produces higher signal-to-noise ratios for small electrodes, given that they record less noise. On the other hand, with a small electrode the chance of being very close to a neuron is smaller. Moreover, considering the equivalent electrical circuit of the electrode (see Figure 1b and Robinson, 1968), small electrodes have a low electrode-to-tissue interface capacitance and a larger spreading resistance, which increases the effective impedance of the electrode, producing an attenuation of the amplitude of the spikes (as shown in Nelson et al., 2008). So although it has been shown that the effect of this equivalent electrode filter can be neglected for low-frequency signals (like the ones in LFP; Nelson & Pouget, 2010), in the frequency range of our simulations, the signal-to-noise (SNR) ratio decreased for small electrodes. Therefore, the combination of these effects gives the optimum response for medium-size electrodes. In order to estimate the optimum electrode size, we generated a set of simulations (of 5 minutes each) with increasing diameters ($d = 10\ \mu$m, 20 μm, 40 μm, 60 μm, 80 μm, 100 μm, 120 μm, 140 μm, 160 μm, 180 μm and 200 μm) and different proportions of active neurons ($P = 2\%$, 4%, 6%, 8%, 10%, 12%, and 14%). The curves in Figure 6b show the average SNR of 20 different simulations for each combination of d and P. The SNRs were calculated as the mean amplitude of spikes of the single neurons identified in the recording (after spike sorting), divided by the standard deviation of the noise, estimated from the expression $\sigma_n = median\{\frac{|x|}{0.6745}\}$, where x is the bandpass filtered signal as in Quian Quiroga et al. (2004). Note that the maximum SNR is reached for electrode diameters between 30 and 50 μm, and therefore we chose the optimal electrode size to be 40 μm.

3.2 Estimation of the Ratio of Active Neurons. Next, we compared the SNR values obtained from real recordings and our simulations to estimate the proportion of active neurons. For this, we generated 30 seconds of simulated extracellular recordings with an electrode diameter of 40 μm (the same diameter used in the real recordings) and 6 rates of firing neurons (2%, 4%, 6%, 8%, 10%, and 12%). Figure 7a shows the average SNR values (in each case across 10 independent simulations) obtained as a function of the proportion of active neurons. Basically the larger the number of active neurons, the lower the SNR, which is due to an increase of the noise level. This result was compared with real recordings obtained with 40 μm width intracranial electrodes from the human MTL (see section 2). The analysis of these data gave an average SNR of about 12, thus corresponding to about 7% of active neurons.

This data set was also used to estimate the number of single units that can be detected after spike sorting (using Wave_clus). Given the density of neurons used in our model, the dashed line in Figure 7b represents the number of neurons within a radius of 50 μm from the electrode that should in principle be detected as single units (Buzsaki, 2004). The solid line shows the number of neurons that were detected by the spike sorting algorithm,

