

Resolution of Spike Overlapping by Biogeography-Based Optimization

Original

Resolution of Spike Overlapping by Biogeography-Based Optimization / Chiarion, Giovanni; Mesin, Luca. - In: ELECTRONICS. - ISSN 2079-9292. - ELETTRONICO. - 10:12(2021), p. 1469. [10.3390/electronics10121469]

Availability:

This version is available at: 11583/2908072 since: 2021-08-30T17:39:27Z

Publisher:

MDPI

Published

DOI:10.3390/electronics10121469

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

Article

Resolution of Spike Overlapping by Biogeography-Based Optimization

Giovanni Chiarion , Luca Mesin * 

Mathematical Biology and Physiology, Department Electronics and Telecommunications, Politecnico di Torino, 10129 Turin, Italy; giovanni.chiarion@polito.it

* Correspondence: luca.mesin@polito.it; Tel.: +39-0110904085

Abstract: There are many cases in which the separation of different sources from single channel recordings is important, for example, in fluorescence spectral overlap compensation, electrical impedance signaling, intramuscular electromyogram decomposition or in the case of spike sorting of neuron potentials from microelectrode arrays (MEA). Focusing on the latter, the problem can be faced by identifying spikes emerging from the background and clustering into different groups, indicating the activity of different neurons. Problems are found when more spikes are superimposed in overlapped waveforms. We discuss the application of Biogeography-Based Optimization (BBO) to resolve this specific problem. Our algorithm is compared with three spike-sorting methods (SpyKING Circus, Common Basis Pursuit and Klusta), showing statistically better performance (in terms of F1 score, True Positive Rate—TPR and Positive Predictive Value—PPV) in resolving overlaps in realistic, simulated data. Specifically, BBO showed median F1, TPR and PPV of 100%, 100% and about 75%, respectively, considering a simulated noise with the same spectral density as the experimental one and a similar power with highly statistically significant improvements of at least two performance indexes over each of the other three tested algorithms.



Citation: Chiarion, G.; Mesin, L. Resolution of Spike Overlapping by Biogeography-Based Optimization.

Electronics **2021**, *10*, 1469.

<https://doi.org/10.3390/electronics10121469>

electronics10121469

Academic Editor: Hugo Proença

Received: 10 May 2021

Accepted: 16 June 2021

Published: 18 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: microelectrode arrays (MEA); spike sorting; Biogeography-Based Optimization (BBO)

1. Introduction

Technological progresses have made it possible to record the activity of living neuron cells in vitro [1] and in vivo [2]. For example, microelectrode (or multi-electrode) arrays (MEAs) can measure the action potential of neurons, using a certain number of electrodes placed in vitro [3]. Depending on the position of the neurons with respect to the electrode, a different waveform is recorded, allowing to distinguish the activity of each sensed neuron (i.e., spike sorting [4]). In this scenario, a serious complication is that the signal generated over time by each channel is often very noisy, so only the activity of nearby neurons can be discriminated from the potential recorded by an electrode. Experimentally, a few neurons can be correctly detected by modern spike-sorting algorithms on each channel [5,6]; the more neurons are sorted, the higher the probability for generating false positives and false negatives [7].

The correct identification of the activity of each neuron perceived by an electrode is very important. For example, a certain number of false positive spikes assigned to a neuron could significantly change some physiologically important parameters, such as the firing rate (FR) and the variability of the inter-spike intervals (ISI).

One of the most annoying yet unsolved problems in spike sorting is the discrimination of overlaps that occur when different neurons in the detection volume of a channel fire at close time instants [8]. A single neuron can hardly generate overlapping fires alone since it remains refractory for a few ms after firing [9,10]; moreover, the maximal observed FRs are usually lower than 50 Hz, whereas the generated spikes have a duration of a few ms [11]. Instead, it is pretty common to have two or more different neurons that fire in

a narrow time window, generating a resulting potential that is a mixture of more single neuron spike waves.

Most modern spike-sorting algorithms do not take into consideration the overlapping spikes problem by ignoring and discarding the recognized overlapping waves [12]. In fact, some separate computations performed after the main spike-sorting algorithm are required to distinguish the activity of two or more sources firing together. A few algorithms do face the problem, usually by template matching (e.g., SpyKING Circus [13], KiloSort [14], MountainSort [15] and KlustaKwik [16]), or other techniques, such as sparse optimization (Continuous Basis Pursuit [17]), but performance is often not good [18,19].

We propose an innovative approach to solve this problem by an integer optimization method. Specifically, an overlapping spike wave is optimally reconstructed by the convolution of prototypes of single neuron spikes (estimated by clustering and averaging not overlapped waves) with solution vectors obtained by Biogeography-Based Optimization (BBO) [20]. This evolutionary-based algorithm was chosen because of the big dimension of the solution space. BBO is appreciated for the capability of finding good solutions in reasonable time and when the cardinality of the solution space makes very intensive use of classical integer optimization approaches [20]. Compared to the popular genetic algorithm (GA) [21], BBO is a more recent approach that has outperformed GA in many different scientific and engineering fields [22,23].

