In press with "5th German Workshop on Artificial Life - Abstracting and Synthesizing the Principles of Living Systems",Eds. D. Polani, J. Kim, T. Martinetz (Eds.), IOS Press Infix.

Biologically realistic simulation of a part of hippocampal CA3: Generation of testdata for the evaluation of spike detection algorithms

Kerstin M. L. Menne, Thomas Malina, Ulrich G. Hofmann

Institut für Signalverarbeitung und Prozeßrechentechnik
Seelandstraße 1a
D-23569 Lübeck, Germany
[menne/malina/hofmann]@isip.mu-luebeck.de

Abstract

Biologically realistic simulations on the neuronal network level are frequently used in computational neuroscience in order to explain and understand biological brain functions. The presented simulation mimicks a tiny part of the CA3 region of the hippocampus. But new insights about the functionality of CA3 are not the ultimate goal in this case. Rather we want to show that sound extracellular potential data can be generated within a network close to reality. The data is used for the evaluation of spike detection methods.

1 Introduction

Current neurophysiology research benefits strongly from advances made in microstructural technologies [8],[16] and neurophysiological methods [6],[4],[9],[22] and thus recordings from an ever increasing number of neuronal sources become more and more ubiquitously available. However, this seems to come at the expense of thoroughness in raw data analysis, since ”visual” data screening, e.g. to correctly assign spikes on incoming data, requires almost supernatural stamina in cases of data sets originating from dozens or even hundreds of recording sites over several hours of experimental time. Having noticed such problems the solution is obvious: Utilize automated analysis tools, which never get tired and always perform perfectly well. Unfortunately, to validate the performance of such tools, some kind of trustworthy test data has to be used. Real data is not quite suitable for test purposes, since it is always subjected to interpretations concerning contained spike trains. On the other hand, simulated data is controllable but at the same time usually very artificial. One frequently applied method to generate synthetic data is to extract spikes and chunks of noise from extracellular recordings, to distribute the spikes randomly in time and to add the noise sample [2], [18]. The disadvantage of this method is quite obvious. The randomly distributed spikes do not reflect any biologically realistic network behaviour, including cell interactions and synchronous spikings, and even no biologically realistic single cell behaviour. Nakatani [14] employs a peripheral nerve model in order to generate synthetic data. He adds a Gaussian noise process to the simulated signal. This method is comparable to ours, but lacks the presence of different interacting neurons within an elaborated network cicuitry. The data we provide, simulated in a biologically realistic network simulation, is intended to bridge the gap between biological relevance and control, and thus is predestinated for the test of spike analysis methods like for example spike detection algorithms.

To be published in "5th German Workshop on Artificial Life - Abstracting and Synthesizing the Principles of Living Systems",Eds. D. Polani, J. Kim, T. Martinetz (Eds.), IOS Press Infix.

2 Network simulation

The network simulation is implemented in GENESIS 2.2, the GEneral NEural SImulation System [3]. The network simulation we present has nothing to do with artificial neural networks. Our network simulation is based on cell models that try to recapitulate the knowledge about anatomy and electrophysiological behaviour of hippocampal CA3 pyramidal cells and interneurons. What hides behind the simulation is mathematically speaking a system of differential equations. This is transfered into a system of difference equations and numerically solved with the implicite Crank-Nicholson method with a step size of 2.5e-5.

We applied a pyramidal cell model and an interneuron model as suggested by Traub [19], [20]. The pyramidal cell model consists of 66 compartments, representing branching dendritic trees, a soma and an axon (Figure 2(b)). Active ion channels are assigned to the compartments, namely Na^+, Ca^{2+}, K_DR, K_AHP, K_A and K_C channels. The interneuron model comprises 48 compartments and the same active ion channels as the pyramidal cell model. The network is built up from 72 pyramidal cell models, arranged in a 6 times 12 array, and 9 feedforward as well as 9 feedback interspersed interneurons (Figure 2(a)). Pyramidal cells have a distance of 10 μm +/- 3 μm from each other in x- and y-direction. z-coordinates are randomized in a range from -50 to +50 μm. The rather high cell density compared to earlier implementations [12] was made necessary by the rather small horizon of the simulated electrodes, see below. Each cell model is randomly rotated 0 to 180 degrees clockwise about its z-axis in an attempt to compensate the fact that we have identical cell morphologies (”clones”). Feedforward and feedback interneurons are derived from the same model. They differ exclusively in the way they are connected to other neurons. Synapses are simulated by AMPA, NMDA, GABA_A and GABA_B receptors. The interneurons have only AMPA receptors. Figure 1 reveals the network circuitry at a glance. Cell connections can be all-

