

# Spike Sorting with R and GGobi

Christophe Pouzat, Andrea Ridolfi and Pascal Viot

February 27, 2006

Neurobiologists are more and more interested in studying neuronal populations. To this aim, one of the most popular experimental techniques is Multi-Electrode Array (MEA) recording which collects the activity of several neurons from each of many electrodes. But, in order to be really informative this technique requires the difficult “spike sorting” problem to be solved. The raw recorded data are indeed a mixture of signals (action potentials or “spikes”) originating from many neurons. Spike sorting techniques aim at un-mixing these signals to get the single neurons spike trains. They in fact require the application of three classical data processing stages: 1) Pre-Processing, the spikes must be detected and the dimension of the space in which they are represented must be made as small as possible. 2) Clustering, individual spikes must be grouped into clusters corresponding to the different neurons whose number is not known a priori. 3) Classification, some detected events correspond in fact to two spikes from two neurons occurring almost simultaneously, they must be identified and classified as such; recordings are moreover long (several hours) and stable, the time consuming stage is therefore typically performed on the first minutes of data and subsequent parts are directly classified.

R and GGobi provide wonderful tools to address all these issues. After several years of software development in pure C interfaced with Scilab and then R, we are presently rewriting our software, SpikeOMatic<sup>1</sup> as an R package using default and contributed packages like mclust, e1071, XML. S4 type of classes and methods allow both easy and intuitive data manipulation, because our objects behave in most cases like matrices, and a full access to R functionalities. Many analysis steps in spike sorting, that is, in clustering and/or classification, can also be trivially parallelized thanks to Rmpi and/or rpvm together with snow. Finally the powerfull data visualization features of R and GGobi provide crucial help to the neurophysiologist trying to understand the method he is using and/or to evaluate the trustworthiness of his results.

---

<sup>1</sup>[http://www.biomedicale.univ-paris5.fr/phycserv/C\\_Pouzat/newSOM/newSOMtutorial/newSOMtutorial.html](http://www.biomedicale.univ-paris5.fr/phycserv/C_Pouzat/newSOM/newSOMtutorial/newSOMtutorial.html)