Open-Source Tools for Processing and Analysis of In Vitro Extracellular Neuronal Signals

Mufti Mahmud and Stefano Vassanelli

Abstract The recent years have seen unprecedented growth in the manufacturing of neurotechnological tools. The latest technological advancements presented the neuroscientific community with neuronal probes containing thousands of recording sites. These next-generation probes are capable of simultaneously recording neuronal signals from a large number of channels. Numerically, a simple 128-channel neuronal data acquisition system equipped with a 16 bits A/D converter digitizing the acquired analog waveforms at a sampling frequency of 20 kHz will generate approximately 17 GB uncompressed data per hour. Today's biggest challenge is to mine this staggering amount of data and find useful information which can later be used in decoding brain functions, diagnosing diseases, and devising treatments. To this goal, many automated processing and analysis tools have been developed and reported in the literature. A good amount of them are also available as open source for others to adapt them to individual needs. Focusing on extracellularly recorded neuronal signals in vitro, this chapter provides an overview of the popular open-source tools applicable on these signals for spike trains and local field potentials analysis, and spike sorting. Towards the end, several future research directions have also been outlined.

Keywords Neuroengineering · Neuronal activity · Neuronal spikes · Local field potentials · Neuronal signal processing and analysis

---

M. Mahmud (✉)
Computing and Technology, School of Science and Technology, Nottingham Trent University, Nottingham, UK
e-mail: mufti.mahmud@ntu.ac.uk

S. Vassanelli
NeuroChip Lab, Department of Biomedical Sciences, University of Padova, Padova, Italy

M. Chiappalone et al. (eds.), *In Vitro Neuronal Networks*,
Advances in Neurobiology 22, https://doi.org/10.1007/978-3-030-11135-9_10

1 Introduction

Recent technological advancements allowed scientists to have unprecedented access to biological data. These data come from different organs in the body and have been effectively utilized in screening and diagnosis of diseases and their treatment (Mahmud et al. 2018). Brain is the most complex organ in the mammals. Though it is analog, its capability in decision-making and pattern recognition is higher than any existing computing machines (Mahmud et al. 2017). To understand brain's functionality, diagnose disease, and devise treatments, scientists have been investigating it using different approaches (Mahmud and Vassanelli 2016). In last two decades, micro- and nanotechnology underwent an exponential growth in terms of developing novel miniaturized devices and this allowed neuroscientists to target large populations of neurons and record from them to decode the activities of the brain cells (Mahmud et al. 2017; Vassanelli 2011; Vassanelli et al. 2012b,a; Vassanelli 2014; Schröder et al. 2015; Thewes et al. 2016; Jun et al. 2017). However, these novel techniques to acquire neuronal signals generate huge amount of data. And, analyzing this data and mining relevant information is a big challenge. To this goal, individual research groups have contributed towards the development of automated, efficient, and intelligent processing methods and disseminated them to the neuroscientific community (Mahmud and Vassanelli 2016). The interdisciplinary "Neuroengineering" community (Vassanelli and Mahmud 2016) used these tools to mine useful information from these large datasets (Mahmud and Vassanelli 2016). Targeting different applications and needs, these methods deal with processing and analysis of data coming from single or multiple channels. However, with today's increasing number of recording sites accommodated in a single probe, many of these methods are difficult to rescale and fit to analyze these data. Therefore, the community is still in need of novel analysis tools targeting multichannel neurophysiological data coming from high-resolution neuronal probes. This chapter aims in introducing the reader with available open-source toolboxes capable of performing processing and analysis of multichannel extracellular neuronal signals recorded in vitro.

2 State of the Art of Extracellular Neuronal Signal Analysis

Modern neuroscience research has emerged as a data-driven discipline where both experimental and computational approaches go hand-in-hand (see Fig. 1) (Mahmud et al. 2018). The two approaches of neuroscience research are bridged through a relatively new discipline, called data science, which mainly deals with the analytics of the acquired data, and interpretation of the simulated results and design novel experiments suggested by the obtained results.

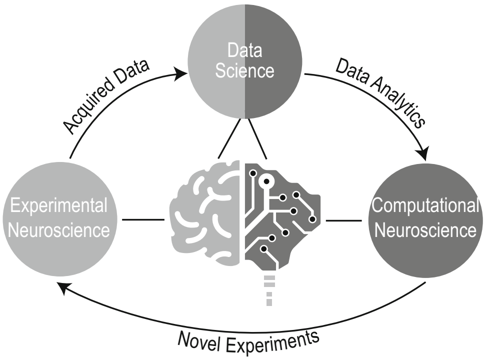

Fig. 1 Overview of modern neuroscience research which comprises of computational and experimental approaches to neuroscience through the data science domain

Focusing on the extracellular neuronal signals acquired in vitro, this chapter is going to deal with the data analytics part, and provide an overview of the available open-source data analytics resources which can be exploited to process and analyze the recorded signals. Using the in vitro experimental paradigm (see Fig. 2), mainly two types of signals can be recorded: (1) neuronal spikes and their trains, and (2) field potentials. Therefore, the following subsections will contain the state-of-the-art open-source resources categorized by signal types, e.g., field potentials, and spikes. Despite that the majority of the available resources are application and signal specific, there exist a few toolboxes which provide methods applicable on multiple types of signals. For the sake of simplicity, the tools are divided into following two categories:

- Toolboxes for spike trains and field potentials analysis and
- Toolboxes for spike sorting

It is also worthy to note that majority of the available toolboxes are developed using MATLAB (Mathworks Inc., Natick, USA; www.mathworks.com) and python (www.python.org) programming languages due to their diffused usage in the neuroscience community. In addition, popular programming languages including C, C++, C#, Delphi7, Java, and R have also been employed in some of the toolboxes.