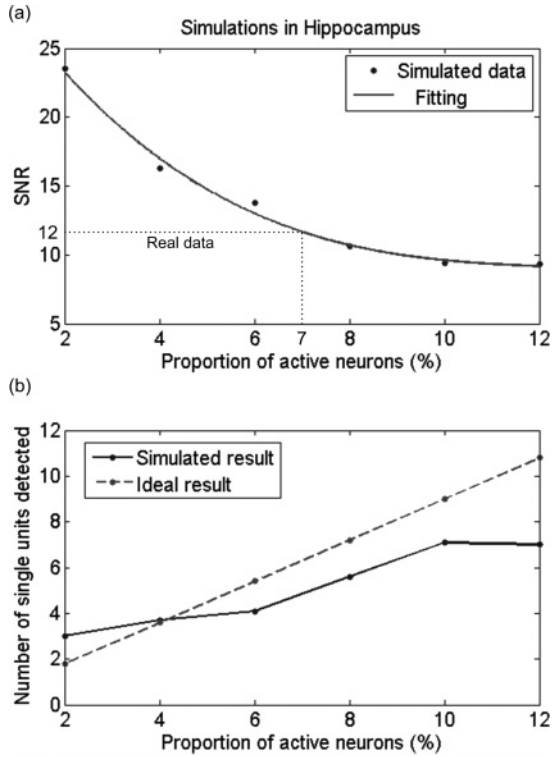

Figure 7: (a) Signal-to-noise ratio (SNR) as a function of the proportion of active neurons obtained with a 40 μm electrode. The larger the number of active neurons, the lower the SNR. With real data (human MTL recordings), an average SNR = 12 was obtained, which in the model corresponds to 7% of firing neurons. (b) Given the density of neurons used in our model, the dashed line indicates the number of neurons within a radius of 50 μm from the electrode, which should in principle be detected as single units. The solid line shows the mean number of single neurons detected with Wave_clus for different proportions of active neurons.

which saturated at a value around 7, thus highlighting a limitation of spike sorting methods (a similar limitation for this and other algorithms was described in detail in Pedreira et al., 2012).

3.3 Simulation of Tetrode Recordings. One of the most interesting advantages of our model is the possibility of generating realistic simulations of tetrode recordings. Figure 8 shows the tetrode configuration used to generate synthetic data sets. Each single electrode contact had a diameter of 40 μm, and the distance between contacts was 50 μm. These data sets

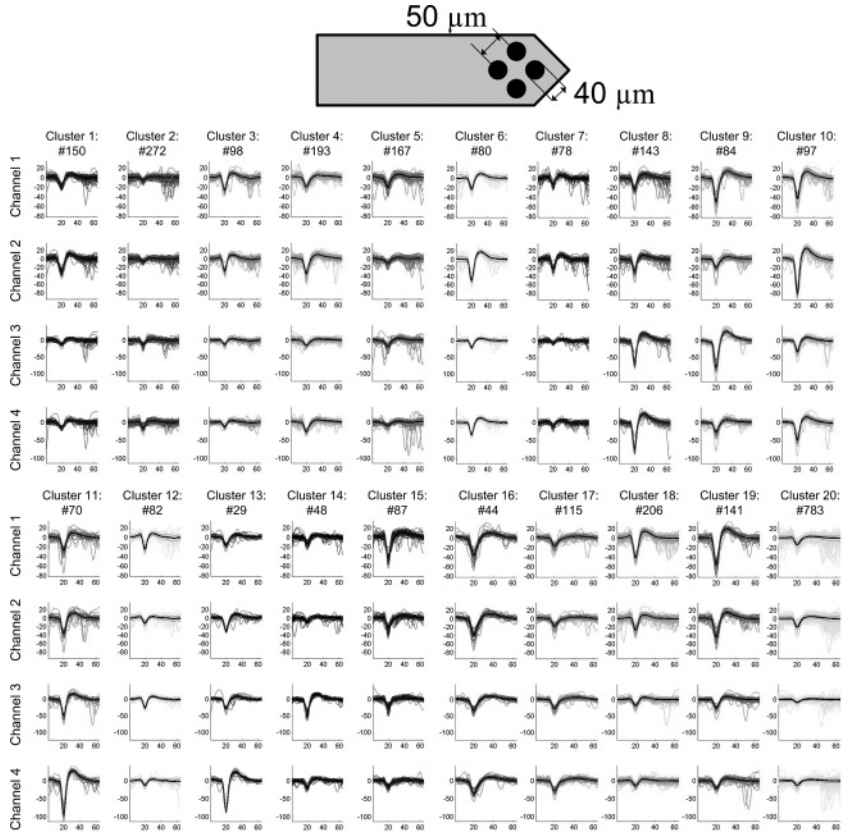

Figure 8: Simulation of a tetrode recording. Each single electrode contact had a diameter of 40 μm, and the distance between contacts was 50 μm (top). Clusters obtained after spike sorting, with 20 identified units (bottom).

were processed with Wave_clus, and the clusters obtained are shown in the bottom part of Figure 8. From this recording, it was possible to detect 20 different single units, a much larger value than the one obtained with single electrodes. Note, in particular, that for each cluster of spikes, not only the amplitude but also the shape of the spike is different in the four channels of the simulated tetrode.