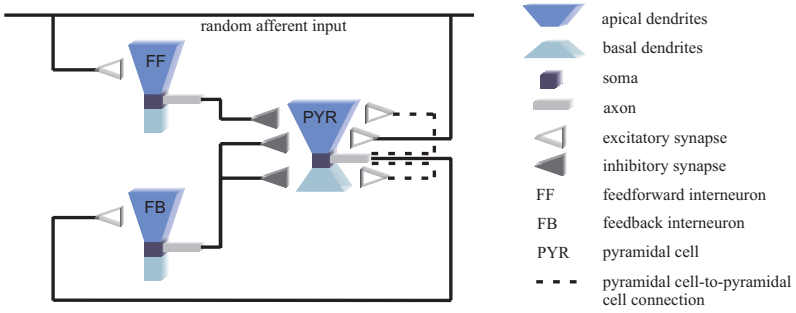

Figure 1: Network circuitry. Compared to the network circuitry published in [12], [13], pyramidal cells do not get inhibitory inputs on their basal dendrites from feedforward interneurons anymore and in turn do not excite feedback interneurons at basal dendrites.

to-all, but are established with a certain probability only. Feedforward interneurons are excited by random afferent input onto their apical dendrites. Feedback interneurons are excited in the same region, but the input originates from pyramidal cells. Pyramidal cells receive in turn feedback inhibition in the perisomatic region and feedforward inhibition at their apical dendrites. Recurrent excitation among the pyramidal cells targets onto the perisomatic region, whereas random afferent input, representing mossy fiber input, is received at the apical dendrites. CA3 projections into CA1 are not modeled. Slight changes in network circuitry compared to [12], [13] proved to be advantageous for the overall network behaviour. The distribution af activity across the pyramidal cell array is best visualized by means of a movie. A 3D representation of the array (Figure 2(b)) is generated with help of the Visualization Toolkit VTK (Kitware Inc.). Colour is assigned to each

To be published in "5th German Workshop on Artificial Life - Abstracting and Synthesizing the Principles of Living Systems",Eds. D. Polani, J. Kim, T. Martinetz (Eds.), IOS Press Infix

neuron dependend on the somatic membrane potential. Colour changes mirror activity changes. The network elicits both, simultaneous bursts as well as isolated spikes.

3 Simulation of extracellular recordings

Single-site recording probes get simulated with the help of GENESIS ”efield” objects arranged at arbitrary positions within the above described network. The ”efield” object is an implementation of an equation by Nunez [17] for the calculation of extracellular field potentials:

$$F = \frac{1}{4 \cdot \pi \cdot s} \sum_{i=1}^{n} \frac{I_i}{r_i}. \tag{1}$$

Transmembrane currents I_i of n compartments are added up with respect to their distance r_i from the ”electrode”. s denotes conductivity. A homogeneous resistivity and no capacitance are assumed for the extracellular space. The Nunez equation does not provide a direction characteristic: Signals are recorded throughout the whole surrounding space. In our simulation, this isotropic detection characteristic has to be changed to a more directional sensitive by exclusively taking into account the transmembrane currents of cells situated within a certain sector as seen from the probe (Figure 2(a)). An opening angle of 90° proved to be too small. Contrary to experimental results,