2.1 Toolboxes for Spike Trains and Field Potentials Analysis

With the growing amount data acquired via simultaneously recorded channels from an increasing number of neurons, the neuroengineering community has developed automated toolboxes addressing the required processing and analyses. The following subsections describe—in alphabetical order—popular publicly available toolboxes. Table 1 summarizes the different packages with their representative features.

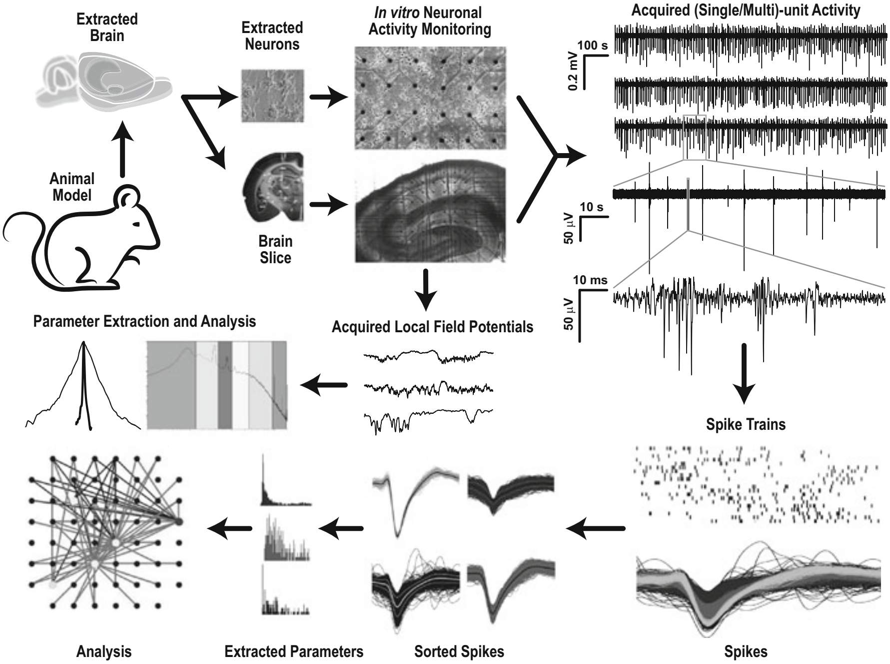

Fig. 2 Exemplary overview of the in vitro neuronal signal acquisition and analysis paradigm

2.1.1 Brain System for Multivariate AutoRegressive Time Series (BSMART)

"BSMART" toolbox is written in MATLAB/C for spectral analysis of neurophysiological signals (Cui et al. 2008). It provides multi- or bivariate AutoRegressive modeling, spectral analysis through coherence and Granger causality, and network analysis. The main analyses include: adaptive (bi-/multi-)variate autoregressive model, fast Fourier transform, Granger causality (Granger 1969), coherence, coherent network analysis, and Granger causality network analysis. Available at http://www.brain-smart.org/.

2.1.2 Chronux

"Chronux" toolbox is developed in MATLAB for the analysis of both point process and continuous data (Bokil et al. 2010). It provides spike sorting, and local regression and multitaper spectral analysis of neural signals. The main analyses include: hierarchical clustering method (Fee et al. 1996), locally weighted sum of squares (Cleveland 1979), local regression fitting and density estimation (Loader 1999), multitaping method (Thomson 1982), coherence, and spike field coherence. Available at http://chronux.org/.

Table 1 Popular spike train and field potential processing and analysis toolboxes with their representative features

Toolbox	Features					
	Lang	PF	GUI DV	DIE	AR	PDP
BSMART	Matlab/C	OSML	Yes	Yes	No	No
Chronux	Matlab	LMW	Yes	No	No	No
DATA-MEAns	Delphi7	W	Yes	No	No	No
FIND	Matlab	OSML	Yes	Yes	No	No
ibTB	Matlab	LMW	No	No	No	No
KNSNDM	C++	LMW	Yes	Yes	Yes	No
MeaBench	C++/Matlab	L	Yes	No	Yes	No
MVGC	Matlab	OSML	No	No	No	No
nSTAT	Matlab	OSML	No	No	No	No
PANDORA	Matlab	LMW	No	Yes	No	EMP
QSpikeTools	Matlab	ML	No	No	Yes	EMP
SigMate	Matlab	OSML	Yes	Yes	Yes	No
sigTOOL	Matlab	OSML	Yes	Yes	No	No
SpiCoDyn	C#	W	Yes	Yes	No	No
SPKTool	Matlab	OSML	Yes	Yes	No	No
STAToolkit	Matlab/C	LMW	Yes	No	No	Yes
ToolConnect	C#	W	Yes	Yes	No	No

Lang Language; *PF* Platform; *GUI DV* GUI and data visualization; *DIE* Data import/export; *AR* Artifact removal; *PDP* Parallel data processing; *KNSNDM* Klusters, NeuroScope, NDManager; *L* Linux; *U* Unix; *M* Mac; *W* Windows; *OSML* Operating system supported by Matlab; *EMP* Embarrassingly parallel

2.1.3 DATA-MEAns

"DATA-MEAns" is a toolbox developed in Borland Delphi 7 (Embarcadero Technologies Inc., Austin, USA) and MATLAB (Bonomini et al. 2005). It provides data visualization, basic analysis (i.e., autocorrelations, perievent histograms, rate curves, PSTHs, ISIs, etc.), and nearest neighbor or k-means clustering. The analyses include: poststimulus time and perievent histogram estimation, auto- and cross-correlation, Fano factor and coherence calculation, event synchrony (Quian Quiroga et al. 2002), and nearest neighbor (Cover and Hart 1967) and K-means (MacQueen 1967) clustering. Available at http://cortivis.umh.es/.