4 Discussion

Given the lack of ground truth with real data, simulations of extracellular recordings, in which the time and identity of each spike is known, are critical for developing and optimizing spike sorting algorithms. This allows

testing and comparing the performance and reliability of different methods, particularly nowadays, when many new spike sorting strategies have been proposed (Lewicki, 1998; Quian Quiroga et al., 2004; Harris et al., 2000; Rutishauser et al., 2006; Quian Quiroga, 2007; Gibson et al., 2012; Kwon et al., 2012). An alternative approach to test spike sorting methods is to use simultaneous intra- and extracellular recordings, but the number of single units that can be validated this way is limited to only one (Harris et al., 2000) or two (Wehr et al., 1999) at a time.

A first approach for generating synthetic extracellular recordings consists of superposing previously recorded spikes into background noise (Lewicki, 1994; Smith & Mtetwa, 2007; Martinez et al., 2009), normalizing the amplitude of the spikes by the distance to the electrode. However, especially for close-by neurons, the position of the electrode affects not only the amplitude but also the shape of the spikes. A more realistic approach would be to use detailed compartmental neuron models (Holt & Koch, 1999; Gold et al., 2006, 2007; Pettersen et al., 2008; Pettersen & Einevoll, 2008), generating the extracellular recording by adding the contribution of each of the simulated neurons. Some preliminary results have been obtained following this strategy (Menne et al., 2002; Mamlouk et al., 2005), but given the high computational load of such calculations, these studies considered small networks of only 90 cells, which is not sufficient to replicate the characteristics of real recordings in general. As a compromise between detailed modeling and computational load, we have presented in this letter a fast approach to simulate extracellular recordings, using detailed compartmental models for neurons in the proximity of the electrode and previously recorded spike shapes for more distant neurons. Using this model, we could simulate extracellular recordings generated by tens of thousands neurons in a 1 mm^3 volume of brain tissue without an excessive computational load. In fact, our modeling approach reduced the processing time by about two orders of magnitude.

We should remark that in this study we did not try to replicate the behavior of the extracellular recording in lower-frequency components (below 300 Hz). Other interesting studies have modeled local field potentials (LFPs), considering the effects of synaptic currents and synchronization patterns within the studied network (Bedard, Kroger, & Destexhe, 2004; Linden et al., 2011). Also, there has been some debate about the presence of transient monopolar components in LFPs (Riera et al., 2012; Destexhe & Bedard, 2012). All of these effects, however, influence frequencies that are much lower than the ones we considered in this study for the testing of spike sorting algorithms and the study of electrode configurations to optimize the identification of the single unit firing.

Concerning the behavior of the extracellular medium for the propagation of electrical potentials, in this model we assumed that it is purely ohmic, which is in agreement with some measurements (Logothetis et al., 2007) but not with others (Gabriel, Lau, & Gabriel, 1996; Bedard et al., 2004;

Bedard, Rodrigues, Roy, Contreras, & Destexhe, 2010). Assuming a resistive medium (and not considering low-pass filtering effects given by the cable properties of the neurons shown in Pettersen & Einevoll, 2008), for distant neurons we varied only the amplitude (and not the shape) as a function of the distance. The model could be extended to take into account non-ohmic (i.e., capacitive) extracellular media by describing the characteristics of the propagation medium in the NEURON models for close-by neurons and implementing a (low-pass) filter transfer function to calculate the shapes of the spikes generated by distant neurons after propagating through the medium.

