Spike Sorting Errors: Statistical Differences of Cortical 'Single Unit' and 'Single Neuron' Activity.

Martin Paul Nawrot[1,2], Clemens Boucsein[3,4]

[1] Neuroinformatics and Theoretical Neuroscience, Freie Universität Berlin, Germany
[2] Bernstein Center for Computational Neuroscience Berlin
[3] Neurobiologie und Biophysik, Institut für Biologie III, Albert-Ludwigs-Universität Freiburg, Germany
[4] Bernstein Center for Computational Neuroscience Freiburg

Extracellular recording techniques are the preferred experimental means to monitor neuronal spiking activity in the nervous system of animals. Extracting the so-called 'single-unit' activity (SUA) from the extracellular signal involves two steps. First, 'spikes' that are assumed to reflect action potentials (APs) of nearby neurons are detected, e.g. by thresholding. Second, the spike sorting procedure assigns each individual spike to a particular 'unit'. All spikes of one unit are supposed to correspond to the APs generated by one single neuron. However, detection and sorting are error-prone. Spikes may be falsely assigned to a particular unit (false positives, FPs), and some spikes are missed and not assigned to the respective unit (false negatives, FNs). Indeed, several publications have estimated the amount of errors obtained in the process of spike sorting and FP/FN rates on the order of 10-15% (e.g. Pouzat et al., 2004; Joshua et al., 2007) seem plausible.

We investigated the effect of spike sorting errors on statistical properties of cortical spike trains. In particular we focused on 3 statistical features of cortical spike trains, namely (1) the negative serial correlation of neighboring intervals (Lebedev & Nelson, 1996, Nawrot et al., 2007, Engel et al., 2008), (2) the interval variability as measured by the coefficient of variation (CV), and (3) the spike count variability quantified by the Fano factor. To test the effect of FPs and FNs on these statistical measures, we used two methods of generating surrogate data sets with FPs and FNs. Firstly, we used a point process model with a realistic interval distribution and serial interval correlations (Farkhooi et al., 2008) and randomly inserted or deleted spikes from numeric realizations. Secondly, we used in vivo intracellularly recorded spike trains from cortical neurons (Nawrot et al., 2007) and mixed spikes of different independent recordings.

Our results demonstrate that (1) serial correlation is lost and becomes insignificant for a FP rate of about 10-15%, (2) the coefficient of variation monotonically increases with increasing FP or FN rate, and (3) the Fano factor increases even more strongly than the CV as negative serial correlation is lost which reduces the spike count variance in single neurons (Nawrot et al., 2007; Farkhooi et al., 2008). Thus, we conclude that a realistic rate of spike sorting errors can severely alter the statistics of the true single neuron spike train. Our results suggest that spike sorting may lead to a general over-estimation of single neuron variability, and that it conceals serial spike train statistics that are observed in true single neuron spike trains.

This research receives funding from the BMBF to BCCNs Berlin and Freiburg

Farkhooi, Strub-Blosse, Nawrot (2008) Phys Rev E (under revision)

Lebedev and Nelson (1996) Exp Brain Res 111: 313-325

Joshua, Elias, Levine, Bergman (2007) J Neurosci Meth 163: 267-282
Nawrot, Boucsein, Rodriguez-Molina, Aertsen, Grün, Rotter (2007) Neurocomp 70: 1717-1722

Pouzat, Delescluse, Viot, Diebolt (2004) J Neurophysiol 91: 2910–2928