Article

Duplicate Detection of Spike Events: A Relevant Problem in Human Single-Unit Recordings

Gert Dehnen [1,†], Marcel S. Kehl [1,†], Alana Darcher [1], Tamara T. Müller [2,3], Jakob H. Macke [2,4], Valeri Borger [5], Rainer Surges [1] and Florian Mormann [1,*]

1 Department of Epileptology, University of Bonn Medical Center, Venusberg-Campus 1, 53127 Bonn, Germany; gert.dehnen@ukbonn.de (G.D.); marcel.kehl@ukbonn.de (M.S.K.); alana.darcher@ukbonn.de (A.D.); rainer.surges@ukbonn.de (R.S.)

2 Computational Neuroengineering, Department of Electrical and Computerengineering, TU Munich, 80333 Munich, Germany; tamara.mueller@tum.de (T.T.M.); jakob.macke@uni-tuebingen.de (J.H.M.)

3 Institute for Artificial Intelligence and Informatics in Medicine, TU Munich, 80333 Munich, Germany

4 Machine Learning in Science, University of Tübingen, 72076 Tübingen, Germany

5 Department of Neurosurgery, University of Bonn Medical Center, Venusberg-Campus 1, 53127 Bonn, Germany; valeri.borger@ukbonn.de

* Correspondence: florian.mormann@ukbonn.de

† Both authors contributed equally to this work.

Abstract: Single-unit recordings in the brain of behaving human subjects provide a unique opportunity to advance our understanding of neural mechanisms of cognition. These recordings are exclusively performed in medical centers during diagnostic or therapeutic procedures. The presence of medical instruments along with other aspects of the hospital environment limit the control of electrical noise compared to animal laboratory environments. Here, we highlight the problem of an increased occurrence of simultaneous spike events on different recording channels in human single-unit recordings. Most of these simultaneous events were detected in clusters previously labeled as artifacts and showed similar waveforms. These events may result from common external noise sources or from different micro-electrodes recording activity from the same neuron. To address the problem of duplicate recorded events, we introduce an open-source algorithm to identify these artificial spike events based on their synchronicity and waveform similarity. Applying our method to a comprehensive dataset of human single-unit recordings, we demonstrate that our algorithm can substantially increase the data quality of these recordings. Given our findings, we argue that future studies of single-unit activity recorded under noisy conditions should employ algorithms of this kind to improve data quality.

1. Introduction

The opportunity to record from single neurons in the brain of behaving human subjects increasingly contributes to the advances of cognitive and systems neuroscience. These recordings allow researchers to investigate complex brain functions, such as perception [1–9], memory [10–14], emotion [15,16], or decision making [17,18].

In these studies, humans are implanted with intracranial electrodes solely for medical purposes, such as the identification of the seizure onset zone in patients with pharmacologically intractable epilepsy [19], the treatment of movement disorders [20,21], or the management of treatment-resistant depression [22]. In some medical centers, it is currently possible to record the activity of single neurons from these patients while performing cognitive experiments. As the opportunity to perform such experiments is exceedingly rare, it is imperative that researchers optimize the quality of the recorded data. In a clinical setting, there are many different external sources of noise, such as medical instruments close to the patient, and only limited possibilities to control the setting [23]. The signal

quality can be increased during the recordings by eliminating local noise sources, as well as after the recording by detecting artifacts using advanced spike-sorting algorithms. In animal studies, the development of polytrodes has increased the signal-to-noise ratio and thereby the reliability of single-unit recordings [24–26]. Since polytrodes have not yet been accredited for use in humans, most medical centers currently use microwire bundles, which have a considerably lower signal-to-noise ratio. To take full advantage of these datasets, it is essential that researchers identify as many artifacts and sources of noise as possible in microwire recordings.

Despite the limitations mentioned above, various advances in electrophysiological technology have led to a rapid increase in the total number of neurons recorded in a given experiment. To deal with this increasing amount of data, automated spike-sorting algorithms have become crucial to efficiently extract and cluster the neuronal spike events. In the past two decades, numerous spike-sorting algorithms have been developed (e.g., [26–44]) with explicit focus on the reliability of extracted units and the quality of their separation. Some automated sorting algorithms try to increase data quality by using different criteria to detect artificial events (e.g., [27–35]). However, even algorithms adjusted to human single-unit recordings include only elementary methods for the identification and removal of artificial events [27–29,41]. An important and commonly overlooked artifact is the recording of spike events simultaneously in multiple channels. So far, no method has been developed to address this issue in human microwire recordings. We therefore propose a new method for detecting events occurring simultaneously in multiple channels with similar features. We present our method as an open-source and freely available MAT-LAB module that allows researchers to further improve the results of spike sorting from existing cluster algorithms (Combinato Spike Sorting [29]; Wave_clus [27,41]) by removing duplicate spike events.

1.1. Motivation

Recording single units in humans entails a clinical setting and thus a noisy environment from which it is impossible to eliminate all sources of artifacts. In order to optimize the data quality and unit yield, it is crucial to distinguish neural spike events from artificial spike events. Below we discuss prominent sources of noise and the characteristics of these artificial events.

Often, technical aspects of the recording set-up itself can lead to artificial spike events, such as electrical interference, head and cable movements, or broken wires [23]. All of these technical issues originate from non-neural sources and can produce artifacts that may be recorded on several channels simultaneously.

Additionally, duplicate spike events can also originate from physiological sources. For arrays [45], or bundles of microwires, individual micro-electrodes can end up rather close to one another. Since their individual recording volumes can overlap, the same unit can be recorded on different channels [46,47]. Another potential physiological cause of duplicate spike events is a biphasic shape of the signal that is extracted from the same channel exceeding the positive and negative extraction thresholds [31,33,47]. Moreover, a spike event may be recorded on a reference wire and thus appear inverted on multiple channels referenced against this wire [23]. Complicating the matter, physiological coincident spikes may also occur naturally as a reflection of temporal coding in local neuronal networks [48]. Therefore, it is crucial to distinguish between artificial and physiological or pathological coincident spike events.

1.1.1. Artifact Detection Methods

In extracellular recordings of action potentials ('spikes'), various approaches have been used to identify artificial spike events. Most of these approaches can be classified into the following categories:

1. *Artifact detection based on spike shape*

As we have a fairly good understanding of the spike shapes of physiological action potentials [49], the shape of a spike event can provide information on its origin. For example, artificial spike events produced by electrical interference often exhibit a sinusoidal shape [29]. Furthermore, artificial events exhibiting an unphysiologically high amplitude [29,31] can be easily detected.

2. *Artifact detection based on spike timing*

The time course of spike events in a potential unit appearing as a cluster in a spike-sorting algorithm can provide further indications of its origin. For instance, if this cluster contains a high proportion of inter-spike intervals below the physiological refractory period, then these spikes cannot originate from one single unit but instead must be contributions by other units' activity or by artifacts [29]. A different approach is to analyze simultaneous spike events across different channels. Several spike-sorting algorithms identify redundant clusters (e.g., [31,33]) and remove noise outliers in the frequency domain depending on the spatial resolution of electrodes in an array [33].

3. *Manual artifact detection based on spike shape and timing*

Experienced operators can combine spike-shape and spike-timing information to label clusters as artifacts. However, manual evaluation of a large number of clusters is rather time-consuming, especially for a large number of channels, and typically limited to a given recording channel. Effects and interactions occurring simultaneously in different channels are therefore typically disregarded. Recently, there have also been approaches using deep-learning classifiers to automate this process [34,40,43,44,50].