Our method was tested on different simulated data with different levels of noise. The performance was compared to those obtained with three other spike-sorting algorithms, which face the problem of temporal overlapping: SpyKING Circus (SC, [13]), KlustaKwik (KK, [16]) and Continuous Basis Pursuit (CBP, [17]). All of these algorithms are open source and available online in Github repositories [24–26].

2. Methods

In this section, the simulated ground truth signal is first described. Then, a simple spike identification algorithm, used to find the spikes and estimate prototypes, is described. The proposed solution of the overlapping spike problem by BBO is then introduced. Finally, the performance indexes are described.

Our algorithm was implemented using Matlab 2020b and run on an 8-core AMD Ryzen 2700x CPU of a desktop computer with a Windows 10 operating system. CBP was also run on the same system, whereas the tests with SC and KK were performed using Python 3.8 with a Conda virtual environment [27].

2.1. Ground-Truth Signal Generation

To create a reasonable ground-truth signal, we defined three essential parameters:

1. The number of neurons that independently fire, N ;
2. The FR of each neuron, f_i (chosen randomly in the range 10–20 Hz);
3. The coefficient of variation (CoV) of the ISI of each neuron, ISI_i (Gaussian random distributions of the ISI were assumed; CoV of ISI was 15% for each firing pattern).

Then, we generated N vectors of zeros with length $f_s T$, where f_s is the sampling frequency (10 kHz) and T the signal duration in s (90 s). Each of the N zeros vectors was then populated with ones, randomly chosen to follow the parameters defined above. This way, we got the firing pattern of each neuron. To obtain the firing trains of each neuron, such vectors were then convoluted with appropriate kernels, each representing the prototype of the wave measured by the electrode when the corresponding neuron fires.

The N kernels were the templates extracted by the SpyKING Circus by processing a real dataset with default parameters [28]. Using the Phy Python GUI [29], we searched for a channel with more than two clusters detected. Channel 9, with three different neurons detected, was chosen and we imported in Matlab the corresponding three templates (thus, $N = 3$). The templates were normalized, dividing by the maximum value M_k of the highest peak. This way, each template had an amplitude less than or equal to 1.

The ground-truth signal was generated as the summation of N firing trains and random noise.

To better simulate the noise present in a real MEA acquisition, from the same dataset, we extracted the residual noise and then characterized it. Specifically, given the firing matrix obtained by SpyKING Circus, a window of 5 ms centered on each spike was removed. The remaining sample series was considered as the residual noise of the original signal. A white Gaussian noise was then filtered, using the coefficients of an autoregressive Burg's model (of order 20) of the residual noise. Thus, a colored Gaussian process with approximately the same power spectral density (PSD) as the experimental residual noise was obtained.

Different noise levels were added to the simulated signal. An estimation of the noise level present in the dataset was obtained calculating the standard deviation of the residual signal normalized by M_k , obtaining 0.09. Then, we selected 0.05 and 0.1 as different levels of the noise to be applied over our simulated signal, in order to consider a lower noise level, and another a bit higher than the experimental one.

2.2. Spike Sorting Algorithm

We developed a simple spike-sorting algorithm to identify the neuron firings, build the prototype waveforms of their action potentials and find the overlaps. The resolution of those overlaps will allow to measure the possible improvement provided by the introduction of the BBO in the pipeline.

At first, the input signal was band-pass filtered in the range 300–3000 Hz [7] (5th order low-pass and high-pass Chebyshev filters of type I, forward and reverse application to get zero-phase).

An appropriate threshold proportional to the noise was then applied to the signal. Specifically, the mean absolute deviation was used to estimate the noise (since it is more robust than standard deviation). A proportional factor of $k = 5$ was chosen [7].

A 5 ms time window centered on the positive identified peaks was chosen to select each waveform.

Principal component analysis (PCA) was applied to the dataset of waveforms, extracting the first components that allowed to keep 60% variance included in the data (a few tests showed that increasing the represented energy did not improve the performance of our method for spike identification). The resulting transformed dataset was then passed to a k-means algorithm. In order to avoid bad clustering, due to an unfortunate initial seed, the k-means was run 150 times. The best clusterization in terms of inter-cluster distance from the centroids was then selected. At last, waves belonging to the same clusters were averaged to obtain the N prototypes. Those clusters included also a few overlapping waves. Overlaps were identified as those waves with low cross-correlation with the cluster prototype (i.e., the cross-correlations between the prototype and all spikes of the cluster were computed; mean and standard deviations of these cross-correlations were computed; the overlaps were the spikes showing cross-correlation with the prototype lower than the mean minus 2.5 times the standard deviation).