2.1.4 Finding Information in Neural Data (FIND)

"FIND" is a platform-independent framework for the analysis of neuronal data based on MATLAB (Meier et al. 2008). It provides a unified data import function from various proprietary formats simplifying standardized interfacing with analysis tools and provides means for analysis of discrete series of spike events, continuous time series, and imaging data. Also, it allows simulating multielectrode activity using point-process-based stochastic model. The analyses include: co-variance estimation, point process modeling, pair-wise cross-correlation, asymmetric Savitzky–Golay filter calculation (Savitzky and Golay 1964), response latency differences estimation (Nawrot et al. 2003), and spike detection. Available at http://find.bccn.uni-freiburg.de/.

2.1.5 Information Breakdown ToolBox (ibTB)

“ibTB” is a MATLAB-based toolbox which implements information theory methods for spike, LFP, and EEG analysis (Magri et al. 2009). It provides information breakdown technique to decode the encoding of sensory stimuli by different groups of neurons. The major supported analyses include: direct method, quadratic extrapolation (Strong et al. 1998), Panzeri and Treves method (Panzeri and Treves 1996) for bias correction, shuffling procedure (Montemurro et al. 2007), bootstrap bias correction (Optican et al. 1991), and Gaussian method (Misra et al. 2005). The source code can be obtained from the publisher’s website (http://static-content.springer.com/esm/art%3A10.1186%2F1471-2202-10-81/MediaObjects/1471-2202-10-81-S1.zip).

2.1.6 Klusters, NeuroScope, and NDManager

“Klusters,” “NeuroScope,” and “NDManager” are three integrated modules bundled together for processing and analysis of spike and field potential signals (Hazan et al. 2006). Klusters performs spike sorting using KlustaKwik (see Sect. 2.2.3) and displays 2D projection of features, spike traces, correlograms, and error matrix view. NeuroScope allows inspection, selection, and event editing of spike signals as well as local field potentials (LFPs). NDManager facilitates experimental and preprocessing parameter management. The major analyses include: auto- and cross-correlation estimation, spike detection and sorting, and classification expectation–maximization (Celeux and Govaert 1992). Available at http://neurosuite.sourceforge.net/.

2.1.7 MeaBench

“MeaBench” is a toolbox written mainly in C++ with certain parts written in Perl[1] and MATLAB. It is intended for data acquisition and online analysis of commercial multielectrode array recordings from Multichannel Systems GmbH (Reutlingen, Germany) (Wagenaar et al. 2005). It allows real-time data visualization, line and

---

[1] https://www.perl.org/.

stimulus artifact suppression, and spike and burst detection and validation. Available at www.danielwagenaar.net/res/software/meabench/.

2.1.8 Multivariate Granger Causality Toolbox (MVGC)

"MVGC" is a toolbox written in MATLAB that implements Wiener–Granger causality (G-causality) on multiple equivalent representations of a vector autoregressive model in both time and frequency domains (Barnett and Seth 2014). The main analyses supported by the toolbox include: ordinary least squares estimation, Wiener (root mean square) method (Levinson 1946), vector autoregressive maximum likelihood estimator method, cross-power spectral density, multitaper method, fast Fourier transform, and unconditional Granger causality. It can be applied to neuroelectric, neuromagnetic, and fMRI signals and can be obtained from http://www.sussex.ac.uk/sackler/mvgc/.

2.1.9 nSTAT

"nSTAT" toolbox is coded in MATLAB and performs spike train analysis in time domain (e.g., Kalman Filtering), frequency domain (e.g., multitaper spectral estimation), and mixed time–frequency domain (e.g., spectrogram) (Cajigas et al. 2012). The supported analyses include: point process generalized linear model (Paninski et al. 2007), generalized linear model-based peristimulus time histogram estimation, Akaike's and Bayesian information criteria, state-space generalized linear model, Kalman filtering, multitaper method, and spectrogram. Available at www.neurostat.mit.edu/nstat/.

2.1.10 PANDORA

"PANDORA" is a MATLAB-based toolbox that extracts user-defined characteristics from spike train signals and creates numerical database tables from them (Gunay et al. 2009). Further analyses (e.g., drug and parameter effects, spike shape characterization, histogramming and comparison of distributions, cross-correlation, etc.) can then be performed on these tables. However, spike detection and feature extraction can also be performed. The supported analyses include: rational database creation from datasets, extraction of spike shape characteristics, Kullback–Leibler divergence measure (Kullback and Leibler 1951) estimation, and resistor-average distance (Johnson et al. 2001) estimation. It is available at https://github.com/cengique/pandora-matlab.

2.1.11 QSpike Tools

"QSpike Tools" is a Linux/Unix-based cloud-computing framework, modeled using client–server architecture and developed in MATLAB/Bash scripts,[2] for processing and analysis of extracellular spike trains (Mahmud et al. 2014). It performs batch preprocessing of CPU-intensive operations for each channel (e.g., filtering, multiunit activity detection, spike sorting, etc.), in parallel, by delegating them to a multicore computer or to a computers cluster. The main analyses include: spike detection and validation, poststimulus time and perievent histogram estimation, burst detection and validation, and spike sorting through Wave_Clus package (see Sect. 2.2.11). It can be obtained from https://sites.google.com/site/qspiketool/.