The use of multiple close-by recording sites—stereotrodes (McNaughton, Okeefe, & Barnes, 1983), tetrodes (Okeefe & Recce, 1993; Gray, Maldonado, Wilson, & McNaughton, 1995), or more recently, high-density multielectrode arrays (Segev, Goodhouse, Puchalla, & Berry, 2004; Blanche, Spacek, Hetke, & Swindale, 2005; Frey, Egert, Heer, Hafizovic, & Hierlemann, 2009; Lambacher et al., 2011)—has been proposed to increase the number of identified neurons. The basic idea of these electrodes is to improve spike sorting by observing the neurons from different locations, given that spike shapes that are hard to differentiate in a given channel can be separated by considering their differences in other channels. Stereotrodes, and in particular tetrodes, are widely used now and there is a clear consensus that they improve spike sorting performance and reliability (Gray et al., 1995). Nevertheless, the development and optimization of spike sorting methods for tetrode data is still in very early stages, as typically only the peak amplitude, or the first two principal components of the four channels is considered. Moreover, it is not known which polytrode configuration (number of electrodes, diameter, and distance) gives the best results. So both the development of optimal polytrodes and the spike sorting methods to analyze these data cry out for realistic simulations of these types of recordings. The easiest strategy to generate synthetic polytrode recordings consists of scaling the spike amplitudes for each electrode. However, this approach does not consider the fact that the position of the neuron with respect to each electrode determines not only the spike amplitude but also its shape. In this respect, we have shown that it is possible to apply our model to calculate the spike shapes (and amplitudes) generated by each neuron at different electrode positions. This provides ground truth data to develop new spike sorting methods specifically designed for polytrode recordings and, in turn, to systematically study the performance obtained with different polytrode configurations. Compared to a previous study by our group (Martinez et al., 2009), the main advantage of the new modeling approach described here is that we have not only considered the variation of spike amplitudes but also the variation of spike shapes for close-by neurons, a feature that is critical for simulations of polytrode recordings. Moreover, we have also introduced an electrode model to study the effect of finite-size electrodes in the recordings.

In the first paragraph of his by-now classic paper describing the properties of microelectrodes, Robinson (1968) states, "Despite this prevalence [of the use of microelectrodes for single-cell recordings] there remains a good deal of mystery about how best to make these electrodes and how to interpret the extracellular potentials that they record. The attitude of many practical users is the sensibly pragmatic one of the biological assay. When one finds some method of making microelectrodes that successfully isolate units in a given neuronal structure, one 'freezes the design' and attends to the more important task of collecting neural data." Surprisingly, more than four decades after Robinson's seminal work, there are still relatively few papers describing the influence of the size and shape of the electrodes on the extracellular recordings (Schmidt, 1999; Moffitt & McIntyre, 2005; Massobrio, Massobrio, & Martinoia, 2007; Nelson et al., 2008; Nelson & Pouget, 2010; Lempka et al., 2011; Mechler & Victor, 2012). It is possible that simple optimizations regarding the shape and sizes of the electrodes may lead to a much better single-neuron yield. In this respect, we integrated Robinson's electrode model in our simulations and have analyzed the influence of the electrode size on the quality of the recordings. A diameter of 40 μm provided the best signal-to-noise ratio. We should stress, however, that this result should be taken as a first approximation given the simplifications in our model (e.g., we considered a circular exposed area, without considering any coating).

Converging evidence suggests that many neurons remain inactive or have very small firing rates in awake states—what has been named the "dark matter" problem in neuroscience (Shoham et al., 2006). Neurons with such extremely low baseline firing have indeed been shown to be of relative importance for encoding and generating new memories in the human medial temporal lobe (Quian Quiroga, Reddy, Kreiman, Koch, & Fried, 2005; Quian Quiroga, 2012b). To get a first estimation of the number of silent neurons in recordings in the human medial temporal lobe, we compared the SNR ratio obtained in real human MTL recordings with the one obtained in the simulated recordings, using the same neural density and firing distribution as the one observed in the real data. Given the mean SNR value obtained in the real data, we estimated that this corresponds to about 7% of active neurons, which is consistent with the estimations in previous studies (Shoham et al., 2006).

To facilitate further research, we have created a downloadable package with a software interface that controls all the Matlab and NEURON code used to generate extracellular recordings where the simulation parameters can be changed by the user. It is available at www.le.ac.uk/csn/neurocube.

Acknowledgments

This work was supported by BBSRC (BB/H014047/1), MRC, and EPSRC.

References

Bedard, C., Kroger, H., & Destexhe, A. (2004). Modeling extracellular field potentials and the frequency-filtering properties of extracellular space. *Biophysical Journal*, *86*, 1829–1842.

Bedard, C., Rodrigues, S., Roy, N., Contreras, D., & Destexhe, A. (2010). Evidence for frequency-dependent extracellular impedance from the transfer function between extracellular and intracellular potentials. *Journal of Computational Neuroscience*, *29*(3), 389–403.

