



Contents lists available at ScienceDirect

Journal of Physiology - Paris

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

Review Paper

Recent progress in multi-electrode spike sorting methods

Baptiste Lefebvre^{a,b,*,#}, Pierre Yger^{a,#}, Olivier Marre^a^a Sorbonne Universités, UPMC Univ Paris 06, INSERM, CNRS, Institut de la Vision, 17 rue Moreau, 75012 Paris, France^b Laboratoire de Physique Statistique, UPMC-Sorbonne Universités, CNRS, ENS-PSL Research University, 24 rue Lhomond, 75005 Paris, France

ARTICLE INFO

Article history:

Received 30 June 2016

Received in revised form 9 November 2016

Accepted 27 February 2017

Available online xxxxx

Keywords:

Electrophysiology

Signal processing

Spike sorting

Template matching

Multi-electrode array

Microelectrode array

ABSTRACT

In recent years, arrays of extracellular electrodes have been developed and manufactured to record simultaneously from hundreds of electrodes packed with a high density. These recordings should allow neuroscientists to reconstruct the individual activity of the neurons spiking in the vicinity of these electrodes, with the help of signal processing algorithms. Algorithms need to solve a source separation problem, also known as spike sorting. However, these new devices challenge the classical way to do spike sorting. Here we review different methods that have been developed to sort spikes from these large-scale recordings. We describe the common properties of these algorithms, as well as their main differences. Finally, we outline the issues that remain to be solved by future spike sorting algorithms.

© 2017 Published by Elsevier Ltd.

Contents

1. Introduction	00
2. The challenge posed by large-scale multi-electrode recordings to classical approaches	00
3. Improvements of the clustering	00
3.1. Improved spike detection	00
3.2. Pre-clustering	00
3.3. Main issues associated with clustering	00
3.3.1. Mathematical definition and non-linear optimization	00
3.3.2. Overlapping spikes	00
4. Template matching approaches	00
4.1. Template extraction	00
4.2. Finding the spike trains	00
4.3. Approaches with binary amplitudes	00
4.4. Approaches with graded amplitudes	00
4.5. Different algorithms correspond to different assumptions about spike amplitude distributions	00
4.6. Caveats when minimizing an objective function	00
4.7. Assumptions behind the template decomposition	00
5. Conclusion: challenges ahead	00
Acknowledgments	00
References	00

1. Introduction

Progress in neuroscience relies to a large extent on the ability to record simultaneously from large populations of cells, in order to understand how information is represented among neurons. One of the most popular techniques to measure such an activity is

* Corresponding author.

E-mail address: baptiste.lefebvre@ens.fr (B. Lefebvre).

These authors contributed equally to this work.

the use of arrays of extracellular electrodes. With these devices, each electrode records the extracellular field in its vicinity and can detect the action potentials emitted by the neighboring neurons. In contrast to intracellular recording, those extracellular recordings do not give a direct access to the neuronal activity: one needs to process the recorded signals to extract the spikes emitted by the different cells around the electrode. This process is termed *spike sorting*, and many algorithms have been suggested to do it efficiently (see Lewicki (1998) or Rey et al. (2015) for a review).

The first extracellular recordings were performed with a single electrode, and could only give access to 3-5 neurons (Gerstein and Clark, 1964). A recent study (Pedreira et al., 2012) highlighted that the maximal number of accessible neurons should lie between 8 and 10 in that case. Over the last decades, there has been a strong effort to increase the number of electrodes, and therefore the number of recorded neurons. Spike sorting algorithms had to be adapted to process this increasingly large amount of data. At first, electrodes were spaced by hundreds of microns such that the spike of one cell could only be detected on a single electrode (Jones et al., 1992; Shoham et al., 2003). In that case, spike sorting on a large amount of electrodes could simply be done by processing each electrode independently. The parallelization of the problem for large amount of independent electrodes was relatively easy to address.

However, devices where electrodes are packed with a high density have also been developed. The spacing between electrodes is much smaller (tens of microns). As a consequence, a spike from a single cell can be detected on several electrodes. Conversely, each electrode will detect the activity of many cells, a property already encountered in the case of single electrode. This increased density helps a lot to resolve single cells (Gray et al., 1995; Franke et al., 2015a), but electrode signals could not be processed independently. Spike sorting algorithms had to be adapted to this new type of data. While for small numbers of electrodes (e.g. tetrodes), methods that could be seen as adaptations of single electrode sorting worked very well (McNaughton et al., 1983; Harris et al., 2000; Gao et al., 2012), this is not the case with new devices designed with hundreds of electrodes all densely packed. CMOS-based devices with thousands of electrodes have been tested and are now frequently used (Berdondini et al., 2005; Fiscella et al., 2012; Müller et al., 2015; Hilgen et al., 2016), calling for new algorithmic methods, largely different from the usual sorting methods.