2.3. Biogeography-Based Optimization

The Biogeography-Based Optimization algorithm (BBO) [20] was used here to separate overlapping spike waves. The BBO is an evolutionary optimization algorithm that tries to minimize an *objective function*, measuring the high suitability index, HSI, that depends on a candidate solution found by the algorithm. HSI is an index of the goodness of the solution, similar to the fitness score used in genetic algorithms [21].

The solutions were searched stochastically and iteratively through hundreds or thousands of iterations, modeling the rationale of the speciation, migration and extinction of species. Considering an ecosystem (the entire set of candidate solutions), species living there are likely to migrate from "islands" where there are poor living conditions to others with better conditions. Each problem solution is considered to be a habitat of our species

and each independent variable of a candidate solution characterizes the habitability of species in that habitat (called the suitability index variable, SIV).

Habitats with great values of HSI have solutions with a high probability of emigration (high emigration rate μ) and a low probability of immigration (low immigration rate λ); the opposite holds for low HSI habitats.

The parameter μ represents how likely a solution will share its features with other solutions; λ , instead, represents how likely a solution will accept features from other solutions. It is low when the number of species in the habitat (and, therefore, HIS) is high. With this rationale, good solutions with high HSI have low probability of changing their own features but a high probability of sharing their features with other solutions.

2.3.1. Solution Encoding

The first essential step is encoding our solution. Since we want to process only a little portion of the signal where an overlap is present, the solution is a vector of dimension $overlap_time \times f_s$, where $overlap_time$ is the time duration of an overlap in s. As $overlap_time$, we chose 5 ms and f_s was 10 kHz, so the solution was a 50×1 vector.

Each bit of the solution represents a time instant where a spike can occur or not; if it occurs, it may come from neurons 1, 2 or 3. Hence, as the solution space, we have all 50×1 vectors, where each bit can be either 0 (if no spike occurred), or one of the following numbers (when one neuron fires): 1, 2 or 3. Notice that a perfect superposition of spikes is not allowed (but the sampling frequency is high enough to make this occurrence unlikely and to reduce its effect in the case that it occurs).

The solution space being spanned by the BBO is huge, as it contains about 4^{50} solutions. In this case, a *brute-force* approach to search the optimal solution for a single overlap would be very expensive. By adding the constraint that each neuron can only fire once within the window, the number of the search space can be drastically reduced. However, the number of feasible solutions is still considerable, i.e., K^N , where N is the number of neurons (3 in our case) and $K = 50$ is the number of samples in a 5 ms window sampled at 10 kHz. The number of feasible solutions, in this case, is more than 10^5 for each overlapping spike.

2.3.2. Objective Function

The choice of a good objective function is crucial. The candidate solution evaluated by the BBO algorithm was used to reconstruct the original overlap, using the templates estimated by the spike-sorting algorithm. As we want that the reconstructed signal to be similar to the original, the following squared error was used as the metric of similarity:

$$Cost(k) = \sum_{i=1}^{50} [s(i) - r(i)]^2 \quad (1)$$

where i is the i th bit of the k th candidate solution, s is the input signal (i.e., the overlap) and r is the reconstructed version of it, using the estimated templates.

2.3.3. Initialization and Parameters

As for GAs, also for BBO, there were several choices to be made and problem-dependent parameters to be set. At first, an initial population of 90 candidate solutions was randomly selected. Then a reasonable feasibility check was made; as already mentioned, since the overlap time is small, a candidate initial solution was accepted only if the number of spikes for each neuron was less than two (i.e., the inter-spike period of a neuron was assumed to be larger than the duration of our window).

As the stop criterion, a maximum number of iterations equal to 50 was chosen and 8 independent trials were made for each overlap signal. The crossover probability was set to 1, and the mutation probability to 0.01.

Among the 8 independent trials, the minimum cost solution was chosen to build the corresponding firing matrix.

A low number of iterations and a small population size were chosen mainly for two reasons.

- It was experimentally seen that most of the relevant changes in the solutions were in the very first generations of each trial. To help the algorithm explore a larger solution space, we preferred to perform a higher number of independent trials.
- We wanted the BBO to find an optimal solution in a reasonable time frame (about 5 s in our PC) for each overlap.

2.4. Performance Tests

After setting up the sampling rate (10 kHz), width of the templates (5 ms) and peak polarity (positive), the default parameters were chosen for SC, KK and CBP. For SC, the parameter indicating the maximum number of clusters was set to 3, according to the generated input signal.

Some performance indexes were calculated to compare the estimated firing matrix with the ground truth. The firing matrix columns of all algorithms were rearranged in order to refer to the same ground-truth template. To match the firing patterns to the correct neuron, the prototype with maximum cross-correlation with the ground truth kernel was selected. If a template had high correlation with more prototypes, the closest relative amplitude was used to make the selection. The correctness of the choice was finally checked by comparison of the estimated FRs with the ground truth.