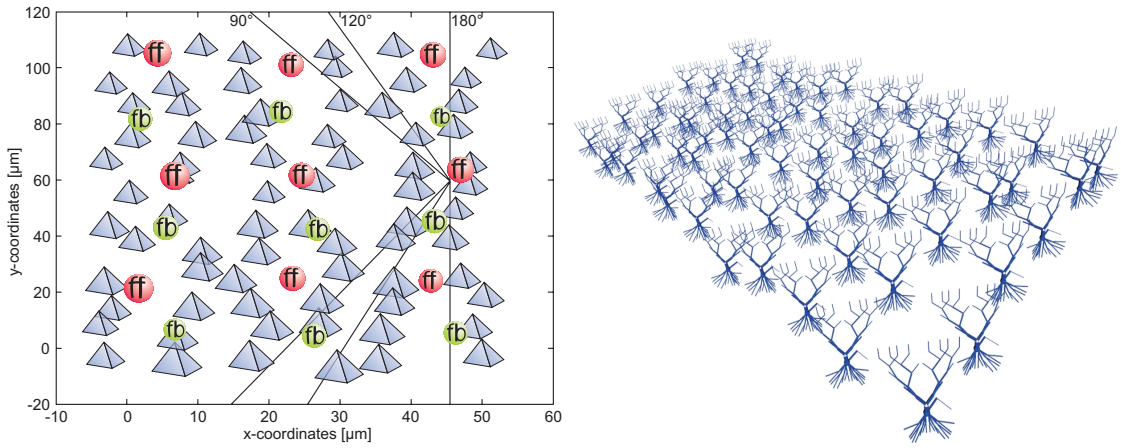

Figure 2: The graph on the left illustrates different opening angles. The discrete nature of the network restricts the possible number of gradations. The graph on the right illustrates only the pyramidal cell array making use of the compartmental model.

only spikes of the same amplitude were contained in the recorded signal. An opening angle of 120° allowed for different amplitude heights of spikes coming from a larger number of neurons than in the 90° case (Figure 3). The recordable sector for recording sites fixed on an insulating carrier most likely comes close to a 120° sector.

Multi-site recording probes can be simulated by the linear arrangement of ”efield” objects. The ”recorded” signals contain low-frequency field potentials and high-frequency spikes that mirror intracellular activity. Many overlapping spikes, originating from synchronous activity of several neurons, are contained in the data - a tough problem for all types of spike train analysis published. If individual recording sites (i.e. ”efield” objects) are arranged close enough, spikes originating

To be published in "5th German Workshop on Artificial Life - Abstracting and Synthesizing the Principles of Living Systems",Eds. D. Polani, J. Kim, T. Martinetz (Eds.), IOS Press Infix.

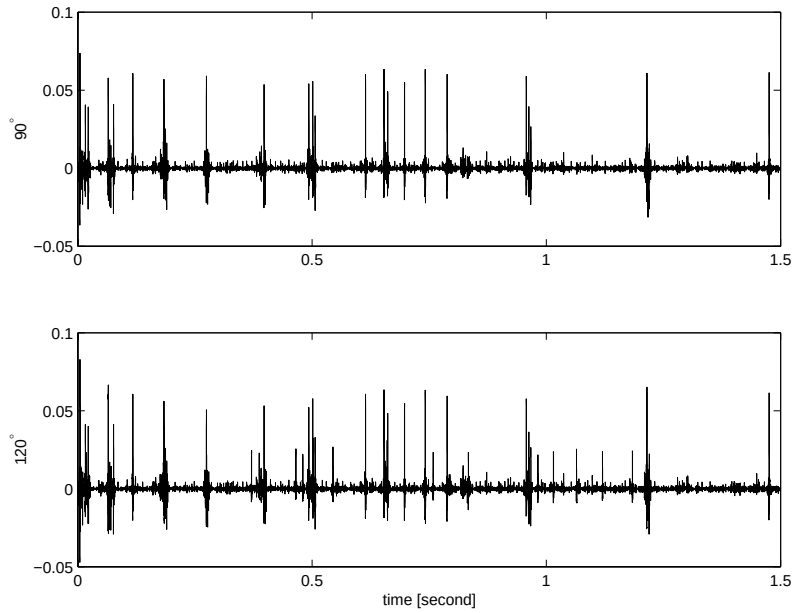

Figure 3: Comparison of extracellularly recorded signals, where the opening angles amount to 90° and 120°, respectively. Signals are given in volt.