Here we review the different spike sorting algorithms that have been proposed to process recordings from these novel high-density devices. We will first explain the limitations of classical spike sorting approaches to process these large-scale, dense recordings. Then, we will outline the main changes introduced by these new algorithms compared to classical spike sorting approaches. We will emphasize that most of these new methods follow the same global strategy, although they have been developed independently by different groups. Therefore, we will outline the common properties shared by these algorithms, before explaining and discussing their main differences. Finally, we will discuss the issues that still need to be resolved by future spike sorting algorithms.

2. The challenge posed by large-scale multi-electrode recordings to classical approaches

Most of the classical approaches to spike sorting can be decomposed in two main steps. First, some specific features of the spike waveforms are extracted from the raw data. This allows each spike to be characterized by a small set of numbers/features. Using these

features, each spike can now be seen as a point in a low dimension space, and the second step consists in clustering all the points in this reduced space.

For the first step, earliest methods only extracted the spike amplitude (Hubel, 1957), and width (Meister et al., 1994) of each spike. More recently, some methods use the full waveform directly when the number of electrodes remains small (Pouzat et al., 2002). Another standard technique is to project each waveform on a set of basis functions (Litke et al., 2004; Quiroga et al., 2004), that are either found by performing a principal component analysis (PCA) on the entire set of waveforms (Egert et al., 2002; Pouzat et al., 2002; Einevoll et al., 2012; Swindale and Spacek, 2015), or by choosing a wavelet basis (Letelier and Weber, 2000; Hulata et al., 2002; Quiroga et al., 2004). For a comparison between PCA and wavelet based analysis, see Pavlov et al. (2007). Note that the two can be combined (Bestel et al., 2012).

Once the dimensionality has been reduced, to tackle the problem of the clustering step, several approaches have been used, but the most standard approach is to fit the clusters with a mixture of Gaussians (Wood et al., 2004; Rossant et al., 2016; Kadir et al., 2014). However, one could also find in the literature approaches such as paramagnetic clustering (Quiroga et al., 2004), mean-shift clustering (Swindale and Spacek, 2014) or even k -means clustering (Atiya, 1992; Chah et al., 2011). Another interesting approach is to consider the most consensual clustering across an ensemble of k -means solutions (Fournier et al., 2016).

Not all standard methods strictly follow this workflow. For example, linear filtering is an alternative approach which identifies the optimal linear filter to distinguish one signal, of unknown temporal position but of known waveform, from a finite number of other signals of known waveforms, observed on noisy electrodes. This approach was first proposed by Roberts and Hartline (1975), then by Gozani and Miller (1994) and more recently by Franke et al. (2010). This method is similar to template matching approaches that we will describe later. An alternative approach is independent component analysis (ICA) where the first step demix blindly the data and extract the individual source signals from which spikes are identified (Takahashi et al., 2003; Brown et al., 2001; Jäckel et al., 2012). Note that variants, such as the convolutional independent component analysis (cICA) of Leibig et al. (2016), has been developed. However, there is no guarantee that the independent components found by those algorithms are indeed isolated neurons.

While all of these methods can be successful when one electrode captures the signals from a only few cells, and when one cell is only recorded by one or a small number of electrodes, it is not trivial to scale them up to process a large number of densely packed electrodes. In recordings performed by large and dense multi-electrode arrays, the spike waveforms live in a high dimensional space, and this makes the clustering challenging. We will review below some suggested improvements to enable clustering on a large number of electrodes.

Finally, a more fundamental problem with clustering-based approach is that the extraction of features from one spike can be distorted by the presence of other spikes nearby. As a consequence, most of the overlapping spikes are not captured by clustering approaches, because they correspond to points in the feature space that are far from the centers of the corresponding clusters. This is a major challenge for clustering techniques (Bar-Gad et al., 2001), that we will explain in more details below. In large scale and dense multi-electrode recordings, overlapping spikes become the rule rather than the exception. Solving this issue is one of the motivation behind new algorithms, based on template matching, that we will review and discuss in a second part.

Download English Version:

<https://daneshyari.com/en/article/5593250>

Download Persian Version:

<https://daneshyari.com/article/5593250>

[Daneshyari.com](https://daneshyari.com)