2.1.12 SigMate

"SigMate" is a MATLAB-based comprehensive framework that allows preprocessing and analysis of EEG, LFPs, and spike signals (Mahmud et al. 2012a). Its main contribution is in the analysis of LFPs which includes data display, file operations, baseline correction, artifact removal, noise characterization, current source density (CSD) analysis, latency estimation from LFPs and CSDs, determination of cortical layer activation order using LFPs and CSDs, and single LFP clustering. The main processing and analyses include: various file operations (e.g., file splitting, concatenation, and column rearranging), latency calculation (Mahmud et al. 2016), detection of cortical layer activation order (Mahmud et al. 2010), current source density analysis (Mahmud et al. 2011), classification of single-trial LFPs (Mahmud et al. 2012c), and spike analysis. The spike analyses are provided through Wave_Clus (see Sect. 2.2.11) toolbox. It can be obtained from https://sites.google.com/site/muftimahmud/codes.

2.1.13 sigTOOL

"sigTOOL" toolbox is written in MATLAB and allows direct loading of a wide range of proprietary file formats (Lidierth 2009). The usable data file formats originate from various hardware vendors such as Alpha Omega, Axon Instruments, Blackrock Microsystems, Cambridge Electronic Design, Heka, MultiChannel Systems, NeuroExplorer, NeuroShare native, and Plexon. The major analyses supported by the toolbox include: auto- or cross-correlation, power spectral analysis, and coherence estimation in addition to usual spike train analysis (i.e., ISI, event auto- and cross-

---

[2] https://en.wikipedia.org/wiki/Bash_(Unix_shell).

correlations, spike-triggered averaging, perievent time histograms, frequencygrams, etc.). Available at http://sigtool.sourceforge.net/.

2.1.14 SpiCoDyn

"SpiCoDyn" is an open-source windows-only graphical user interface-based toolbox focusing on functional-effective connectivity analysis and spiking and bursting dynamics analysis (Pastore et al. 2018). Developed in C#, under the Microsoft .NET platform, it supports HDF5,[3] level 5 MAT files[4] and text files. The toolbox provides optimized implementations of two main transfer entropy algorithms (i.e., delayed transfer entropy and high-order transfer entropy) and provides analysis platform for multiple spike trains originating from large number of electrodes. Available at https://www.nitrc.org/projects/spicodyn/.

2.1.15 Spike Train Analysis Toolkit (STAToolkit)

"STAToolkit" is a MATLAB/C-hybrid toolbox implementing information theoretic methods to quantify how well the stimuli can be distinguished based on the timing of neuronal firing patterns in a spike train (Goldberg et al. 2009). The main analyses include: direct method (Strong et al. 1998), metric space method (Victor and Purpura 1997), binless method (Victor 2002), asymptotically debiased method (Treves and Panzeri 1995), Jackknife debiased method (Thomson and Chave 1991), debiased Ma bound method (Ma 1981), best upper bound method (Paninski 2003), coverage-adjusted method (Chao and Shen 2003), and Bayesian with Dirichlet prior (Wolpert and Wolf 1995). Available at http://neuroanalysis.org.

2.1.16 SPKTool

"SPKTool" is coded in MATLAB for the detection and analysis of neural spiking activity (Liu et al. 2011). It performs spike detection, feature extraction, and manual and semiautomatic clustering of spike trains. Spike detection by thresholding and raw and nonlinear energy of signal, extraction of various spike features (e.g., principal components, peaks, valleys, energy, timestamps, slice of waveforms, etc.), implementation of popular spike-sorting techniques (e.g., K-means, template matching, EM of Gaussian mixed model, valley seeking, manual contour based method, etc.), ISI, Poincare maps, correlation, firing rate histograms, and perievent rasters and histograms. Available at http://spktool.sourceforge.net/.

---

[3] https://support.hdfgroup.org/HDF5/.

[4] https://www.mathworks.com/help/pdf_doc/matlab/matfile_format.pdf.

2.1.17 ToolConnect

"ToolConnect" is a standalone windows-only tool, developed in C#, targeting to infer functional connectivity from spike trains data (Pastore et al. 2016). This toolbox has been optimized for in vitro networks grown on high-density multielectrode arrays. The implemented analyses include: cross- and partial correlation, and joint and transfer entropy. Additionally, it contains several add-ons for the visualization of the functional connectivity graphs as well as extraction of topological features of the network. Available at https://www.nitrc.org/projects/toolconnect/.

2.2 Toolboxes for Spike Sorting

A great amount of efforts have been put towards the development of sophisticated tools capable of accurate spike sorting and analysis. Rey et al. (2015), in their review, outline the primary concepts of spike sorting, the requirements for applying specific techniques, and most importantly, the shortcomings of currently available algorithms. Due to the vast amount of methods and tools available for the purpose, we restrict our discussion only to the popular open-source toolboxes.

2.2.1 EToS

"EToS" or Efficient Technology of Spike sorting is a spike-sorting toolbox aimed mainly at heterogeneous neural population recordings (Takekawa et al. 2012). It is written in C++ implementing multimodality-weighted PCA for feature extraction and variational Bayes for student's t mixture model for clustering. The spike-sorting code is parallelized through OpenMP (www.openmp.org) and available at http://etos.sourceforge.net/.

2.2.2 KiloSort

"KiloSort" is a spike-sorting package developed in MATLAB for handling data coming from MEAs with large number of recording electrodes (Pachitariu et al. 2016). The package uses template matching for detection and clustering of spikes and can handle batch processing using both GPUs and CPUs. The source code can be downloaded from https://github.com/cortex-lab/KiloSort.

2.2.3 KlustaKwik

"KlustaKwik" is a standalone program written in C++ for automatic clustering analysis (Harris et al. 2000) of high-dimensional spiking data by fitting a mixture of Gaussians and hard expectation–maximization algorithm with unconstrained covariance matrices (Kadir et al. 2014; Rossant et al. 2016). The package can be downloaded from https://github.com/klusta-team/klustakwik.