If the spatial configuration of electrodes in polytrodes or arrays is known, then this information can be used to differentiate more reliably between local neural spike events and artifacts [26,30,31,33,35,39,40,42]. Due to the flexible nature of microwire bundles, it is usually not possible to infer their precise spatial configuration.

1.1.2. Characteristics of Coincident Spike Events in Human Single-Unit Recordings

The previous section has demonstrated the need for effective strategies to detect duplicate spike events. To develop our methods, we first looked at the characteristics of simultaneous spike events with respect to their temporal synchrony and shape.

Close examination of our recorded data revealed that simultaneous spike events in different channels often exhibit very similar event shapes (examples in Figures 1A and 2B). Interestingly, this observation was made for simultaneous spike events within the same wire bundle (spike events marked in blue in Figure 1A) as well as spike events occurring in different wire bundles (spike events marked in red in Figure 1A). Furthermore, we found that spike events of the real data occurring within a small time bin show a significantly higher proportion of similar event shapes than spike events from different hemispheres at different time bins ($p = 3.3 \times 10^{-7}$, Wilcoxon signed-rank test, $\Delta t = 50$ μs, median of real population against surrogate population per session, see Section 2.3.2.). This finding demonstrates that simultaneous spike events exhibit more similar shapes.

Regarding the temporal synchrony, Figure 1B illustrates an example of binned event counts in a 15 s data segment pooled across all 80 recording channels. A substantial fraction of the 0.5 ms bins contain two or more spikes. A large proportion of these bins exceed the red line, indicating the mean + 5 standard deviations (σ) of the spike-count-distribution of the original data. It is worth noting that these bins show a temporal clustering (see, e.g., Figure 1B, original data around $t = 7$ s) that likely corresponds to time intervals of poor data quality (e.g., movement artifacts). In order to estimate the rate of coincident spike events, we generated a time-shifted surrogate based on the original data by circularly shifting all extracted spike events, independently for each cluster, by some random offset. This procedure eliminates all effects of simultaneity in the dataset [51]. In this example, none of the bin counts in the time-shifted surrogate exceed the 5σ threshold of the original data, demonstrating that simultaneous spike events occur above chance.
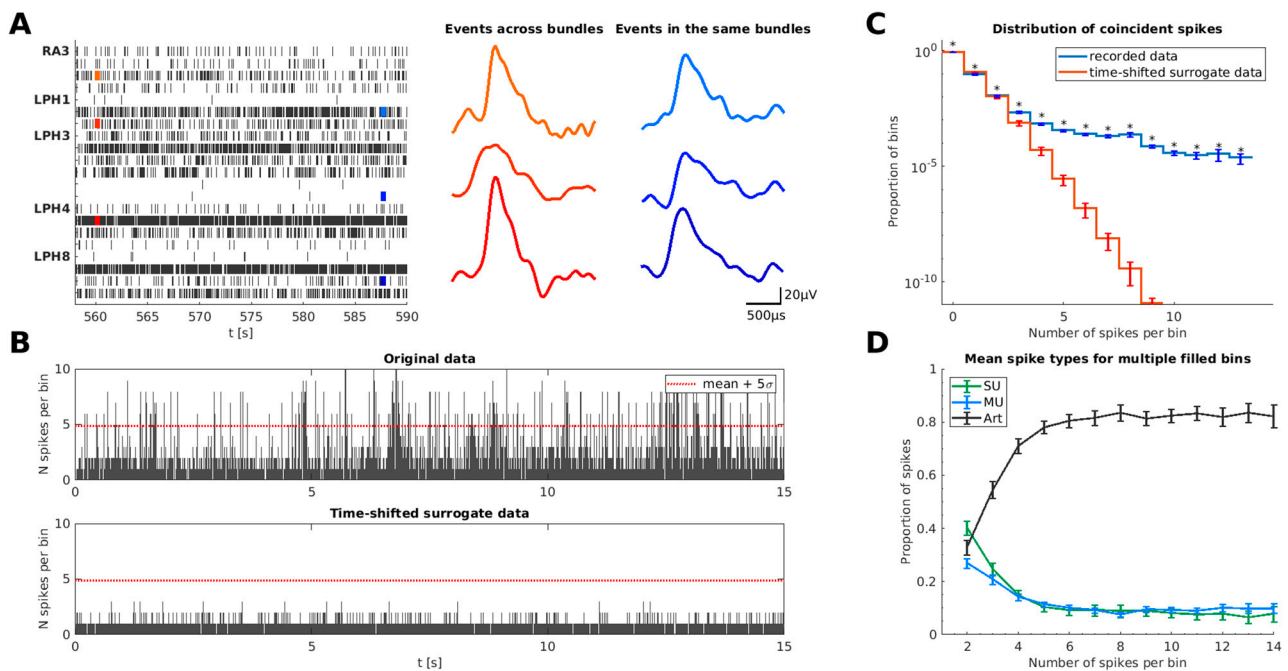
Figure 1. Characteristics of duplicate spike events in human microwire recordings. (**A**) Raster plot of units recorded from 5 microwires located in the right amygdala (RA) and the left posterior hippocampus (LPH). Red and blue lines indicate the timing of simultaneous spike-event shapes drawn next to the raster plot. The three red-shaded spike events occurred simultaneously across bundles while the three blue-shaded events occurred within the same bundle (left hippocampus); (**B**) Binned spike counts of a typical 15 s segment of data. Within each 0.5 ms time bin, the number of spike events across all 80 recording channels was counted. The red dotted horizontal line indicates the mean plus 5 standard deviations (σ) of all bin counts. A substantial fraction of time bins exceeds this 5σ-threshold, indicating a high proportion of simultaneous spike events. The lower panel displays a 15 s segment from the same recording session based on randomly circular time-shifted data for each cluster. In this time-shifted surrogate data, no bin count exceeded the 5σ-threshold of the original data. The *y*-axis is limited to 10 for better visualization; (**C**) Distribution of bins filled with different numbers of spike events across all 51 recording sessions. The blue curve shows the proportion of bins in the original recorded data, indicating a considerable number of bins with 5 and more spikes. The red curve shows the distribution for time-shifted surrogate data of all 51 recordings (1000 permutations, Wilcoxon signed-rank test, $N = 51$, $p < 0.0036$); (**D**) Distribution of spike-event types in filled bins. Spike events in bins with several spikes most frequently originated from artifact clusters. However, some of these events were also found in SU (single units) and MU (multi-unit) clusters (*x*-axis limited to 14 for which more than half of recordings contributed with at least 10 bins).

In order to investigate whether this effect generalizes across recording sessions, we next calculated the frequency of simultaneous spike events across all recorded data used in this study. For this purpose, we counted the number of spike events in each bin across all recordings. Figure 1C shows the abundance of bins filled with different numbers of spike events. Applying the same procedure to 1000 time-shifted surrogate data yielded a distribution for the abundance of bins filled with different numbers of spike events without temporal synchronization of the recorded clusters. For this surrogate dataset, the proportion of these bins decreased rapidly compared to the original dataset. This reveals that our recorded data contain a significantly increased number of simultaneous spike events (see Figure 1C).

In order to better understand from where the spike events in these bins originate, we manually classified clusters based on our spike-sorting algorithm [29] into single units (SU), multi-units (MU), and artifacts (Art) (for details, see Section 2.1.). Figure 1D illustrates the mean proportion of cluster types contributing to the spike events in bins with two or more spikes, averaged across recordings. Time bins with high event counts contain more artifact spikes than bins that are filled with only two simultaneous events. This indicates

that spike events occurring simultaneously in numerous channels most likely belong to artifact clusters.