Blanche, T. J., Spacek, M. A., Hetke, J. F., & Swindale, N. V. (2005). Polytrodes: High-density silicon electrode arrays for large-scale multiunit recording. *Journal of Neurophysiology*, *93*(5), 2987–3000.

Buzsaki, G. (2004). Large-scale recording of neuronal ensembles. *Nature Neuroscience*, *7*(5), 446–451.

Contreras, D., Destexhe, A., & Steriade, M. (1997). Intracellular and computational characterization of the intracortical inhibitory control of synchronized thalamic inputs in vivo. *Journal of Neurophysiology*, *78*(1), 335–350.

Destexhe, A., & Bedard, C. (2012). Do neurons generate monopolar current sources? *Journal of Neurophysiology*, *108*(4), 953–955.

Destexhe, A., Contreras, D., Steriade, M., Sejnowski, T. J., & Huguenard, J. R. (1996). In vivo, in vitro, and computational analysis of dendritic calcium currents in thalamic reticular neurons. *Journal of Neuroscience*, *16*(1), 169–185.

Einevoll, G. T., Franke, F., Hagen, E., Pouzat, C., & Harris, K. D. (2011). Towards reliable spike-train recordings from thousands of neurons with multielectrodes. *Current Opinion in Neurobiology*, *22*(1), 11–17.

Farrant, M., Feldemeyer, D., Takahashi, T., & Cull-Candy, S. G. (1994). NMDA-receptor channel diversity in the developing cerebellum. *Nature*, *368*, 335–339.

Frey, U., Egert, U., Heer, F., Hafizovic, S., & Hierlemann, A. (2009). Microelectronic system for high-resolution mapping of extracellular electric fields applied to brain slices. *Biosensors and Bioelectronics*, *24*(7), 2191–2198.

Gabriel, S., Lau, R. W., & Gabriel, C. (1996). The dielectric properties of biological tissues: III. Parametric models for the dielectric spectrum of tissues. *Physics in Medicine and Biology*, *41*, 2271–2293.

Gibson, S., Judy, J. W., & Markovi, D. (2012). Spike sorting: The first step in decoding the brain. *IEEE Signal Processing Magazine*, *29*(1), 124–143.

Gold, C., Henze, D. A., & Koch, C. (2007). Using extracellular action potential recordings to constrain compartmental models. *Journal of Computational Neuroscience*, *23*(1), 39–58.

Gold, C., Henze, D. A., Koch, C., & Buzsaki, G. (2006). On the origin of the extracellular action potential waveform: A modeling study. *Journal of Neurophysiology*, *95*(5), 3113–3128.

Gray, C. M., Maldonado, P. E., Wilson, M., & McNaughton, B. (1995). Tetrodes markedly improve the reliability and yield of multiple single-unit isolation from multi-unit recordings in cat striate cortex. *Journal of Neuroscience Methods*, *63*(1–2), 43–54.

Harris, K. D., Henze, D. A., Csicsvari, J., Hirase, H., & Buzsaki, G. (2000). Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements. *Journal of Neurophysiology*, *84*(1), 401–414.

Henze, D. A., Borhegyi, Z., Csicsvari, J., Mamiya, A., Harris, K. D., & Buzsaki, G. (2000). Intracellular features predicted by extracellular recordings in the hippocampus in vivo. *Journal of Neurophysiology*, *84*(1), 390–400.

Hill, D. N., Mehta, S. B., & Kleinfeld, D. (2011). Quality metrics to accompany spike sorting of extracellular signals. *Journal of Neuroscience*, *31*(24), 8699–8705.

Hines, M. L., & Carnevale, N. T. (1997). The NEURON simulation environment. *Neural Computation*, *9*(6), 1179–1209.

Holt, G. R., & Koch, C. (1999). Electrical interactions via the extracellular potential near cell bodies. *Journal of Computational Neuroscience*, *6*(2), 169–184.