2.2.4 MClust

"MClust" is a spike-sorting toolbox developed in MATLAB. It supports both manual and automated clustering with possibility to manual feature selection (Redish 2014) on data recorded from single electrode, stereotrode, and tetrode. It allows manual corrections to automated clustering results and can be obtained from http://redishlab.neuroscience.umn.edu/MClust/MClust.html.

2.2.5 NEV2lkit

"NEV2lKit" is a package written in C++ with routines for analysis, visualization, and classification of spikes (Bongard et al. 2014). It is a preprocessor for the analysis of intra- and extracellular spiking neuronal signals. The main purposes served by the toolbox includes: loading ASCII files as well as neural event files (NEV) and extract spike events from them, PCA-based sorting of spikes based on the spike waveform, and saving of spike timestamps, unit information, and the spike signals into ASCII or NEV files. It provides accurate, efficient, and consistency across experiments. Available at http://nev2lkit.sourceforge.net/.

2.2.6 OSort

"OSort" is a template-based, unsupervised, online spike-sorting algorithm written in MATLAB (Rutishauser et al. 2006). It uses residual-sum-of-squares-based distance method and custom thresholds to on-the-fly sort of the recorded spikes. As the algorithm is online, the tool uses a technique where the clusters are built and adapted iteratively over the course of the recording. Available at http://www.urut.ch/new/serendipity/index.php?/pages/osort.html.

2.2.7 SpikeOMatic

"SpikeOMatic" is a spike-sorting package developed in R (Pouzat and Chaffiol 2009). It implements two sophisticated data generation models, namely Gaussian mixture model (GMM) and dynamic hidden Markov model (DHMM). For statistical inference for the abovementioned models, the tool makes use of expectation–maximization for GMM and Markov chain Monte Carlo method for DHMM. The package can be downloaded from http://www.biomedicale.univ-paris5.fr/SpikeOMatic/.

2.2.8 Spyke

“Spyke” is a python toolbox for visualizing, navigating, and spike sorting of high-density multichannel extracellular spikes (Spacek et al. 2009). It uses PCA for dimensionality reduction and modified gradient ascent clustering algorithm (Fukunaga and Hostetler 1975; Swindale and Spacek 2014) to classify the features. The sorting method initially assigns each event to a channel and then these channel-based clusters are subdivided into possible distinguished clusters. Available at http://spyke.github.io/.

2.2.9 SpyKING CIRCUS

“SpyKING CIRCUS” is a python toolbox aiming to provide spike sorting for high-density multichannel extracellular spikes (Yger et al. 2018). This semiautomatic spike-sorting package performs highly parallel code execution to handle large number of recording electrodes. Based on a greedy template matching approach and with the help of a smart clustering technique, the package can efficiently sort spikes from up to 4225 channels. The code is available at https://github.com/spyking-circus/spyking-circus with a ground truth dataset at https://zenodo.org/record/1205233/export/hx#.WrORP3XwaV4.

2.2.10 UltraMegaSort2000

“UltraMegaSort2000” is a MATLAB-based toolbox for spike detection and clustering which implements a hierarchical clustering scheme using similarities of spike shape and spike timing statistics, and provides false-positive and false-negative errors as quality evaluation metrics (Fee et al. 1996; Hill et al. 2011). The toolbox also provides the users with tools to manually correct the automatically generated clusters. Available at http://physics.ucsd.edu/neurophysics/software.php.

2.2.11 Wave_Clus

“Wave_Clus” is probably the most popular spike-sorting package to date. Developed in MATLAB, it uses sophisticated wavelet transformation-based time–frequency analysis for feature selection and a “temperature”-based superparamagnetic clustering (Blatt et al. 1996) method to sort the features into different clusters (Quian Quiroga et al. 2004). It is available for downloading at https://github.com/csn-le/wave_clus.

3 Future Perspectives

In terms of offline processing and analysis of extracellular neuronal signals, the neuroscientific community has seen active participation and contribution from diverse disciplines where individual laboratories developed novel tools and techniques and shared them with the community. As a result of these great efforts, now it has been possible to establish some standardized analyses from these available tools. However, it is noteworthy that, to date, only a few tools are available to the community which are capable of doing the extracellular neuronal signal processing and analysis online which is huge bottleneck for future generation closed-loop real-time systems (Mahmud and Vassanelli 2016).

Also, keeping pace with the technological advancements and speed of the Internet, next-generation online distributed processing and analysis tools are becoming increasingly essential. Exploiting the existing institutional infrastructures, it might be possible to transform the interinstitutional competition into multi-institutional collaboration. In such scenario, a vital requirement is to have secure infrastructures where multi-institutional neuronal signal analysis facilities shall be capable of integrating their data and tools seamlessly (Mahmud et al. 2012b) with the option to protect individual institute's proprietary data. Additionally, these infrastructures should support user-friendly interfaces allowing even experimentalists with minimal information technology skills to explore, navigate, and use provided scientific data and services. Towards these goals, cloud computing and service-oriented architectures might be utilized through the distributed infrastructure. These approaches allow better representation of responsibilities taken by the different users in accordance to their granted privileges. As an early example of such systems, the Spike-Sorting Evaluation Project initiative (http://spike.g-node.org/) aims to lay a platform for evaluating the performance of spike-sorting algorithms through sharing benchmark data and receiving spike-sorting results for comparison.

Furthermore, considering the growing usage of in vitro MEA technology in applications contributing to improved quality of life for patients, such as pharmacological screening and stem-cell-derived neuronal cultures (Tanskanen et al. 2018), it is becoming increasingly important to standardize experimental protocols and analysis procedures, and data and code sharing to foster experimental reproducibility and validation of obtained results for a combined and accelerated maturation of groundbreaking discoveries.