Using these characteristics of spike events in real data, in the following section, we present our algorithm to detect duplicate artificial events.

2. Materials and Methods

2.1. Data and Materials

In order to develop and optimize our algorithm, we used a dataset of 51 recording sessions from 13 patients with pharmacological intractable epilepsy (for details see Table 1). For diagnostic reasons, patients were implanted bilaterally in the medial temporal lobe (MTL) with 9–12 Behnke-Fried depth electrodes (AD-TECH Medical Instrument Corp., Racine, WI, USA). The exact locations of depth electrodes were exclusively defined for clinical diagnostics. Each electrode contained eight high-impedance and one low-impedance platinum-iridium wires spreading out at the end of its tip. Using these eight high-impedance micro-contacts, we were able to record action potentials from single units. The ninth low-impedance wire was used as a recording reference. The data were collected with an ATLAS recording system (Neuralynx Inc., Bozeman, MT, USA). All data were referentially recorded, filtered at a frequency range of 0.1–9000 Hz, and sampled at 32,768 Hz. We analyzed 51 recording sessions lasting about 25 min each (mean: 25.6 min, *SD*: 3.6 min), that were used to screen for visually responsive units [6,8]. After data collection, spike events were automatically extracted, sorted, and manually evaluated using the Combinato software package [29] (see Appendix A, Table A1). Spike events with positive and negative deflections in Combinato were sorted separately using the default parameters from [29]. Each extracted event shape was sampled by 64 data points spanning a time window of 2 ms. Using the Combinato software package, the sorted clusters were automatically labeled as artifacts or multi-units. During manual evaluation the automated sorting results were checked and optimized, and units were classified as single units, multi-units, or artifacts based on their firing characteristics and spike shapes (see e.g., [46,47]).

Units showing characteristic peaks in their inter-spike interval (*ISI*) histograms stemming from electrical sources (e.g., 50 Hz line noise resulting in peaks at multiples of 20 ms), or having an unphysiological spike-event shape, that were not automatically labeled as artifacts, were merged together with all other artifacts into one artifact cluster. To label a cluster as a single unit, several conditions had to be fulfilled: a physiological waveform in the density plot with a well-defined shape and a steep increase; an asymmetrical spike-event shape with respect to the maximum of the mean cluster shape, and an *ISI* < 3 ms for less than 5% of all spike events in a cluster. Units that were not labeled as artifacts and did not meet all of the above criteria were labeled as multi-units.

Table 1. Overview of recorded sessions used to optimize the DER algorithm.

Patients	Recording Sessions	SU	MU	Art	#Events
13	51	2217	2212	4078	22,341,989

As we are specifically interested in artifacts during human single-unit recordings in a clinical setting, in the following regard, these data serve as a gold-standard since there is no ground truth data available for these types of recordings.

2.2. Structure of the Duplicate Event Removal Algorithm

To identify spike events that are spuriously recorded or detected multiple times, we implemented the duplicate event removal algorithm (DER algorithm) consisting of three parts (see Figure 2):

- **Part I**. Detect simultaneous spike events between different bundles.
- **Part II**. Identify duplicate detected biphasic spike events on the same channel and simultaneous spike events on different channels within the same bundle.
- **Part III**. Detect duplicate spike events based on unphysiologically high zero-lag cross-correlation between clusters.

The open-source code of the DER algorithm and further instructions are accessible at GitHub (https://github.com/Geaht/DER, accessed on 20 May 2021).

2.3. Part I—Detection of Artifacts across Bundles

Sources of noise originating from the environment and the clinical setup are often recorded on several bundles simultaneously. As these artifacts can lead to spike events that look very similar to action potentials of neurons (see Figures 1A and 2C), they can be extracted by automatic spike-sorting algorithms but may not be labeled as artifacts. Part I of the DER algorithm identifies these artifacts by detecting spike events of similar shape recorded in different bundles at the same time.

2.3.1. Detection of Simultaneous Artifacts

If more than two spike events (no_{sim}) appeared within a time window of $\Delta t_{max} = 50$ µs (see Section 2.3.2) in two or more different bundles, we compared the shape of each pair of spike events. We extracted the features of each spike-event shape with a five-level discrete wavelet decomposition (Haar wavelets). Next, we reduced the dimensionality of the feature space to the 10 dimensions in which the distribution of the wavelet coefficients differed most strongly from a normal distribution (quantified with a Kolmogorov–Smirnov test statistic). This feature extraction of spike-event shapes is motivated by the feature extraction used by the spike-sorting algorithms Wave_clus and Combinato [27,29]. For each spike-event pair, the Euclidean distance of their selected wavelet coefficients was calculated as a measure of shape similarity. If the median Euclidean distance of all combinations of events within this time window was below the threshold of $d_{thr} = 14.6$ (see Section 2.3.2 and Table 4), all spike events in this time bin were labeled as artifacts, since only these spike events are likely to fulfill both criteria.

2.3.2. Definition of Thresholds of the Euclidean Distance and the Time Window

In Part I and Part II of the algorithm, two thresholds are needed to define spike events appearing simultaneously with a similar shape: a maximum difference in occurrence time (as a measure of simultaneity) and a maximum Euclidean distance (representing the similarity of two event shapes).

The best proxy for actual duplicate artifacts in our data are spike events occurring simultaneously in different bundles that have been manually labeled as artifacts during the clustering process. To define a threshold for the similarity of coincident artificial spike events in different bundles, we compared two populations of previously labeled artifacts.