Ison, M. J., Mormann, F., Cerf, M., Kock, C., Fried, I., & Quian Quiroga, R. (2011). Selectivity of pyramidal cells and interneurons in the human medial temporal lobe. *Journal of Neurophysiology*, *106*, 1713–1721.

Kwon, K. Y., Eldawlatly, S., & Oweiss, K. (2012). NeuroQuest: A comprehensive analysis tool for extracellular neural ensemble recordings. *Journal of Neuroscience Methods*, *204*(1), 189–201.

Lambacher, A., Vitzthum, V., Zeitler, R., Eickenscheidt, M., Eversmann, B., Thewes, R., et al. (2011). Identifying firing mammalian neurons in networks with high-resolution multi-transistor array (MTA). *Applied Physics A—Materials Science and Processing*, *102*(1), 1–11.

Lempka, S. F., Johnson, M. D., Moffitt, M. A., Otto, K. J., Kipke, D. R., & McIntyre, C. C. (2011). Theoretical analysis of intracortical microelectrode recordings. *Journal of Neural Engineering*, *8*(4), 045006.

Lewicki, M. S. (1994). Bayesian modeling and classification of neural signals. *Neural Computation*, *6*(5), 1005–1030.

Lewicki, M. S. (1998). A review of methods for spike sorting: The detection and classification of neural action potentials. *Network: Computation in Neural Systems*, *9*(4), 53–78.

Linden, H., Tetzlaff, T., Potjans, T. C., Pettersen, K. H., Gruen, S., Diesmann, M., et al. (2011). Modeling the spatial reach of the LFP. *Neuron*, *72*(5) 859–872.

Llinas, R. R. (1988). The intrinsic electrophysiological properties of mammalian neurons: Insights into central nervous system function. *Science*, *242*, 1654–1664.

Logothetis, N. K., Kayser, C., & Oeltermann, A. (2007). In vivo measurement of cortical impedance spectrum in monkeys: Implications for signal propagation. *Neuron*, *55*(5), 809–823.

Mamlouk, A. M., Sharp, H., Menne, K. M. L., Hofmann, U. G., & Martinetz, T. (2005). Unsupervised spike sorting with ICA and its evaluation using GENESIS simulations. *Neurocomputing*, *65*, 275–282.

Markram, H., Toledo-Rodriguez, M., Wang, Y., Gupta, A., Silberberg, G., & Wu, C. Z. (2004). Interneurons of the neocortical inhibitory system. *Nature Reviews Neuroscience*, *5*(10), 793–807.

Martinez, J., Pedreira, C., Ison, M. J., & Quian Quiroga, R. (2009). Realistic simulation of extracellular recordings. *Journal of Neuroscience Methods*, *184*(2), 285–293.

Massobrio, P., Massobrio, G., & Martinoia, S. (2007). Multi-program approach for simulating recorded extracellular signals generated by neurons coupled to microelectrode arrays. *Neurocomputing*, *70*(13–15), 2467–2476.

McNaughton, B. L., Okeefe, J., & Barnes, C. A. (1983). The stereotrode: A new technique for simultaneous isolation of several single units in the central nervous system from multiple unit records. *Journal of Neuroscience Methods*, *8*(4), 391–397.

Mechler, F., & Victor, J. D. (2012). Dipole characterization of single neurons from their extracellular action potentials. *Journal of Computational Neuroscience*, *32*(1), 73–100.

Menne, K. M. L., Folkers, A., Malina, T., Maex, R., & Hofmann, U. G. (2002). Test of spike-sorting algorithms on the basis of simulated network data. *Neurocomputing*, *44*, 1119–1126.

Moffitt, M. A., & McIntyre, C. C. (2005). Model-based analysis of cortical recording with silicon microelectrodes. *Clinical Neurophysiology*, *116*(9), 2240–2250.

Nelson, M. J., & Pouget, P. (2010). Do electrode recordings create a problem in interpreting local field potential recordings? *Journal of Neurophysiology*, *103*, 2315–2317.

Nelson, M. J., Pouget, P., Nilsen, E. A., Patten, C. D., & Schall, J. D. (2008). Review of signal distortion through metal microelectrode recording circuits and filters. *Journal of Neuroscience Methods*, *169*(1), 141–157.