from one cell are trackable, but exhibit a characteristic decay over distance, on several recording channels, like in real experiments. We simulated probes comprising 13 different recording sites. The individual recording sites are linearly arranged, perpendicularly to the xy-plane. The lowest z-coordinate is at -55 micrometer, the highest one at 95 micrometer. Thus individual recording sites have a distance of 12.5 micrometer. A spike amplitude decay along the site array is a feature of real multi-site recordings. Figure 4 illustrates the footprints (on 10 channels) of 4 spikes coming from pyramidal cells 34 and 40, respectively. The average positive peak amplitude distribution given on the right of Figure 4 reveals that the highest amplitudes are not always found in one and the same channel. This is due to the contribution of cells that are located at different z-coordinates. Unfortunately, individual electrodes do not have a big recording horizon, since spikes are trackable in two channels only and the simulated recording points have a distance of 12.5 μm. Spikes originating from cells within the horizon are clearly identifiable in extracellularly recorded signals. In the experimental case, depending on the set-up, neural spikes can be tracked for many tens of microns [10], [1]. Therefore the decrease in amplitude seems to be too steep along the simulated linear array. This way too rapid decay in amplitude can be found, when the simulated multi-site recording probe is moved away from the cells, as well. This was investigated by generating a highpass filtered signal of the electrode at z=70 μm at different distances from the soma of the pyramidal cell 40. 17 spikes were detected with a threshold of 0.04 mV at a distance of 4 μm. These spikes were tracked throughout the recordings at increasing distances. The average peak-to-peak amplitude of the 17 spikes was calculated at each distance. The result (Figure 5(a)) confirms the above made observation of a rapid decay. At a cell-electrode distance of 15 μm, the amplitudes lose already more than $2/3$ of their value. Gray [7] for example specifies for tetrode (impedance of 0.5 to 1 MΩ at 1 kHz) recordings in the neocortex, area 17 of cats, a cell-electrode distance of 65 μm as the distance at which a 90% reduction in voltage occurs. He outlines, that this result agrees with theoretical estimates by Rall concerning the decay of extracellular spike voltage. We therefore conclude, that the small horizon of the simulated electrodes casts doubt on the validity

To be published in "5th German Workshop on Artificial Life - Abstracting and Synthesizing the Principles of Living Systems",Eds. D. Polani, J. Kim, T. Martinetz (Eds.), IOS Press Infix.