From these artifacts we randomly chose 10,000 pairs of spike events per recording session from different hemispheres that did not occur within a time window of Δt (surrogate population) and 10,000 spike-event pairs from different bundles that did occur within this time window (real population). The corresponding two distributions of Euclidean distances (see Figure 3A) were used to calculate the ROC (receiver operating characteristics) curve shown in Figure 3B. Going in steps of 0.1 from zero to the highest Euclidean distance of both populations, each point in the ROC curve was calculated by counting the number of true and false positives as well as true and false negatives. The point on the ROC curve with minimal distance from the point (0,1) was chosen as the operating point (see [52]), representing the best threshold (d_{thr}) to separate the two distributions.
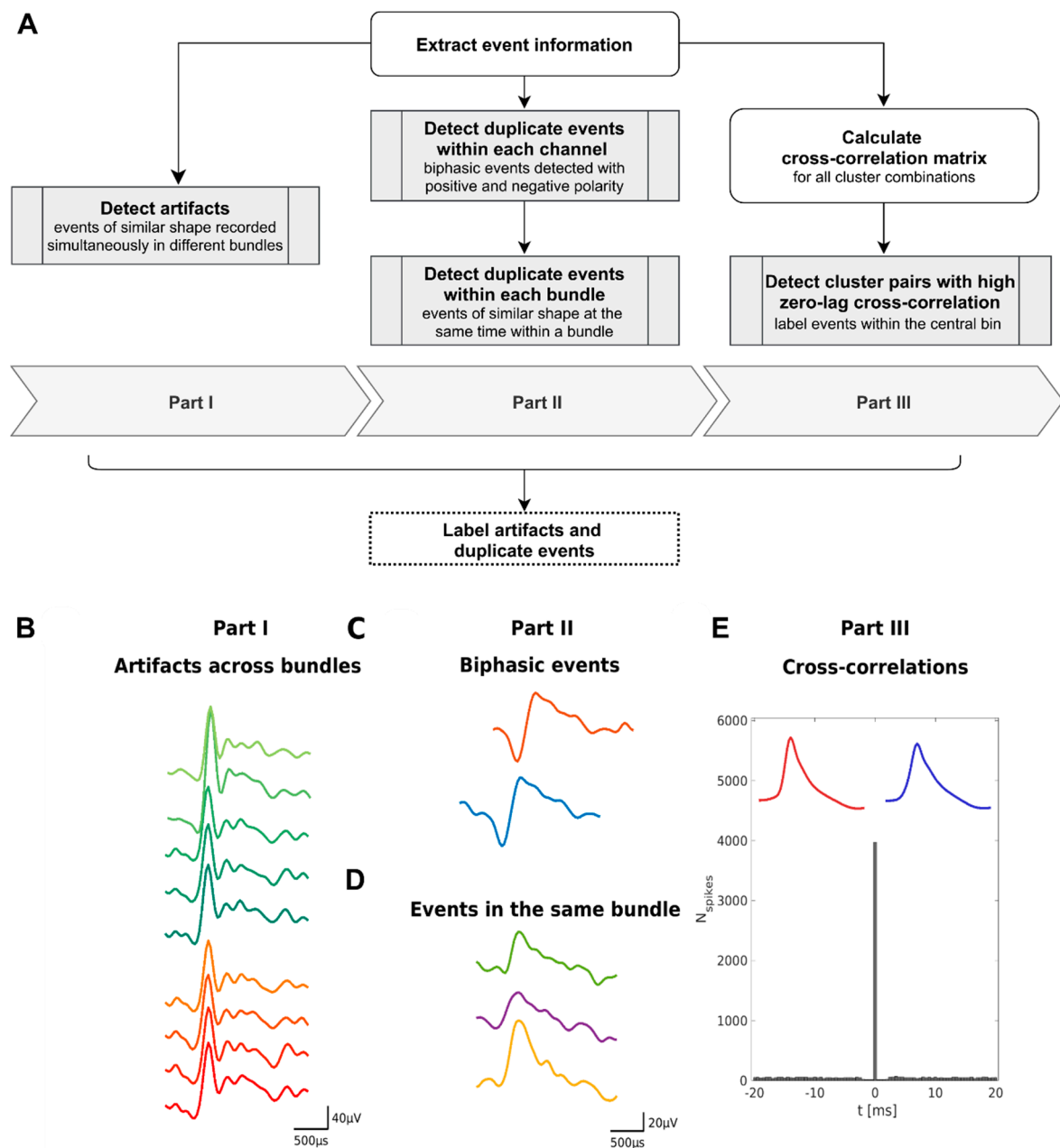
Figure 2. Schematic and illustrative examples for the three parts of the DER algorithm for the detection of spikes events that were recorded or detected multiple times. (**A**) Schematic flow chart of the algorithm. (**B**) In Part I we compare all combinations of the shapes of spike events appearing in a time window of 50 μs on different bundles. An example of event shapes recorded in different bundles is shown (red: left posterior hippocampus (LPH), green: left parahippocampal cortex (LPHC)). (**C,D**) In Part II, we first identify biphasic events that are detected with positive and negative polarity on the same channel within a time window of 650 μs. Panel (**C**) illustrates an example of biphasic event shapes (LPHC 5). Furthermore, similar events within the same wire bundle are identified if they appear within a time window of 50 μs and have highly similar shapes (example in (**D**) from LPHC). In Part III cross-correlograms are calculated for each combination of two clusters. If the central bin exceeds a certain threshold, the spike events within the central bin are considered to be duplicate. Subfigure (**E**) shows an example of the cross-correlograms of two single units that were recorded on two different microwires in the left amygdala. Both units have a large fraction of simultaneous spikes and a similar mean spike shape (shown in red and blue). For each part of our algorithm, we define criteria to determine which event to retain and which to label as artifacts (see also Tables 2 and 3).
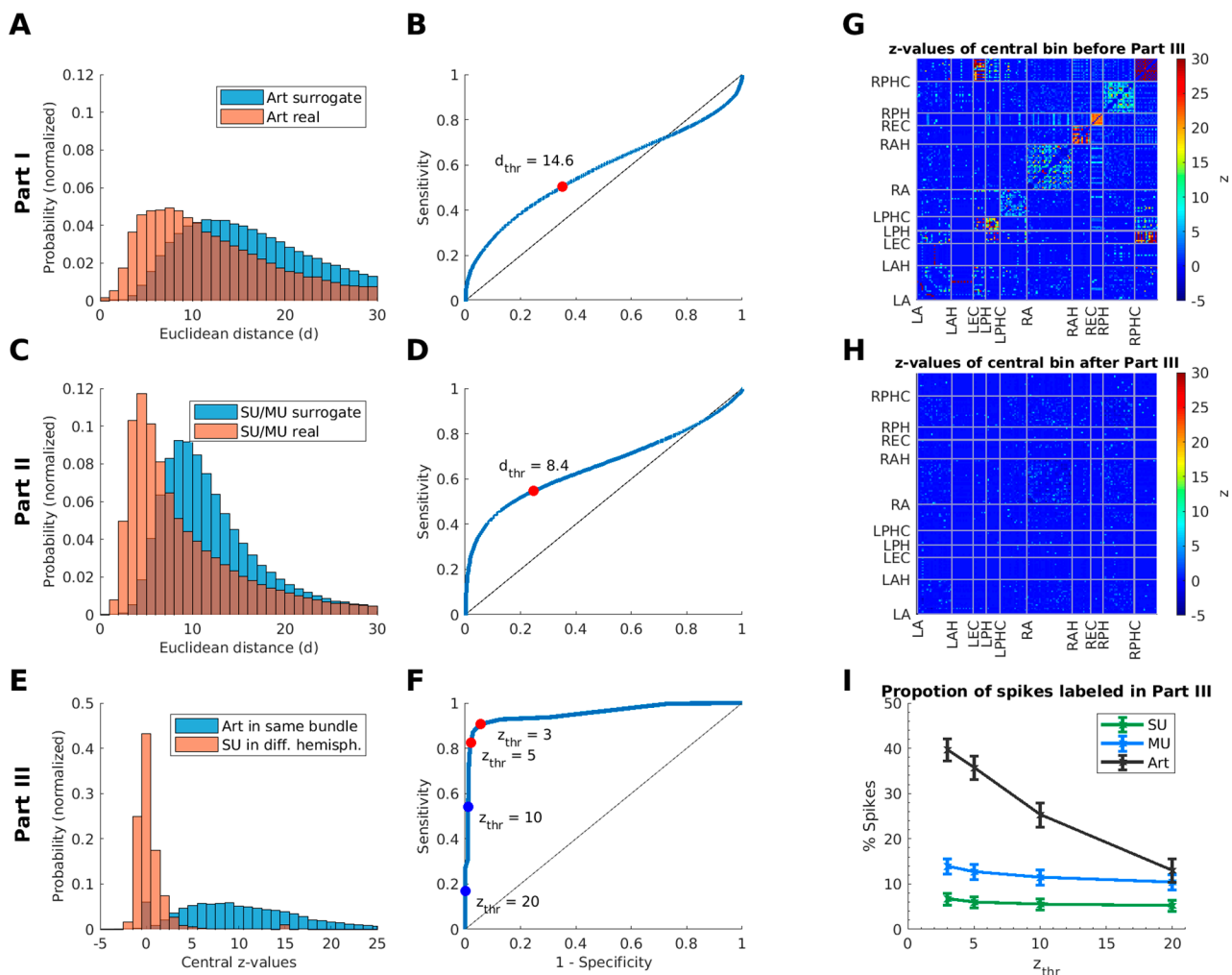
Figure 3. Selection of optimized parameters for the DER algorithm based on the recorded data. (**A**) Distributions of Euclidean distances between simultaneous artifact-cluster events (red) and between non-simultaneous artifact-cluster events randomly drawn from different hemispheres (blue). These distributions were used to determine the threshold of event shape similarity in Part I of the algorithm. (**B**) ROC curve for separating the two distributions of the Euclidean distances in Part I. The operating point on the ROC curve (closest point to the upper left corner) defines the threshold of the Euclidean distance ($d = 14.6$, marked by the red dot). (**C**) Equivalent distributions for joint single- and multi-unit clusters are plotted to define a threshold of the Euclidean distance used in Part II (detection within the same bundle). (**D**) The resulting ROC curve yields a threshold of $d = 8.4$ for Part II (marked by the red dot). (**E**) Distributions of z-values of the central bin in the cross-correlograms between single-unit clusters from different hemispheres (red) and artifact clusters within the same bundle (blue). The corresponding ROC curve is shown in (**F**), including operating points for different thresholds of the central z-value of the cross-correlograms. (**G**) Matrix of central z-values for all cross-correlograms from a recording session with 80 microwires (left (L) or right (R) hemisphere: amygdala (A), anterior hippocampus (AH), entorhinal cortex (EC), posterior hippocampus (PH), parahippocampal cortex (PHC)). Large z-values (red) indicate clusters with a large number of simultaneous spikes; (**H**) Matrix of central z-values for the same recording after removal of all spike events that were detected in Part III of the algorithm ($z_{thr} = 5$). Note that isolated z-values above 5 can result from changes in the background distributions of spike counts in the cross-correlograms. (**I**) Proportion of spikes detected in Part III of the algorithm for different cluster types (SU, MU, and artifacts), averaged across all 51 recording sessions, for different threshold values z_{thr}. Error bars indicate standard error of the mean.