In our opinion, the development is expected towards:

- Development of novel tools capable of performing online analyses,
- Design and implementation of secure and protected systems,
- Advance on cloud-based web applications,
- Facilitate easy deployment of data,
- Reusability and sharing of tools with adaptability to changing requirements, and
- Empower researchers to share data and functionalities that they want to publish.

4 Conclusion

With the tremendous growth of neurotechnologies, scientists can acquire data from multiple levels and multiple sources. This poses a great challenge to the neuroscientific community to automatically process and analyze those data in order to find meaningful conclusions towards understanding brain's functioning and to devise translatable technologies towards autonomous diagnosis and treatment strategies for treating brain diseases. This chapter introduced the reader to the popular open-source automated methods for processing and analysis of extracellularly recorded neuronal signals from in vitro devices. Towards the end, some perspective research lines—where future developments are expected—have also been outlined.

References

Barnett, L., & Seth, A. K. (2014). The MVGC multivariate Granger causality toolbox: A new approach to Granger-causal inference. *Journal of Neuroscience Methods, 223*, 50–68.

Blatt, M., Wiseman, S., & Domany, E. (1996 April). Superparamagnetic clustering of data. *Physical Review Letters, 76*, 3251–3254.

Bokil, H., Andrews, P., Kulkarni, J. E., Mehta, S., & Mitra, P. P. (2010). Chronux: A platform for analyzing neural signals. *Journal of Neuroscience Methods, 192*(1), 146–151.

Bongard, M., Micol, D., & Fernandez, E. (2014). NEV2lkit: A new open source tool for handling neuronal event files from multi-electrode recordings. *International Journal of Neural Systems, 24*(4), 1450009.

Bonomini, M. P., Ferrandez, J. M., Bolea, J. A., & Fernandez, E. (2005). DATA-MEAns: An open source tool for the classification and management of neural ensemble recordings. *Journal of Neuroscience Methods, 148*(2):137–146.

Cajigas, I., Malik, W. Q., & Brown, E. N. (2012). nSTAT: Open-source neural spike train analysis toolbox for Matlab. *Journal of Neuroscience Methods, 211*(2), 245–264.

Celeux, G., & Govaert, G. (1992). A classification EM algorithm for clustering and two stochastic versions. *Computational Statistics and Data Analysis, 14*(3), 315–332.

Chao, A., & Shen, T. J. (2003). Nonparametric estimation of Shannon's index of diversity when there are unseen species in sample. *Environmental and Ecological Statistics, 10*(4), 429–443.

Cleveland, W. S. (1979), Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association, 74*(368), 829–836.

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory, 13*(1), 21–27.

Cui, J., Xu, L., Bressler, S. L., Ding, M., & Liang, H. (2008). BSMART: A Matlab/C toolbox for analysis of multichannel neural time series. *Neural Networks, 21*(8), 1094–1104.

Fee, M. S., Mitra, P. P., & Kleinfeld, D. (1996). Automatic sorting of multiple unit neuronal signals in the presence of anisotropic and non-Gaussian variability. *Journal of Neuroscience Methods, 69*(2), 175–188.

Fukunaga, K., & Hostetler, L. (1975). The estimation of the gradient of a density function, with applications in pattern recognition. *IEEE Transactions on Information Theory, 21*(1), 32–40.

Goldberg, D., Victor, J., Gardner, E., & Gardner, D. (2009). Spike train analysis toolkit: Enabling wider application of information–theoretic techniques to neurophysiology. *Neuroinformatics, 7*, 165–178.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica, 37*(3), 424–438.

Gunay, C., Edgerton, J., Li, S., Sangrey, T., Prinz, A., & Jaeger, D. (2009). Database analysis of simulated and recorded electrophysiological datasets with PANDORA's toolbox. *Neuroinformatics, 7*, 93–111.

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.

Hazan, L., Zugaro, M., & Buzsaki, G. (2006). Klusters, NeuroScope, NDManager: A free software suite for neurophysiological data processing and visualization. *Journal of Neuroscience Methods 155*, 207–316.

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

Johnson, D. H., & Sinanovic, S. (2001). *Symmetrizing the Kullback-Leibler distance*. Available from: http://www.ece.rice.edu/~dhj/resistor.pdf. Accessed 15 April 2016.

Jun, J. J., Steinmetz, N. A., Siegle, J. H., Denman, D. J., Bauza, M., Barbarits, B., et al. (2017 November). Fully integrated silicon probes for high-density recording of neural activity. *Nature, 551*(7679), 232–236.

Kadir, S. N., Goodman, D. F., & Harris, K. D. (2014). High-dimensional cluster analysis with the masked EM algorithm. *Neural Computation, 26*(11), 2379–2394.

Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. *Annals of Mathematical Statistics, 22*(1), 79–86. Available from: http://www.jstor.org/stable/2236703.

Levinson, N. (1946). The wiener (root mean square) error criterion in filter design and prediction. *Journal of Mathematical Physics, 25*(1–4), 261–278.

Lidierth, M. (2009). sigTOOL: A MATLAB-based environment for sharing laboratory-developed software to analyze biological signals. *Journal of Neuroscience Methods, 178*, 188–196.

Liu, X. Q., Wu, X., & Liu, C. (2011). SPKtool: An open source toolbox for electrophysiological data processing. In *International Conference on Biomedical Engineering and Informatics, BMEI 2011* (Vol. 2, pp. 854–857).

Loader, C. (1999). *Local regression and likelihood*. New York: Springer.