A match in the identification of a spike was identified if the estimation had a lag lower than 0.5 ms from the ground truth. The following performance indexes were used: F1 Index, True Positive Rate (TPR) and Positive Predictive Value (PPV). Possible statistically significant differences were assessed by the Wilcoxon signed rank test (paired tests could be applied as the same signals, thus the same overlaps were processed by the algorithms).

A sensitivity analysis of the sampling frequency was performed by repeating our tests on the same data down-sampled at either 8 kHz or 6 kHz.

3. Results

Figure 1 shows the flow chart of our algorithm, together with the representation of a few steps. The signal was filtered and then thresholded. Peaks over the threshold were identified and then split into clusters. Those which were very different from others were selected as overlaps, which were then resolved by BBO.

Figure 2 explains the simulator. Experimental data from a single channel was processed by SpyKING Circus. The prototypes of waveforms indicating the firings of different neurons were identified. Their contributions were then subtracted from the original signal to obtain the residual noise. The prototypes were used to simulate the noise-free signal, to which colored Gaussian noise with same PSD as the experimental one was added.

Figure 3 shows the test on overlap resolution. The same simulated signal was processed by SpyKING Circus (SC), KlustaKwik (KK), Continuous Basis Pursuit (CBP) and our method (BBO). The simulated signal included 100 overlaps for which we focused on testing the performance compared to the ground truth of the different methods in resolving them. Some examples of overlaps are shown in (A). The performance is shown in (B), considering the four methods applied to the same data corrupted by two levels of noise. Three relevant performance indexes are indicated: F1 Index, True positive rate (TPR) and positive predictive value (PPV). The proposed method showed statistical improvements with respect to the other three methods: at least two performance indexes were statistically better for BBO in comparison to each of the other methods; no performance index was statistically better for another method than for BBO.

Figure 4 displays the sensitivity of BBO on the sampling frequency. The same signals used before were down-sampled, considering 8 kHz and 6 kHz (in addition to the original signals, sampled at 10 kHz) with two different levels of noise ($\sigma = 0.05$ or $\sigma = 0.1$, as before). A statistically significant degradation of performance was obtained. However, it was small when using a sampling frequency of 8 kHz; on the other hand, a relevant drop in

the performance was observed if the data were sampled at 6 kHz (some perturbation of the prototypes could be observed with such a low sampling frequency, as they are very spiky). Despite this drawback, the reduction of the sampling frequency has the advantage of speeding up the computations, with an average reduction in the computational times with respect to the original data of 18% and 34.5% for 8 and 6 kHz, respectively.

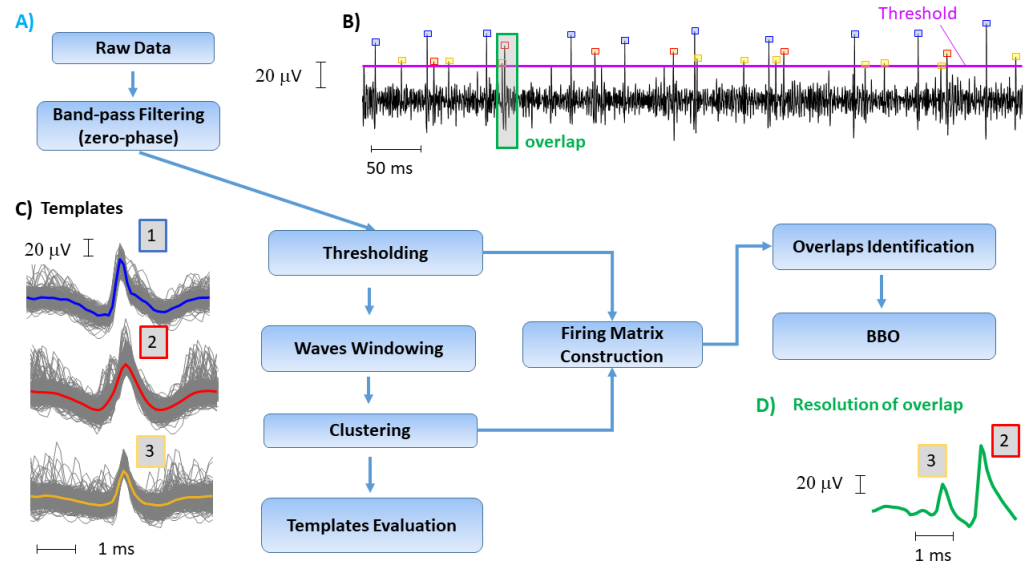


Figure 1. Flow diagram of our method, with example application of some steps. (A) Flow diagram. (B) Example of signal with indication of the selected threshold, the peaks identified (with different color of the markers indicating membership) and an overlap. (C) Clustering of spikes, alignment of windows including them and averaging to get an estimation of prototypes. (D) Resolution of the overlap indicated in (B), with identification of the prototypes superimposed in it.