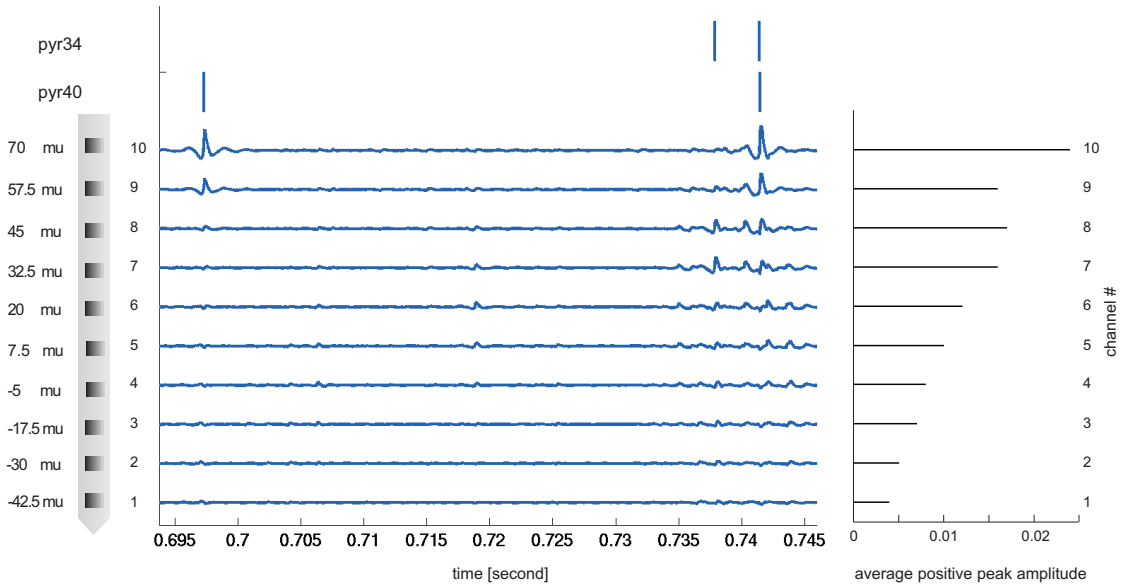

Figure 4: Amplitude decay alongside a multi-site recording probe. Pyramidal cell 40 is 4μm away from the probe, pyramidal cell 34 8.1μm. The calculations of the average spike amplitudes are based on 90 to 120 threshold detected spikes in each case (after highpass filtering).

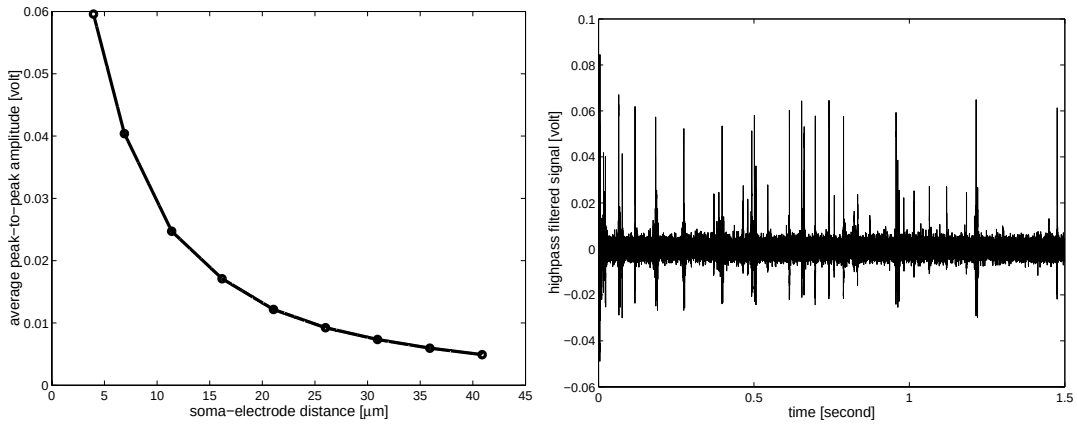

Figure 5: The graph on the left shows the decay of spike amplitude with increasing distance between soma and electrode. The graph on the right shows a simulated extracellularly recorded signal (z=70μm) with added white noise (SNR of 21:1, calculated as ratio of average spike peak-to-peak amplitude and noise rms value (root-mean-squared)).

To be published in "5th German Workshop on Artificial Life - Abstracting and Synthesizing the Principles of Living Systems",Eds. D. Polani, J. Kim, T. Martinetz (Eds.), IOS Press Infix.

of the Nunez equation (1) for the calculation of fast extracellular potentials and requires further simulation and experimental work to be done in the future.

The simulated extracellular potential data per se contains background activity coming from cells farther away from the electrode. In order to make the simulated data resemble experimental data even more, white noise is added after the simulation to represent thermal noise of recording devices, and the signals get highpass filtered with a cut-off frequency of 500 Hz (Figure 5(b)). The addition of noise at this process step allows to play through different noise levels.

4 Results and Discussion

We were able to show the possibility to generate sound extracellular potential data in a biologically realistic network simulation. The simulated data shows a lot of features of experimental data, including spike shapes, amplitude decay over distance and contained noise. Therefore it seems justified to use our simulated data as benchmark set for spike detection algorithms. As far as we know, nobody ever applied realistically simulated data like ours for test purposes. (Our data set as well as additional information can be downloaded from
www.isip.mu-luebeck.de/∼hofmann/Paper/Menne_CD/.)