To define the threshold for the maximum time lag of spike events considered as simultaneous, we used these two populations of artifacts and calculated for different time windows (32 µs–1 ms) the operating point in the ROC curve and the area under the ROC curve (AUC) (see Figure A1). The best separation of these two populations was achieved by maximizing the AUC leading to small time windows, especially Δt_{max} = 32 µs and Δt_{max} = 50 µs. As there was virtually no difference in AUC for these two thresholds, we chose Δt_{max} = 50 µs as detection threshold in time, yielding a threshold of the Euclidean distance of d_{thr} = 14.6 (Table 4).

2.4. Part II—Duplicate Events in the Same Bundle

Part II of our detection algorithm focuses on neuronal spike events that were extracted multiple times on the same channel or recorded simultaneously on different wires in the same bundle.

2.4.1. Same Channel

The positive and negative amplitude of biphasic spikes can cross the positive and negative extraction thresholds leading to the same spike event being extracted twice with opposite polarity. Previous studies [53,54] have distinguished putative interneurons from principal cells based on spike duration, using a classification threshold of 650 µs from a spike-shape's peak to its trough (see all parameters in Table 4). As interneurons tend to have a biphasic spike shape, we used this time window to detect duplicate spike events of biphasic shapes. If two spike events on the same channel are detected within that time window and have opposite signs in their amplitudes, they are labeled as duplicate spikes. To decide which of the two spike events to retain and which to label as artificial, we use the following criteria based on the existing unit classification:

- If one of the two duplicate events was labeled as an artifact by an automated clustering algorithm or by manual reclustering, this event is labeled as artificial (Table 2, Case 1).
- If the two events have different unit labels (i.e., one is a single- and the other a multi-unit), we keep the spike event from the single-unit and mark the other one as an artifact (Table 2, Case 3).
- If both spike events are of the same unit class (both single- or multi-unit), we calculate the signal-to-noise ratio (SNR, peak amplitude/spike extraction threshold) of each event and keep the event corresponding to the higher value (Table 2, Case 4). For further details about the thresholds of spike extraction, see [27,29].

Table 2. Criteria for labelling two coincident spikes events in Part II.

Case	Cluster Combination	Spike Events Labelled as Artifacts
1	pair of artifact and (SU or MU) within the same channel	spike event in the artifact cluster
2	pair of artifact and (SU or MU)	both coincident spike events
3	SU and MU in the same bundle	coincident spikes in the MU
4	two SU or two MU in the same bundle	coincident spikes in the lower SNR

2.4.2. Same Bundle

In the next step, we investigated physiological and non-physiological duplicate recorded spike events within the same bundle.

To detect spike events recorded simultaneously in different channels of the same bundle, we looked again for spike events that appeared within a short time interval. For any pair of events fulfilling this criterion, we compared their shape (the Euclidean distance for each pair) as described in Section 2.3.1. To compare event shapes within the same bundle, we took into account that these spike events might be caused by neural action potentials recorded on more than one microwire. One of the criteria to label a unit

as single- or multi-unit is its cluster shape. If the cluster shape is in accordance with a physiologically expected signal [49], this indicates a neuronal origin. Consequently, single- and multi-units have a smaller variety of cluster shapes than artifacts whose origin may vary (see Section 1.1). Therefore, we calculated the thresholds for the detection of duplicate spike events in the same bundle according to Section 2.3.2, but using two populations of SU and MU (SU/MU within the same bundle compared to SU/MU from different hemispheres). This led to the same maximum time difference as in Part I (Δt_{max} = 50 µs) for two spike events to be considered as duplicate detections, and a resulting threshold for their Euclidean distance of d_{thr} = 8.4 (see Figure 3C,D, Figure A1 and Table 4). If the compared spike events fulfilled these criteria, we decided which spike event to retain using similar criteria as in Section 2.4.1 (see Table 2, Case 2–4).

2.5. Part III—Cross-Correlations

A standard method for assessing relationships in the firing patterns of two recorded units is provided by analyzing their cross-correlation (e.g., [55]). In our dataset we identified multiple cross-correlations exhibiting a prominent increase in the central time bin of the cross-correlograms (see e.g., Figure 2E). Besides physiological synchronization, these cross-correlations might originate from simultaneous artifact events on different channels or from units that are recorded with more than one microwire.

2.5.1. Calculation of all Cross-Correlations

In order to identify potential spurious cross-correlations across all possible pairs of recorded units, we calculated for each recording session a cross-correlation matrix C ($N_{cluster} \times N_{cluster} \times N_{time\text{-}bins}$). The cross-correlogram of two clusters i and j can be obtained from C as $C(i,j,:)$. To achieve nearly linear computational scaling with increasing spike-event counts, our algorithm loops only once through a list of all time-ordered spike times of a recording session. Each spike is checked for subsequent spikes within a maximal time lag of the cross-correlation t_{max} (e.g., 20.25 ms for t_{bin} = 0.5 ms; N_{bins} = 81; $t_{max} = \frac{1}{2} t_{bin} \cdot N_{bins}$). For each spike event detected with a time delay of $\Delta t < t_{max}$, the count in the corresponding time bin of C is increased by 1. Since cross-correlations are skew symmetric ($C_{ij} = -C_{ji}$), it is sufficient to consider only subsequent spike events and complete the cross-correlation matrix afterwards. The magnitude of the spike-event counts within the central time-bin (e.g., from −250 µs to +250 µs) can be assessed by calculating a z-score for the central bin based on the mean and standard deviation of spike-event counts in all other time bins of the cross-correlogram. This method yields one z-score for each combination of recorded clusters (see Figure 3G).