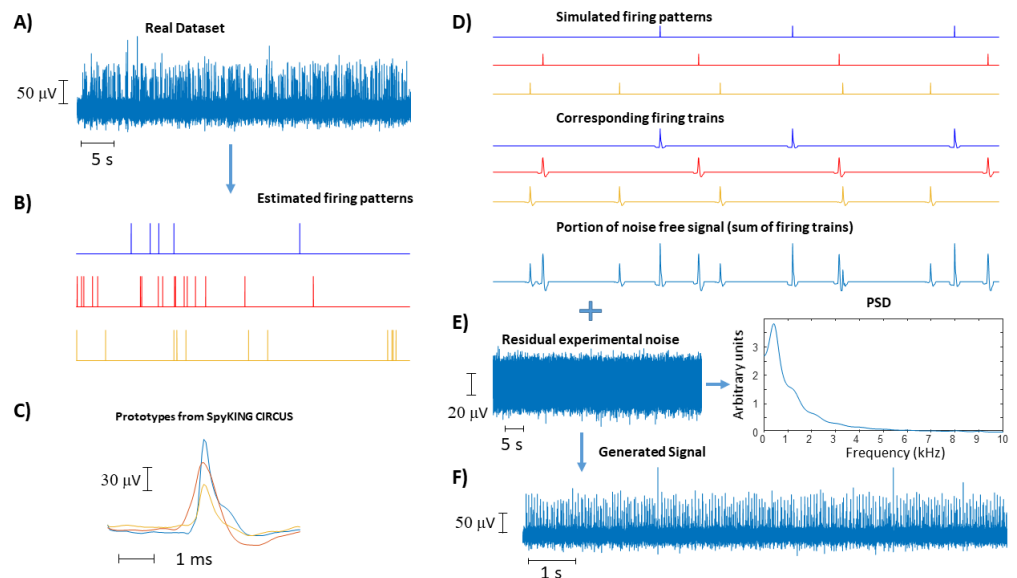


Figure 2. Simulation of data. (A) Portion of experimental data considered to extract prototypes of neuron spikes. (B) Example of spikes identified by SpyKING Circus, separated after sorting. (C) Prototypes identified. (D) Simulation of a portion of data by generating the firing patterns corresponding to each prototype, the firing trains (by convolution) and the overall signal (sum of firing trains). (E) Model of noise: The experimental noise is obtained by subtracting the identified spikes from the experimental data and its power spectral density (PSD) is estimated (by Burg method of order 20). (F) Final simulation of noisy data by summing the noise-free signal (D) and colored Gaussian noise with the same PSD as the experimental one (E).

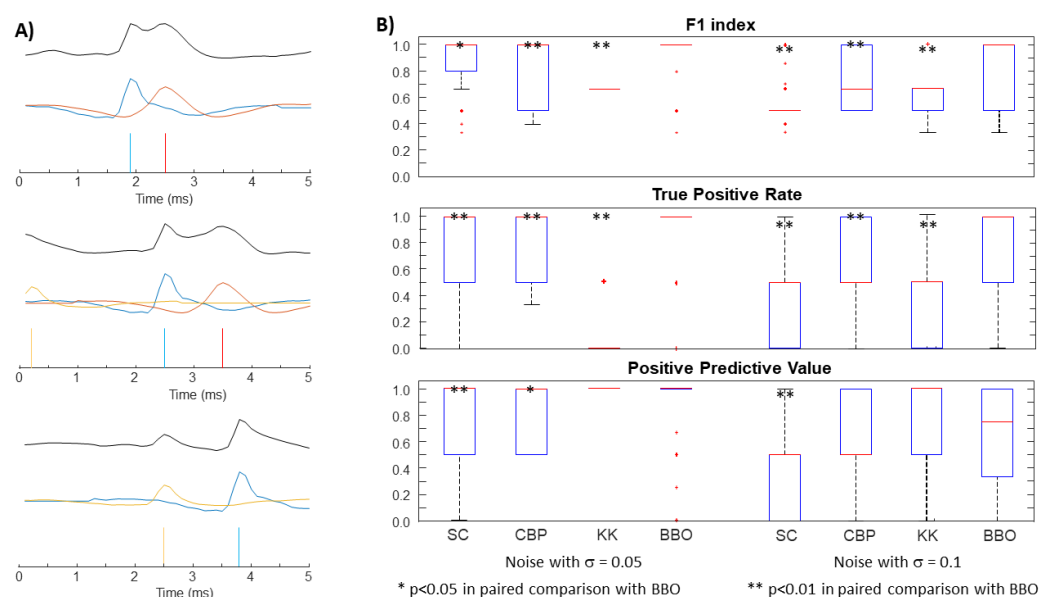


Figure 3. Overall results on the resolution of 100 overlaps randomly occurring in simulated data. (A) Few examples of overlap, with indication of the prototypes superimposed. (B) F1 Index, True Positive Rate (TPR) and Positive Predictive Value (PPV) are shown in box and whiskers plots. The performances of SpyKING Circus (SC), Continuous Basis Pursuit (CBP), KlustaKwik (KK) and Biogeography-Based Optimization (BBO) are shown for two levels of noise. Paired comparisons between one of the three reference methods (SC, CBP and KK) and BBO, providing statistically significant differences, are indicated.