Ma, S. k. (1981). Calculation of entropy from data of motion. *Journal of Statistical Physics, 26*(2), 221–240.

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, 1965/66*(1), 281–297.

Magri, C., Whittingstall, K., Singh, V., Logothetis, N., & Panzeri, S. (2009). A toolbox for the fast information analysis of multiple-site LFP, EEG and spike train recordings. *BMC Neuroscience, 10*(1), 81.

Mahmud, M., Bertoldo, A., Girardi, S., Maschietto, M., & Vassanelli, S. (2012a). SigMate: A Matlab-based automated tool for extracellular neuronal signal processing and analysis. *Journal of Neuroscience Methods, 207*(1), 97–112.

Mahmud, M., Bertoldo, A., Maschietto, M., Girardi, S., & Vassanelli, S. (2010). Automatic detection of layer activation order in information processing pathways of rat barrel cortex under mechanical whisker stimulation. In *Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2010* (6095–6098).

Mahmud, M., Cecchetto, C., Maschietto, M., Thewes, R., & Vassanelli, S. (2017). Towards high-resolution brain-chip interface and automated analysis of multichannel neuronal signals. In *2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)* (pp. 868–872).

Mahmud, M., Cecchetto, C., & Vassanelli, S. (2016). An automated method for characterization of evoked single-trial local field potentials recorded from rat barrel cortex under mechanical whisker stimulation. *Cognitive Computation, 8*(5), 935–945.

Mahmud, M., Kaiser, M. S., Hussain, A., & Vassanelli, S. (2018). Applications of deep learning and reinforcement learning to biological data. *IEEE Transactions on Neural Networks Learning Systems, 29*(6), 1–17. [Epub ahead of print].

Mahmud, M., Kaiser, M. S., Rahman, M. M., Rahman, M. A., Shabut, A., Al-Mamun, S., et al. (2018). A brain-inspired trust management model to assure security in a cloud based IoT framework for neuroscience applications. *Cognitive Computation, 10*(5), 864–873. [Epub ahead of print].

Mahmud, M., Pasqualotto, E., Bertoldo, A., Girardi, S., Maschietto, M., & Vassanelli, S. (2011). An automated method for detection of layer activation order in information processing pathway of rat barrel cortex under mechanical whisker stimulation. *Journal of Neuroscience Methods, 196*(1), 141–150.

Mahmud, M., Pulizzi, R., Vasilaki, E., & Giugliano, M. (2014). QSpike tools: A generic framework for parallel batch preprocessing of extracellular neuronal signals recorded by substrate microelectrode arrays. *Frontiers in Neuroinformatics, 8*, 26.

Mahmud, M., Rahman, M. M., Travalin, D., Raif, P., & Hussain, A. (2012b). Service oriented architecture based web application model for collaborative biomedical signal analysis. *Biomedizinische Technik. Biomedical Engineering, 57*(Suppl 1), 780–783.

Mahmud, M., Travalin, D., Bertoldo, A., Girardi, S., Maschietto, M., & Vassanelli, S. (2012c). An automated classification method for single sweep local field potentials recorded from rat barrel cortex under mechanical whisker stimulation. *Journal of Medical and Biological Engineering, 32*(6), 397–404.

Mahmud, M., & Vassanelli, S. (2016). Processing and analysis of multichannel extracellular neuronal signals: State-of-the-art and challenges. *Frontiers in Neuroscience, 10*(Jun), 248.

Meier, R., Egert, U., Aertsen, A., & Nawrot, M. P. (2008). FIND – A unified framework for neural data analysis. *Neural Networks, 21*(8):1085–1093.

Misra, N., Singh, H., & Demchuk, E. (2005). Estimation of the entropy of a multivariate normal distribution. *Journal of Multivariate Analysis, 92*(2), 324–342.

Montemurro, M. A., Senatore, R., & Panzeri, S. (2007). Tight data-robust bounds to mutual information combining shuffling and model selection techniques. *Neural Computation, 19*(11), 2913–2957.

Nawrot, P. M., Aertsen, A., & Rotter, S. (2003). Elimination of response latency variability in neuronal spike trains. *Biological Cybernetics, 88*(5), 321–334.

Optican, L. M., Gawne, T. J., Richmond, B. J., & Joseph, P. J. (1991). Unbiased measures of transmitted information and channel capacity from multivariate neuronal data. *Biological Cybernetics, 65*(5), 305–310.

Pachitariu, M., Steinmetz, N. A., Kadir, S. N., Carandini, M., & Harris, K. D. (2016). Fast and accurate spike sorting of high-channel count probes with KiloSort. In D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, & R. Garnett (Eds.), *Advances in Neural Information Processing Systems 29* (pp. 4448–4456). Red Hook: Curran Associates, Inc.

Paninski, L. (2003). Estimation of entropy and mutual information. *Neural Computation, 15*(6), 1191–1253.

Paninski, L., Pillow, J., & Lewi, J. (2007). Statistical models for neural encoding, decoding, and optimal stimulus design. *Progress in Brain Research, 165*, 493–507.

Panzeri, S., & Treves, A. (1996). Analytical estimates of limited sampling biases in different information measures. *Network: Computation in Neural Systems, 7*(1), 87–107.

Pastore, V. P., Godjoski, A., Martinoia, S., & Massobrio, P. (2018). SpiCoDyn: A toolbox for the analysis of neuronal network dynamics and connectivity from multi-site spike signal recordings. *Neuroinform, 16*(1), 15–30.

Pastore, V. P., Poli, D., Godjoski, A., Martinoia, S., & Massobrio, P. (2016). ToolConnect: A functional connectivity toolbox for in vitro networks. *Frontiers in Neuroinformatics, 10*, 13.