Okeefe, J., & Recce, M. L. (1993). Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus*, *3*(3), 317–330.

Pedreira, C., Martinez, J., Ison, M. J., & Quian Quiroga, R. (2012). How many neurons can we see with current spike sorting algorithms? *Journal of Neuroscience Methods*, *211*(1), 58–65.

Pettersen, K. H., & Einevoll, G. T. (2008). Amplitude variability and extracellular low-pass filtering of neuronal spikes. *Biophysical Journal*, *94*(3), 784–802.

Pettersen, K. H., Hagen, E., & Einevoll, G. T. (2008). Estimation of population firing rates and current source densities from laminar electrode recordings. *Journal of Computational Neuroscience*, *24*(3), 291–313.

Plonsey, R. (1969). *Bioelectric phenomena*. New York: McGraw-Hill.

Quian Quiroga, R. (2007). Spike sorting. *Scholarpedia*, *2*(12), 3583.

Quian Quiroga, R. (2012a). Quick guide: Spike sorting. *Current Biology*, *22*, R45–R46.

Quian Quiroga, R. (2012b). Concept cells: The building blocks of declarative memory functions. *Nature Reviews Neuroscience*, *13*, 587–597.

Quian Quiroga, R., Nadasdy, Z., & Ben-Shaul, Y. (2004). Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. *Neural Computation*, *16*(8), 1661–1687.

Quian Quiroga, R., Reddy, L., Koch, C., & Fried, I. (2007). Decoding visual inputs from multiple neurons in the human temporal lobe. *Journal of Neurophysiology*, *98*, 1997–2007

Quian Quiroga, R., Reddy, L., Kreiman, G., Koch, C., & Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature*, *435*, 1102–1107.

Riera, J., Ogawa, T., Goto, T., Sumiyoshi, A., Nonaka, H., Evans, A., et al. (2012). Pitfalls in the dipolar model for the neocortical EEG sources. *Journal of Neurophysiology*, *108*(4), 956–975.

Robinson, D. A. (1968). The electrical properties of metal microelectrodes. *Proceedings IEEE*, *56*(6), 1065–1071.

Rudolph, M., Pelletier, J. G., Pare, D., & Destexhe, A. (2005). Characterization of synaptic conductances and integrative properties during electrically induced EEG-activated states in neocortical neurons in vivo. *Journal of Neurophysiology*, *94*(4), 2805–2821.

Rutishauser, U., Schuman, E. M., & Mamelak, A. N. (2006). Online detection and sorting of extracellularly recorded action potentials in human medial temporal lobe recordings, in vivo. *Journal of Neuroscience Methods*, *154*(1–2), 204–224.

Schmidt, E. M. (1999). Electrodes for many single neuron recordings. In M.A.L. Nicolelis (Ed.), *Methods for neural ensemble recordings*. Boca Raton, FL: CRC Press.

Schmitzer-Torbert, N., Jackson, J., Henze, D., Harris, K. D., & Redish, A. D. (2005). Quantitative measures of cluster quality for use in extracellular recordings. *Neuroscience*, *131*(1), 1–11.

Segev, R., Goodhouse, J., Puchalla, J., & Berry, M. J. (2004). Recording spikes from a large fraction of the ganglion cells in a retinal patch. *Nature Neuroscience*, *7*(10), 1154–1161.

Shoham, S., O'Connor, D. H., & Segev, R. (2006). How silent is the brain: Is there a "dark matter" problem in neuroscience? *Journal of Comparative Physiology A—Neuroethology Sensory Neural and Behavioral Physiology*, *192*(8), 777–784.

Smith, L. S., & Mtetwa, N. (2007). A tool for synthesizing spike trains with realistic interference. *Journal of Neuroscience Methods*, *159*(1), 170–180.

Wehr, M., Pezaris, J. S., & Sahani, M. (1999). Simultaneous paired intracellular and tetrode recordings for evaluating the performance of spike sorting algorithms. *Neurocomputing*, *26–27*, 1061–1068.

Received September 7, 2012; accepted December 2, 2012.