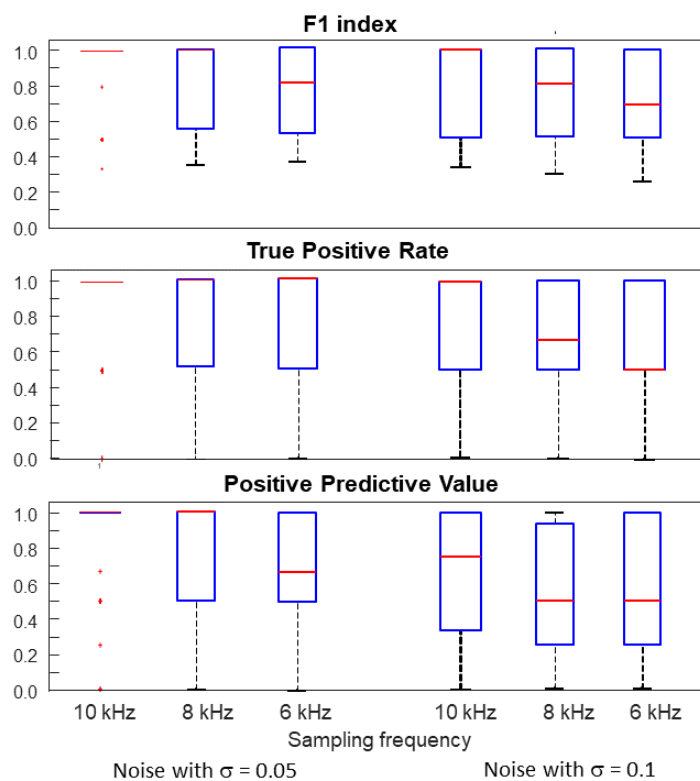


Figure 4. Sensitivity test on the sampling frequency. Same data as in Figure 3 are considered. F1 Index, True Positive Rate (TPR) and Positive Predictive Value (PPV) are shown in box and whiskers plots. The performance of BBO is shown for two levels of noise and three different sampling frequencies.

4. Discussion

There are many applications requiring the separation of the activity of different sources from single channel recordings, for example, in the identification of single motor unit (MU) activity from an intramuscular electromyogram (EMG) [30] or in spike sorting of single neuron potentials from microelectrode arrays (MEA) [7]. One of the most important difficulties is separating two or more source spikes occurring at close time instants, thus generating a waveform that is the superposition of different contributions. Due to phase cancellation, superimposed noise and lack of precise information on the prototype waveforms, which could be superimposed, the problem is under-determined. It is then converted into an optimization problem, e.g., minimizing the mean squared error in reconstructing the overlap by superimposing delayed versions of prototypes. However, the global solution may even be not correct. For example, an artifact or the superposition of other waveforms not included in the set of prototypes cannot be recognized as such, but they can be approximated using the available waveforms; this would indicate the activity of a neuron that has not actually fired.

Most of the spike sorting algorithms do not resolve overlaps but simply discard them [12]; others, in large part, use algorithms based on template matching [13–16]. Here, we propose an innovative method for resolving spike overlapping, based on the solution of an integer optimization problem. The search for an optimal solution of this problem can be approached with different methods (but excluding the exhaustive search, due to the prohibitive computational cost). We selected an evolutionary algorithm, i.e., the Biogeography-Based Optimization (BBO) [20]. It has found many applications in different research fields and it is here applied for the first time on a problem of overlap resolution in data from spiking neurons. BBO is appreciated for its good performance and relatively low computational cost [20,22,23]. Some constraints are imposed: a waveform should be included with the correct amplitude (i.e., a neuron either fires or not, with no possibility of including an amplitude scaled prototype, which could be accommodated by a template matching approach, instead); repetitive firings of the same neuron with a non-physiological lag are not allowed.

The performance of our algorithm was tested on MEA data simulated with an approach driven by an experimental signal. Specifically, a single channel recording was considered; our simulations were obtained by summation of real spikes (extracted by a sorting algorithm) and random noise with same PSD as in the experiment. Our results indicated significant improvements in resolving overlaps with respect to three popular spike-sorting algorithms used as reference, i.e., SpyKING Circus, KlustaKwik and Continuous Basis Pursuit.