Pouzat, C., & Chaffiol, A. (2009). Automatic spike train analysis and report generation. An implementation with R, R2HTML and STAR. *Journal of Neuroscience Methods, 181*(1), 119–144.

Quian Quiroga, R., Kreuz, T., & Grassberger, P. (2002 October). Event synchronization: A simple and fast method to measure synchronicity and time delay patterns. *Physical Review E, 66*, 041904.

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.

Redish, A. D. (2014). *MClust Free-ware spike sorting*. Available from: http://redishlab.neuroscience.umn.edu/MClust/MClust.html. Accessed 07 January 2016.

Rey, H. G., Pedreira, C., & Quian Quiroga, R. (2015). Past, present and future of spike sorting techniques. *Brain Research Bulletin, 119*(Pt B), 106–117.

Rossant, C., Kadir, S. N., Goodman, D. F., Schulman, J., Hunter, M. L., Saleem, A. B., et al. (2016). Spike sorting for large, dense electrode arrays. *Nature Neuroscience, 19*(4), 634–641.

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.

Savitzky, A., & Golay, M. J. E. (1964 July). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry, 36*(8), 1627–1639.

Schröder, S., Cecchetto, C., Keil, S., Mahmud, M., Brose, E., Dogan, O., et al. (2015). CMOS-compatible purely capacitive interfaces for high-density in-vivo recording from neural tissue. In *Proceedings of 2015 IEEE Biomedical Circuits and Systems Conference (BioCAS)* (pp. 1–4).

Spacek, M., Blanche, T., & Swindale, N. (2009). Python for large-scale electrophysiology. *Front Neuroinform, 2*, 9.

Strong, S. P., Koberle, R., de Ruyter van Steveninck, R. R., & Bialek, W. (1998 Janaury). Entropy and information in neural spike trains. *Physical Review Letters, 80*, 197–200.

Swindale, N. V., & Spacek, M. A. (2014). Spike sorting for polytrodes: A divide and conquer approach. *Frontiers in Systems Neuroscience, 8*(6).

Takekawa, T., Isomura, Y., & Fukai, T. (2012). Spike sorting of heterogeneous neuron types by multimodality-weighted PCA and explicit robust variational Bayes. *Frontiers in Neuroinformatics, 6*, 5.

Tanskanen, J. M. A., Kapucu, F. E., & Hyttinen, J. A. K. (2018). A line of MEA signal analysis methods for human stem cell-derived and other dynamic neuronal cultures. *Frontiers in Cellular Neuroscience, 20*. https://doi.org/10.3389/conf.fncel.2018.38.00020.

Thewes, R., Bertotti, G., Dodel, N., Keil, S., Schroder, S., Boven, K. H., et al. (2016). Neural tissue and brain interfacing CMOS devices – An introduction to state-of-the-art, current and future challenges. In *Proceedings of the IEEE International Symposium on Circuits and Systems* (Vol. 2016-July, pp. 1826–1829).

Thomson, D. J. (1982). Spectrum estimation and harmonic analysis. *Proceedings of the IEEE, 70*(9), 1055–1096.

Thomson, D. J., & Chave, A. D. (1991). Jackknifed error estimates for spectra, coherences, and transfer functions. In *Advances in Spectrum Analysis and Array Processing* (pp. 58–113), Upper Saddle River, NJ: Prentice-Hall.

Treves, A., & Panzeri, S. (1995). The upward bias in measures of information derived from limited data samples. *Neural Computation, 7*(2), 399–407.

Vassanelli, S. (2011). Brain-chip interfaces: The present and the future. *Procedia Computer Science, 7*, 61–64.

Vassanelli, S. (2014). Multielectrode and multitransistor arrays for in vivo recording. In M. De Vittorio, L. Martiradonna, & J. Assad (Eds.), *Nanotechnology and Neuroscience: Nano-electronic, Photonic and Mechanical Neuronal Interfacing* (pp. 239–267). Springer: New York.

Vassanelli, S., Felderer, F., Mahmud, M., Maschietto, M., & Girardi, S. (2012a). CyberRat probes: High-resolution biohybrid devices for probing the brain. In *Living Machines 2012: Biomimetic and Biohybrid Systems*. LNAI (Vol. 7375, pp. 274–285).

Vassanelli, S., & Mahmud, M. (2016). Trends and challenges in neuroengineering: Toward "intelligent" neuroprostheses through brain-"brain inspired systems" communication. *Frontiers in Neuroscience, 10*(Sep), 438.

Vassanelli, S., Mahmud, M., Girardi, S., & Maschietto, M. (2012b). On the way to large-scale and high-resolution brain-chip interfacing. *Cognitive Computation, 4*(1), 71–81.

Victor, J. D. (2002 November). Binless strategies for estimation of information from neural data. *Physical Review E, 66*, 051903.

Victor, J. D., & Purpura, K. P. (1997). Metric-space analysis of spike trains: Theory, algorithms and application. *Network: Computation in Neural Systems, 8*(2), 127–164.

Wagenaar, D., DeMarse, T. B., & Potter, S. M. (2005). MeaBench: A toolset for multi-electrode data acquisition and on-line analysis. In *Proceedings of the 2nd International IEEE EMBS Conference on Neural Engineering* (pp. v–viii).

Wolpert, D. H., & Wolf, D. R. (1995 December). Estimating functions of probability distributions from a finite set of samples. *Physical Review E, 52*, 6841–6854.

Yger, P., Spampinato, G. L., Esposito, E., Lefebvre, B., Deny, S., Gardella, C., et al. (2018). A spike sorting toolbox for up to thousands of electrodes validated with ground truth recordings in vitro and in vivo. *eLIFE, 7*, e34518.