So far, we tested the following spike detection methods on the basis of our simulated data: positive voltage threshold (pt) [10, e.g.], pt plus peak-to-peak amplitude threshold (pt+ppa) [15, e.g.], pt+ppa plus peak-to-peak time (pt+ppa+pp-time) [1, e.g.], peak-to-peak amplitude only (ppa-window) [2] , thresholding of the second derivative of a signal (sec_dev+pt) [11],[21], of the energy (energy+pt) [15] or the by means of Discrete Wavelet Transform (DWT) denoised signal (dwt-denoising+pt) [5]. Figure 6 illustrates the results. The two bars on top each case result from comparisons of the detection results with the spike times of two pyramidal cells only (70 spikes in total), the two bars on bottom from comparisons with spike times of three pyramidal cells that are definitely contributing to the simulated signal (112 spikes in total).The spike amplitudes of the third pyramidal cell are hardly above the noise level and therefore difficult to detect. In the case of the pt method e.g., 24% of the 70 spikes originating from two pyramidal cells are missed, compared to 45% of 112 spikes originating from three pyramidal cells. However, even 24% of missed spikes are an alarmingly high value. The investigation of peak-to-peak amplitude and peak-to-peak time in addition to the positive peak amplitude does not improve the performance. The ppa-window, sec_dev+pt and energy+pt methods predict much more than 50% false positives an thus should not be used as stand-alone detection methods. Thresholding of the denoised signal, however, yields even better results than thresholding of the raw signal. The presented results were achieved for data shown in Figure 5(b). We tested the thresholding of raw data, a commonly applied approach in spike detection, for two lower SNRs. The results presented in Figure 7 clearly reveal, that a positive threshold method is completely unsuitable under low SNRs. The number of false positives is by far dominating and would falsify any conclusions based on the detection results. In addition, in the case of the SNR of 6:1, there are almost 50 intolerable percent false negatives.

The test of further spike detection algorithms is pending, but the best method will finally find its way into our realtime data acquisition system (www.vsamuel.de). Special attention will be paid to approaches that take advantage of multi-site recording data and to the performance on overlapping spikes.

To be published in "5th German Workshop on Artificial Life - Abstracting and Synthesizing the Principles of Living Systems",Eds. D. Polani, J. Kim, T. Martinetz (Eds.), IOS Press Infix.