2.5.2. Detection of Suspicious Cross-Correlations

If the spiking of two recorded neurons were independent, the cross-correlogram would be expected to be flat without an asymmetry or a central peak. However, there are several physiological and technical reasons that could cause increased simultaneous firing of two units: A possible physiological reason for increased simultaneous firing are two neurons receiving a direct synaptic input from a third neuron (e.g., [55]). An asymmetry in the cross-correlogram can be caused by a direct or indirect synaptic connection between the two neurons [56]. Moreover, sensory inputs may also cause increased simultaneous firing of different neurons, even without direct synaptic connections [57]. Beside physiological reasons, there are several technical reasons that can also lead to an increased number of simultaneous spike events (see Section 1.1). This can result in a high z-score of the central bin of the cross-correlogram.

In order to identify spike events originating from one of the sources described above, we propose the following procedure: First, identify cluster pairs that exceed a given z-threshold for the spike-event count in the central bin ($z_{central} > 5$, see Section 2.5.3). Depending on the cluster pair combination, spike events within the central bin of the cross-correlogram are labeled according to the scheme in Table 3. Spike events labeled this

way might be considered for later deletion as they are likely caused by artifacts (Cases 1 & 2) or represent duplicate recordings of the same neuronal spike events (Cases 3 & 4).

Table 3. Criteria for labelling coincident spikes events in Part III.

Case	Cluster Combination	Spike Events Labelled as Artifacts
1	pair of artifact and (SU or MU)	all coincident spike events
2	two clusters in different wire bundles	all coincident spikes
3	SU and MU in the same bundle	coincident spikes in the MU
4	two SU or two MU in the same bundle	coincident spikes in the lower SNR cluster

2.5.3. Threshold for the Central Bin of the Cross-Correlogram

For the method described above, a threshold z_{thr} must be determined. Coincident spike events in two given clusters which exceed this threshold are considered artificial or duplicate. Such a threshold can be derived from our recorded data by using the distributions of central z-values across different cluster pairs. Artifact clusters within the same wire bundle were recorded from the same anatomical region and shared the same processing electronics (wires, connectors, etc.). Therefore, they are expected to contain a high proportion of coincident spike events. In contrast, single units recorded from different hemispheres should not contain spurious duplicate spike events. In our dataset, we indeed found that the central z-values of SU pairs originating from different hemispheres were significantly smaller than the z-values from pairs of artifact clusters within the same region ($p = 8.9 \times 10^{-16}$, Wilcoxon signed-rank test of the medians across sessions, $N_{Sessions} = 51$).

We analyzed how well these two distributions (Figure 3E) can be separated by a single threshold z_{thr}. All counts below this threshold may be considered to belong to the SU pairs, while all the others are assigned to the artifact pairs. This procedure allowed us to calculate the number of true and false positive as well as true and false-negative counts for different threshold values z_{thr}. The resulting sensitivity and specificity values are plotted as a ROC curve (Figure 3F). The minimal distance to the point (0,1) is given by the operating point at $z = 3$ (as used in Figure 3B,D, see e.g., [52]). This indicates that cluster pairs with a central z-value above 3 in the cross-correlogram resemble pairs of artifact clusters within the same wire bundle rather than independent units. To avoid false positive detections of artifacts (i.e., spuriously discarding neuronal spikes as artifacts), we chose a more conservative threshold of $z_{thr} = 5$ for our algorithm (Table 4). Spike events occurring within this central bin of the cross-correlogram are labeled following the description in Table 3.

For further validation, we analyzed the type of events detected in Part III of our algorithm by identifying from which type of cluster (SU/MU/artifact—as determined by the spike-sorting procedure) they originated. Figure 3I displays the percentage of artifacts, MU, and SU that were detected, averaged across all 51 recording sessions, using different thresholds for the central z-value. For instance, for $z = 5$, about 35% of all manually clustered artifact events are detected by Part III of our algorithm, while only 13% of the MU and just 9% of the SU events were labeled as duplicate spike events. Remarkably, the percentage of artifacts labeled by the algorithm strongly decreases for higher z thresholds while the percentage of labeled MU and SU decreases only marginally. These observations demonstrate that most of the detected spike events correspond to artificial events.

Table 4. Default parameters of the DER algorithm.

Part	Parameter	Threshold	Value
I	min. number of simultaneous spike events	no_{sim}	3
	max. time difference	Δt	50 µs
	max. Euclidean distance of spike-event shapes	d_{thr}	14.6
II	max. time difference in the same channel	$\Delta t_{same\ channel}$	650 µs
	max. time difference in the same bundle	$\Delta t_{same\ bundle}$	50 µs
	max. Euclidean distance of spike-event shapes	d_{thr}	8.4
III	width of time bins in the cross-correlograms	t_{bin}	500 µs
	max. z-value of central bin count in cross-correlograms	z_{thr}	5

3. Results

In this section, we report the performance of our algorithm on real data. First, we present two examples of data that are contaminated by duplicate spike events and demonstrate the improvement achieved by the DER algorithm. We then illustrate the proportions of detected duplicate spike events for the different types of clusters as well as for the different parts of the DER algorithm.

3.1. Examples of Improved Data Quality

Figure 4A shows three raster plots of the same 30 s data segment for all clusters (including clusters labeled manually as artifacts) from two wire bundles in the left posterior hippocampus and entorhinal cortex. The left panel of this figure shows the original data before the DER algorithm. Note the simultaneous spike events on every wire, occurring approximately seven seconds after the beginning of the recording. These spike events are likely caused by external noise sources as they appear on different bundles in parallel. The middle panel illustrates spike events marked by the different parts of the DER algorithm. The suspicious spike events around the seventh second in the raster (and several others) are detected by the algorithm. To illustrate the resulting raster plot we show in the right panel all remaining spike events. A comparison of the raster plots before (left) and after (right) duplicate event removal indicates that the majority of suspicious synchronous spike events were removed, and the data quality was enhanced.

A second example of the improvement in data quality is shown in Figure 4B. We noticed on several occasions that units of similar shape from different microwires within the same wire bundle responded to the same stimuli [2]. Illustrating this phenomenon, Figure 4B (upper two panels; spike events are marked according to the different parts of the DER algorithm using the same colors as in Figure 4A) shows two single units recorded in the right posterior hippocampus (RPH2 and RPH4) within the same bundle, but on different microwires. The first unit (RPH2) shows a clear response to the image of the German comedian Otto Waalkes (stimulus 2). The second unit (RPH4) primarily responds to the German singer Helene Fischer (stimulus 1), but also increases its firing-rate in response to Otto Waalkes. Note that the spikes of the second unit occur at similar time points as the spikes of the first unit during presentation of the second stimulus. This overlap in the response behavior and the similarity of their cluster shapes (see density plots in Figure 4B) hint towards this being a duplicate recording of neuronal spike events.