Some specific choices were made in the generation of the simulated signal and in the test. For example, the length of the overlap window (5 ms) and the sampling frequency (10 kHz) are two parameters that were fixed. Our choices have the following rationale. The window duration is related to the maximum expected duration of an overlap: as we considered three neurons, each producing a spike of duration of about 1.5–2.0 ms, 5 ms allows to get non-overlapping spikes in the case of maximal separation among them. Thus, enlarging the window is not very useful (and increases the computational cost) and shortening it exposes the risk of not including the entire overlap waveform. Thus, some variations of the duration of the window can be allowed, but in a small range (e.g., about 4–6 ms). About the sampling frequency, it is important to fix it close to the Nyquist limit so as to avoid increasing the computational cost by oversampling. A test on the effect of decreasing the sampling frequency is provided in Figure 4: a decrement in the performance was observed, mostly when the sampling frequency was so small so as to introduce some perturbations in the spike prototypes.

As a limitation of our study, we assumed that the signals were stationary: thus, possible variations of the prototype waveforms were not allowed. Moreover, we considered a basic method to identify spikes (focusing only on resolving overlaps): thus, prototypes may not be optimally estimated. However, our preliminary tests documented that using

BBO for overlap resolution improves performances over other reference algorithms. Thus, the integration of the overlap resolution by BBO into an efficient algorithm for spike identification can be a promising direction.

5. Conclusions and Further Work

We documented on simulations some improvements of the spike overlap resolution with respect to some popular algorithms by solving approximately an integer optimization problem by Biogeography-Based Optimization (BBO). Our promising results, if confirmed on a larger dataset, suggest the integration of the overlap resolution by BBO into efficient spike-identification algorithms. The promising results we achieved suggest the possibility to apply this method also to other applications. For example, fluorescence spectral overlap compensation [31], electrical impedance signaling [32] and the intramuscular EMG decomposition problem [30] are suggested for future work.

Author Contributions: Conceptualization, G.C., L.M.; methodology, G.C., L.M.; software, G.C.; validation, G.C., L.M.; data preparation, G.C.; investigation, G.C., L.M.; writing—original draft preparation, G.C., L.M.; writing—review and editing, L.M.; visualization, G.C., L.M.; supervision, L.M. Both authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The experimental data presented in this study are openly available from <https://zenodo.org/record/1205233#YMqqaUwRWWhd> (accessed on 1 June 2021).

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

PSD	Power Spectral Density
BBO	Biogeography-Based Optimization
FR	Firing Rate
ISI	Inter-Spike Interval
SC	SpyKING CIRCUS
KK	KlustaKwik
CBP	Continuous Basis Pursuit

References

1. Darlington, C.L.; Smith, P.F.; Hubbard, J.I. Neuronal activity in the guinea pig medial vestibular nucleus in vitro following chronic unilateral labyrinthectomy. *Neurosci. Lett.* **1989**, *105*, 143–148.
2. Broussard, G.J.; Liang, Y.; Fridman, M.; Unger, E.K.; Meng, G.; Xiao, X.; Ji, N.; Petreanu, L.; Tian, L. In vivo measurement of afferent activity with axon-specific calcium imaging. *Nat. Neurosci.* **2018**, *21*, 1272–1280.
3. Welkenhuysen, M.; Hoffman, L.; Luo, Z.; De Proft, A.; Van den Haute, C.; Baekelandt, V.; Debyser, Z.; Gielen, G.; Puers, R.; Braeken, D. An integrated multi-electrode-optrode array for in vitro optogenetics. *Sci. Rep.* **2016**, *2*, 20353.
4. Quiroga, R.Q.; Nadasdy, Z.; Ben-Shaul, Y. Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. *Neural Comput.* **2004**, *16*, 1661–1687.
5. Shoham, S.; O'Connor, D.H.; Segev, R. How silent is the brain: Is there a “dark matter” problem in neuroscience? *J. Comp. Physiol.* **2006**, *192*, 777–784.
6. Buzsáki, G. Large-scale recording of neuronal ensembles. *Nat. Neurosci.* **2004**, *7*, 446–451.
7. Rey, H.G.; Pedreira, C.; Quiroga, R. Past, present and future of spike sorting techniques. *Brain Res. Bull.* **2015**, *119*, 106–117.
8. Mokri, Y.; Salazar, R.F.; Goodell, B.; Baker, J.; Gray, C.M.; Yen, S.C. Sorting Overlapping Spike Waveforms from Electrode and Tetrode Recordings. *Front. Neuroinform.* **2017**, *11*, 53.
9. Sukiban, J.; Voges, N.; Dembek, T.A.; Pauli, R.; Visser-Vandewalle, V.; Denker, M.; Weber, I.; Timmermann, L.; Grun, S. Evaluation of Spike Sorting Algorithms: Application to Human Subthalamic Nucleus Recordings and Simulations. *Neuroscience* **2019**, *414*, 168–185.
10. Shimamoto, S.A.; Ryapolova-Webb, E.S.; Ostrem, J.L.; Galifianakis, N.B.; Miller, K.J.; Starr, P.A. Subthalamic nucleus neurons are synchronized to primary motor cortex local field potentials in Parkinson's disease. *J. Neurosci.* **2013**, *33*, 7220–7233.
11. Negri, J.; Menon, V.; Young-Pearse, T.L. Assessment of Spontaneous Neuronal Activity In Vitro Using Multi-Well Multi-Electrode Arrays: Implications for Assay Development. *eNeuro* **2020**, *7*, 80–119.