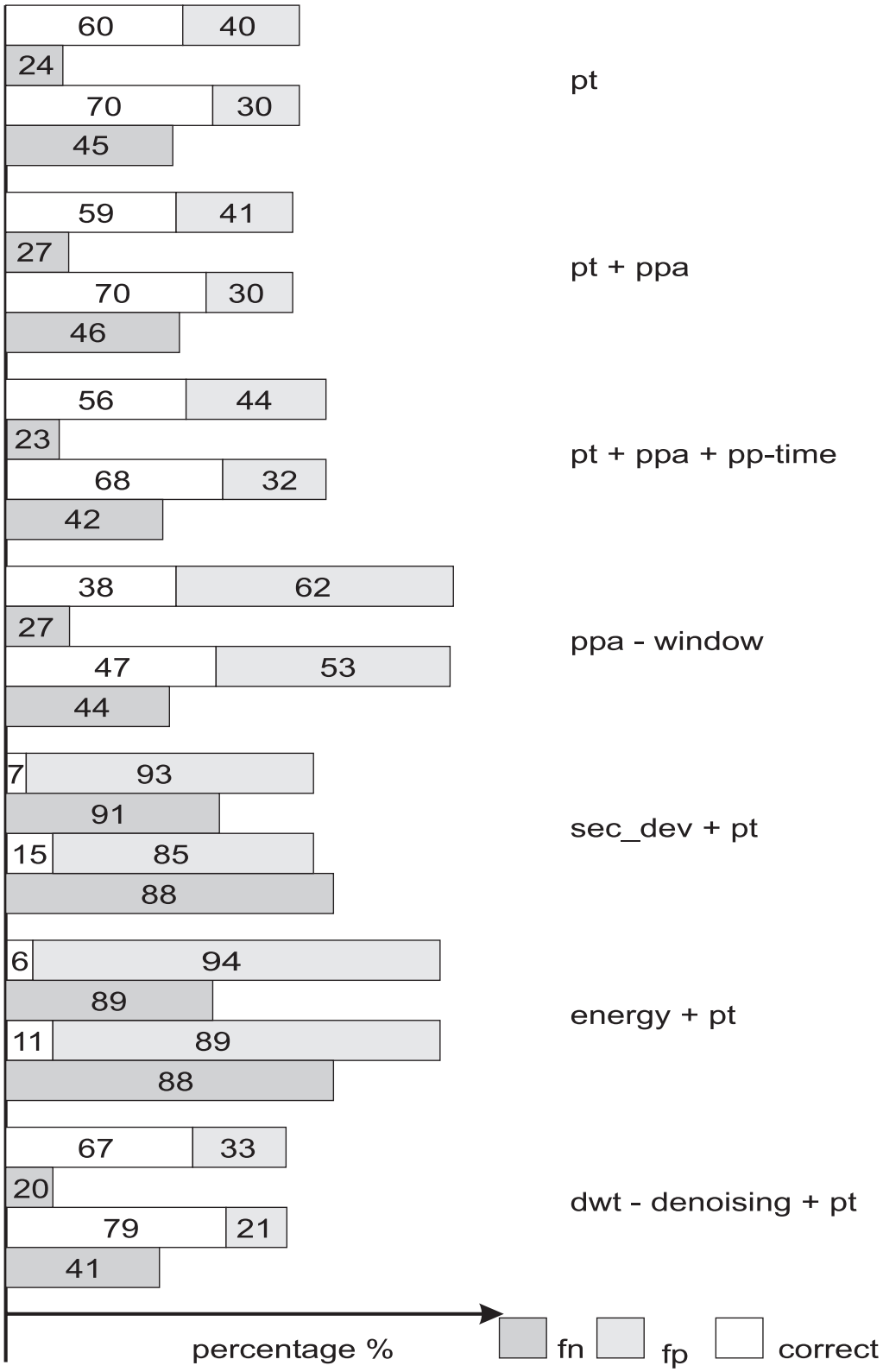

Figure 6: Percentage of correctly detected spikes (correct) and false positives (fp) are calculated on the basis of all detected spikes. False negatives (fn) are given as percentage of 70 or 112 spikes, respectively, see text for methods.

To be published in "5th German Workshop on Artificial Life - Abstracting and Synthesizing the Principles of Living Systems",Eds. D. Polani, J. Kim, T. Martinetz (Eds.), IOS Press Infix.

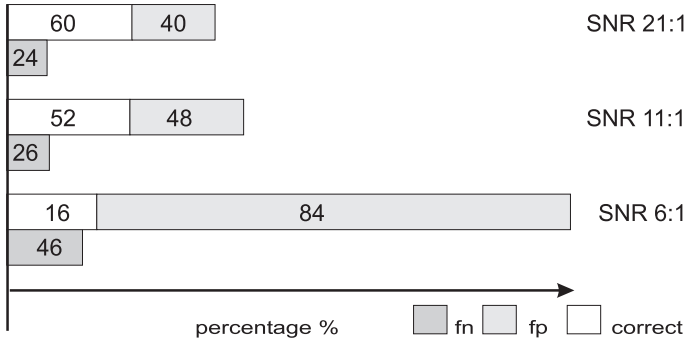

Figure 7: Results of positive threshold method for simulated data with different SNRs.

5 Acknowledgements

We thank Reinoud Maex, Born-Bunge Foudation, University of Antwerp, for his advice concerning the GENESIS simulation.

This work was in part supported by EU-grant IST-1999-10079 and the German research ministry BMBF grant 16SV1433.

References

[1] Moshe Abeles and Moise H. Goldstein. Multispike train analysis. *Proceedings of the IEEE*, 65(5):762–773, 1977.

[2] Amir F. Atiya. Recognition of multiunit neural signals. *IEEE Transactions on Biomedical Engineering*, 39(7):723–729, 1992.