After applying the DER algorithm to these data, most spike events during the response period in the second unit are detected as duplicate events and are therefore removed (see red framed raster in Figure 4B), while both primary responses remain unchanged. This example highlights the importance of detecting duplicate spike events when investigating the response behavior (e.g., selectivity) of concept cells. Furthermore, statistics in single-unit studies are commonly performed across the population of all recorded units, and single units are often analyzed independently (e.g., [1–12,15–18]).
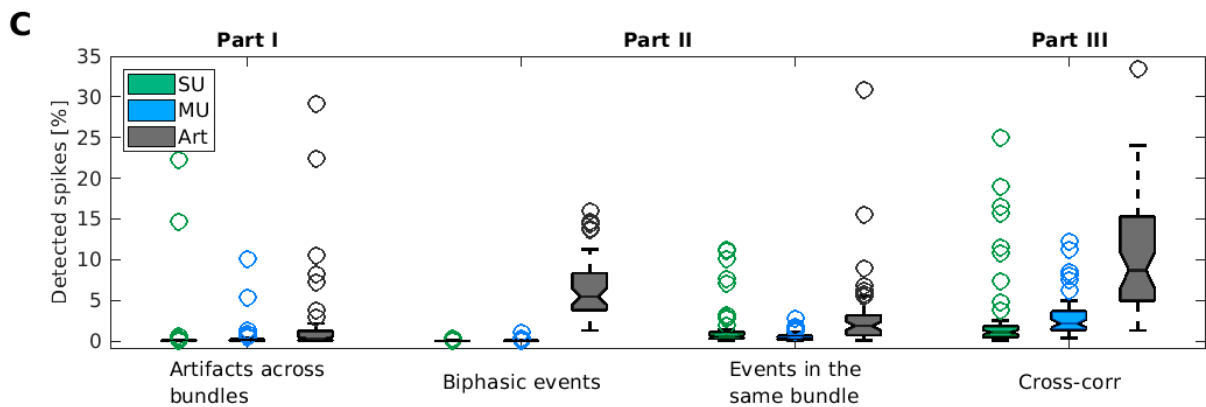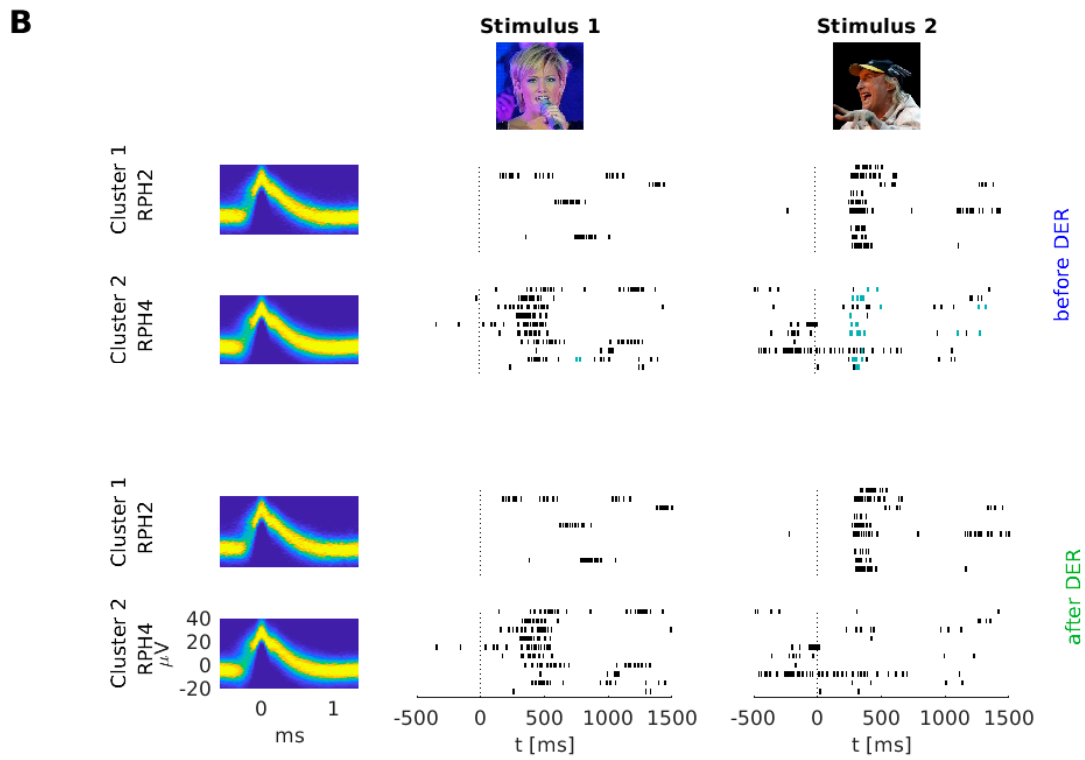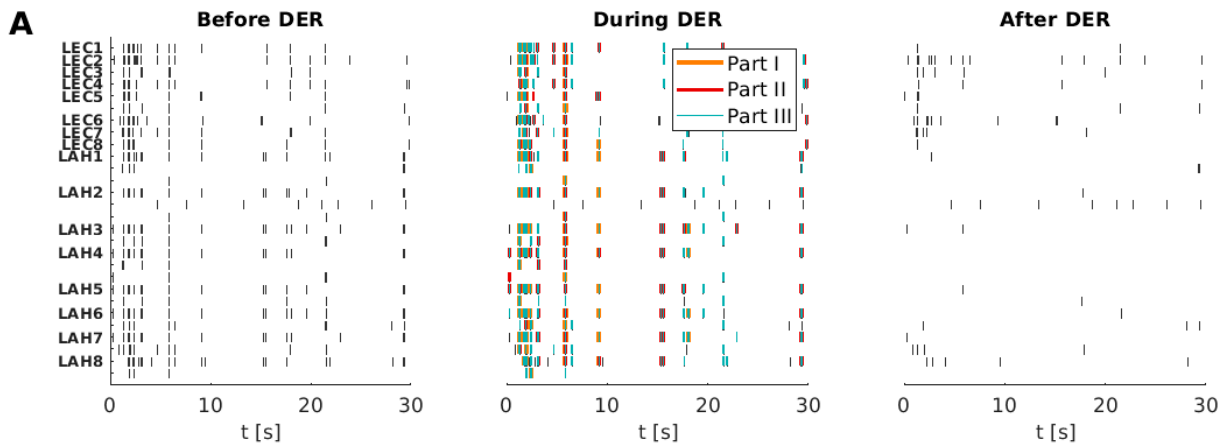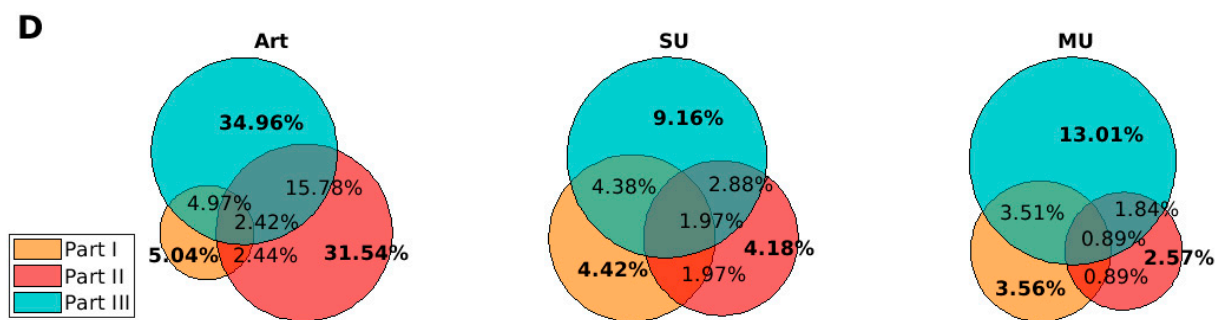
Figure 4. *Cont.*