12. Pillow, J.W.; Shlens, J.; Chichilnisky, E.J.; Simoncelli, E.P. A model-based spike sorting algorithm for removing correlation artifacts in multi-neuron recordings. *PLoS ONE* **2013**, *8*, e62123.
13. Yger, P.; Spampinato, G.L.; Esposito, E.; Lefebvre, B.; Deny, S.; Gardella, C.; Stimberg, M.; Jetter, F.; Zeck, G.; Picaud, S.; et al. A spike sorting toolbox for up to thousands of electrodes validated with ground truth recordings in vitro and in vivo. *Elife* **2018**, *20*, e34518.
14. Pachitariu, M.; Steinmetz, N.; Kadir, S.; Carandini, M.; Harris, K.D. Kilosort: Realtime spike-sorting for extracellular electrophysiology with hundreds of channels. *bioRxiv* **2016**, doi:10.1101/061481.
15. Chung, J.E.; Magland, J.F.; Barnett, A.H.; Tolosa, V.M.; Tooker, A.C.; Lee, K.Y.; Shah, K.G.; Felix, S.H.; Frank, L.M.; Greengard, L.F. A fully automated approach to spike sorting. *Neuron* **2017**, *95*, 1381–1394.
16. Rossant, C.; Kadir, S.; Goodman, D.; Schulman, J.; Hunter, M.L.; Saleem, A.B.; Harris, K.D. Spike sorting for large, dense electrode arrays. *Nat. Neurosci.* **2016**, *19*, 634–641.
17. Ekanadham, C.; Tranchina, D.; Simoncelli, E.P. A unified framework and method for automatic neural spike identification. *J. Neurosci. Methods* **2014**, *222*, 47–55.
18. Bar-Gad, I.; Ritov, Y.; Vaadia, E.; Bergman, H. Failure in identification of overlapping spikes from multiple neuron activity causes artificial correlations. *J. Neurosci. Methods* **2001**, *107*, 1–13.
19. Caro-Martín, C.R.; Delgado-García, J.M.; Gruart, A.; Sánchez-Campusano, R. Spike sorting based on shape, phase, and distribution features, and K-TOPS clustering with validity and error indices. *Sci. Rep.* **2018**, *8*, 17796.
20. Simon, D. Biogeography-Based Optimization. *IEEE Trans. Evol. Comput.* **2008**, *12*, 702–713.
21. Holland, J.H. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*; MIT Press: Cambridge, UK, 1992.
22. Singh, U.; Kumar, H.; Kamal, T.S. Design of Yagi-Uda Antenna Using Biogeography Based Optimization. *IEEE Trans. Antennas Propag.* **2010**, *58*, 3375–3379.
23. Zheng, Y.J.; Ling, H.F.; Shi, H.H.; Chen, H.S.; Chen, S.Y. Emergency Railway Wagon Scheduling by Hybrid Biogeography-based Optimization. *Comput. Oper. Res.* **2014**, *43*, 1–8.
24. SpyKING CIRCUS. Available online: <https://github.com/spyking-circus/spyking-circus> (accessed on 8 May 2021).
25. KlustaKwik. Available online: <https://github.com/klusta-team/klustakwik> (accessed on 8 May 2021).
26. Continuous Basis Pursuit. Available online: <https://github.com/chinasaur/CBPSpikesortDemo> (accessed on 8 May 2021).
27. SpyKING CIRCUS. Available online: <https://spyking-circus.readthedocs.io/en/latest/introduction/install.html> (accessed on 8 May 2021).
28. Spampinato, G.L.B.; Esposito, E.; Yger, P.; Duebel, J.; Picaud, S.; Marre, O. Ground Truth Recordings for Validation of Spike Sorting Algorithms (Version 1.0). Available online: <https://zenodo.org/record/1205233#.YMqqaUwRWhd> (accessed on 1 June 2021).
29. Phy GitHub. Available online: <https://github.com/cortex-lab/phy> (accessed on 8 May 2021).
30. Stashuk, D. EMG signal decomposition: How can it be accomplished and used? *J. Electromyogr. Kinesiol.* **2001**, *11*, 151–173.
31. Bagwell, C.B.; Adams, E.G. Fluorescence spectral overlap compensation for any number of flow cytometry parameters. *Ann. N. Y. Acad. Sci.* **1993**, *677*, 167–184.
32. Liu, J.; Qiang, Y.; Alvarez, O.; Du, E. Electrical impedance microflow cytometry with oxygen control for detection of sickle cells. *Sens. Actuators B Chem.* **2018**, *255*, 2392–2398.