[3] James M. Bower and David Beeman. *The Book of GENESIS*. Springer-Verlag, second edition, 1998.

[4] J. Chapin, K. Moxon, et al. Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex. *nature neuroscience*, 2:664–670, 1999.

[5] J. Fang, G.C. Agarwal, and B.T. Shahani. Decomposition of multiunit electromyographic signals. *IEEE Transactions on Biomedical Engineering*, 46:685–697, 1999.

[6] W. Freiwald, H. Stemman, et al. Stimulus representation in rat primary visual cortex: multi-electrode recordings and estimation theory. in press with Neurocomputing, 2001.

[7] Charles M. Gray, Pedro E. Maldonado, et al. Tetrodes markedly improve the reliability and yield of multiple single-unit isolation from multi-unit recordings in cat striate cortex. *Journal of Neuroscience Methods*, 63:43–54, 1995.

[8] U.G. Hofmann, E. De Schutter, et al. *On the design of multi-site microelectrodes for neuronal recordings*. in MICRO.tec 2000, VDE Verlag, Hannover, 2000.

[9] M. Laubach, J. Wessberg, et al. Cortical ensemble activity increasingly predicts behaviour outcomes during learning of a motor task. *Nature*, 405:567–571, 2000.

[10] Michael S. Lewicki. A review of methods for spike sorting: the detection and classification of neural action potentials. *Network: Computational Neural Systems*, 9:R53–R78, 1998.

[11] K.C. McGill, K.L. Cummins, and L.J. Dorfman. Automatic decomposition of the clinical electromyogram. *IEEE Transactions on Biomedical Engineering*, 32:470–477, 1985.

[12] Kerstin M. L. Menne, Ulrich G. Hofmann, et al. Test of spike sorting algorithms on the basis of simulated network data. In *Neurocomputing, Proceedings of the CNS 2001, Monterey*. To be published.

[13] Kerstin M. L. Menne, Reinoud Maex, and Ulrich G. Hofmann. Extracellular potential data generated in a network simulation. In *The Neurosciences at the Turn of the Century, Proceedings of the 4th Meeting of the German Neuroscience Society 2001*, volume 1, page 602. Georg Thieme Verlag, 2001.

[14] H. Nakatani, T. Watanabe, and N. Hoshimiya. Detection of nerve action potentials under low signal-to-noise ratio condition. *IEEE Transactions on Biomedical Engineering*, 48(8):845–849, 2001.

[15] Miguel A.L. Nicolelis. *Methods for Neural Ensemble Recordings*. CRC Press LLC, 1999.

[16] P. Norlin, M. Kindlundh, et al. A 32-site neural recording probe fabricated by double-sided deep reactive ion etching of silicon-on-insulator substrates. In *12th Micromechanics Europe Workshop (MME-2001), Cork, Ireland*, 2001.

[17] P.L. Nunez. *Electric Fields of the Brain: The Neurophysics of EEG*. Oxford University Press, Oxford, first edition, 1981.

[18] Karim G. Oweiss and David J. Anderson. Noise reduction in multichannel neural recordings using a new array wavelet denoising algorithm. *Neurocomputing*, 38-40:1687–1693, 2001.

[19] Roger D. Traub, John G. R. Jefferys, et al. A branching dendritic model of a rodent CA3 pyramidal neuron. *Journal of Physiology*, 481.1:79–95, 1994.

[20] Roger D. Traub and Richard Miles. Pyramidal cell-to-inhibitory cell spike transduction explicable by active dendritic conductances in inhibitory cell. *Journal of Computational Neuroscience*, 2:291–298, 1995.

[21] Shiro Usui and Itzhak Amidror. Digital low-pass differentiation for biological signal processing. *IEEE Transactions on Biomedical Engineering*, 29:686–693, 1982.

[22] J. Wessberg, C.R. Stambaugh, et al. Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. *Nature*, 408:361–365, 2000.

To be published in *Fifth German Workshop on Artificial Life - Abstracting and Synthesizing the Principles of Living Systems*,D. Polani, J. Kim, T. Martinetz (Eds.), IOS Press Infix.