Figure 4. Performance of the DER algorithm applied to human single-unit recordings. (**A**) Exemplary raster plot of a recording before and after detection of artificial events using the DER algorithm. In the middle raster plot, all detected events are colored according to the part of the algorithm that detected them. (**B**) Raster plots of two units recorded from the same microwire bundle (right posterior hippocampus, RPH) that responded to two different stimuli (left: Stimulus 1—Helene Fischer; right: Stimulus 2—Otto Waalkes). The raster plots are plotted once with all extracted spike events (before DER—upper raster plots; spike events are marked according to the different parts of the DER algorithm using the same colors as in **A**) and once after removal of the duplicate spike events (after DER—lower raster plots). On the left, the spike waveforms of both clusters are shown as density plots. The first unit (RPH 2) responded to the image of the German comedian Otto Waalkes (Stimulus 2). This response displays no notable changes after applying the DER algorithm. The second unit on channel RPH 4 increased its firing rate to stimulus 1 and 2 in the original data. However, after deleting all events detected by the DER algorithm, the increase in firing rate to stimulus 2 was eliminated (red framed raster plot). (**C**) Boxplot of detected spike events (percentage of all spikes) separated by unit class and by the part of the DER algorithm. Part II is subdivided into the detection within the same channel and the same bundle. The *y*-axis is limited to 35% for visualization. (**D**) Percentage of detected spike events per unit class, separated into the three parts of the DER algorithm. Bold numbers indicate the percentage of spikes found in each individual part. Non-bold numbers in the intersections show the percentage of spike events detected by several parts. For the purpose of visualization, intersecting areas are not to scale.

Our algorithm systematically reduces dependencies across recorded units which may originate from duplicate detected spike events.

3.2. Overall Performance of the DER Algorithm

To assess the overall performance of the DER algorithm, we applied it to the complete dataset of 51 recording sessions and analyzed how many spike events were detected by each part of the algorithm. We further separated the detected spike events based on the type of cluster to which they belonged (SU, MU, artifacts).

Altogether, our algorithm marked more than a fifth of all recorded spike events (22.05%) as duplicate and/or artificial (Figure 4C and Table 5). This proportion is rather high since we included all artifact clusters as well as spike detection with positive and negative deflection. In particular, spike events with negative deflections in Combinato were significantly more often detected than spike events with positive defections (50.44% vs. 16.13%, $p = 8.9 \times 10^{-16}$, Wilcoxon signed-rank test across recordings; $N = 51$).

Interestingly, the detection of biphasic spikes within the same channel (Part II—Same channel) deleted 6.46% of all spikes, which were almost exclusively (98.25%) found in artifact clusters, with only 0.55% in SU and 1.20% in multi-unit clusters. This indicates that a high portion of artifacts exhibit a biphasic shape (e.g., a sine-wave-like shape).

Moreover, our algorithm was able to detect more than half (50.77%) of all the events in clusters that were manually labeled as artifacts. The DER algorithm found a smaller but relevant proportion of duplicate spikes in clusters marked as single units (10.50%).

As a control, we analyzed the performance of our algorithm in the same dataset with clusters automatically labeled by Combinato and without manual evaluation. In this dataset, which consisted only of multi-units and artifacts, we detected 19.84% of all spike events (see also Figure A2). Most of the detected spike events in the automatically labeled clusters were also detected in the manually labeled dataset (95.91%). The small difference in the detected spike events affected mostly multi-units (80.32%) in the automatically labeled

dataset. This difference was caused by the criterion to label spike events depending on unit classes (Tables 2 and 3) since the number of clusters labeled as artifacts is lower in the automatically clustered data. This demonstrates that the DER algorithm is suitable for manually labeled as well as automatically sorted data. Therefore, the DER algorithm can be easily integrated in existing fully automated spike-sorting routines.

As the three parts of the algorithm can be run independently on the data, it is possible that spike events are detected in more than one part. To visualize the interactions, we created Venn diagrams of detection overlaps, separating the number of detected spike events into single, multi-units, and artifacts (see Figure 4D). As expected, the highest fraction of duplicate spike events was identified in clusters manually labeled as an artifact. While all three parts show an overlap in the detected duplicate spike events, there is an individual fraction of spike events that is detected within each part, underlining the importance of each step of the algorithm. Most duplicate spike events were detected based on the cross-correlations between clusters (Part III). This is the only step of the algorithm that uses the correlation of clusters across the entire time of recording and can therefore also identify a different population of duplicate spike events than the other parts that only compare simultaneous spike events. Note that Part I has a large overlap with the other two parts. This might be caused by artificial events (e.g., electrical noise) recorded simultaneously across several bundles which also fulfil the detection criteria of Part II or III.

Table 5. Total percentage of spike events detected within each part of the DER algorithm (percentage of all extracted 22 341 989 spike events).

	Part I	Part II		Part III
Unit Class	**Different Bundle**	**Same Channel**	**Same Bundle**	**Cross-Correlation**
Artifacts	1.34%	6.35%	2.26%	9.27%
Multi-units	0.94%	0.08%	0.61%	3.45%
Single units	2.08%	0.04%	1.93%	4.30%

3.3. Estimation of False-Positive Rate

In order to obtain an estimate for the false-positive rate of the DER algorithm, we used three different datasets (manually sorted original data, cluster-wise time-shifted surrogate data, and simulated data [58]). Figure 5 shows the percentage of detected spike events for these three datasets separately for different unit classes and for the different parts of the DER algorithm in which the spike events were detected. Percentages of detected events within the manually sorted data (original, blue bars) are the same as in Table 5 and are shown to visualize the comparison to actual false positives.

To estimate the false positive rate, we altered the manually sorted data by shifting each cluster circularly by a random offset time (cf. Figure 1B). The detected spike events are shown in red in Figure 5, indicating that the false positive rate is in a range of 0.01% with the notable exception of biphasic spike events in artifact clusters. This is due to the merging of artifact clusters in Combinato. Extracting spike events using both a positive and a negative threshold leads to biphasic spike events being extracted and sorted twice, but if both belong to an artifact cluster, they are merged into one cluster as there is only one artifact cluster per channel. Therefore, shifting the data in time by a constant random offset per cluster will not affect the detection of biphasic artifacts within each channel. All other comparisons between different clusters in the same channel are affected, which is reflected by an absolute decrease of 2.3% for artifacts detected by this part of the DER algorithm.

Finally, we estimated the false-positive rate using simulated data [58]. We used 80 individually simulated channels, each containing the activity of 2 to 20 neurons. Single units were computed using a Poisson distribution with a mean firing rate of 0.1–2 Hz (randomly selected), whereas the mean firing rate of multi-units was set to 5 Hz. No artifacts were included in this dataset (70.03% SU & 29.97% MU spike events). The mean firing rate of these simulated channels was 15.78 Hz. The number of detected spike events using our DER algorithm is shown in Figure 5 as yellow bars. The resulting percentages

of false-positive detections are similar to those of the cluster-shifted data except for the biphasic events and the SU spike events detected by Part III. The simulation included only positive spike events, leading to zero detections of biphasic events in Part II of the DER algorithm. The higher fraction of spike events of single-unit clusters detected in Part III is due to the significantly higher firing rates in the simulated data compared to our original dataset (15.78 Hz vs. 4.64 Hz on average) and the larger number of SU clusters in the simulation. This does not affect the detection of multi-units in Part III because of the smaller fraction of multi-units in the simulated data. Thus, both the time-shifted surrogate data and the simulated data demonstrate that our algorithm produces a rather low percentage of false positive detection and thus operates at a rather high specificity.
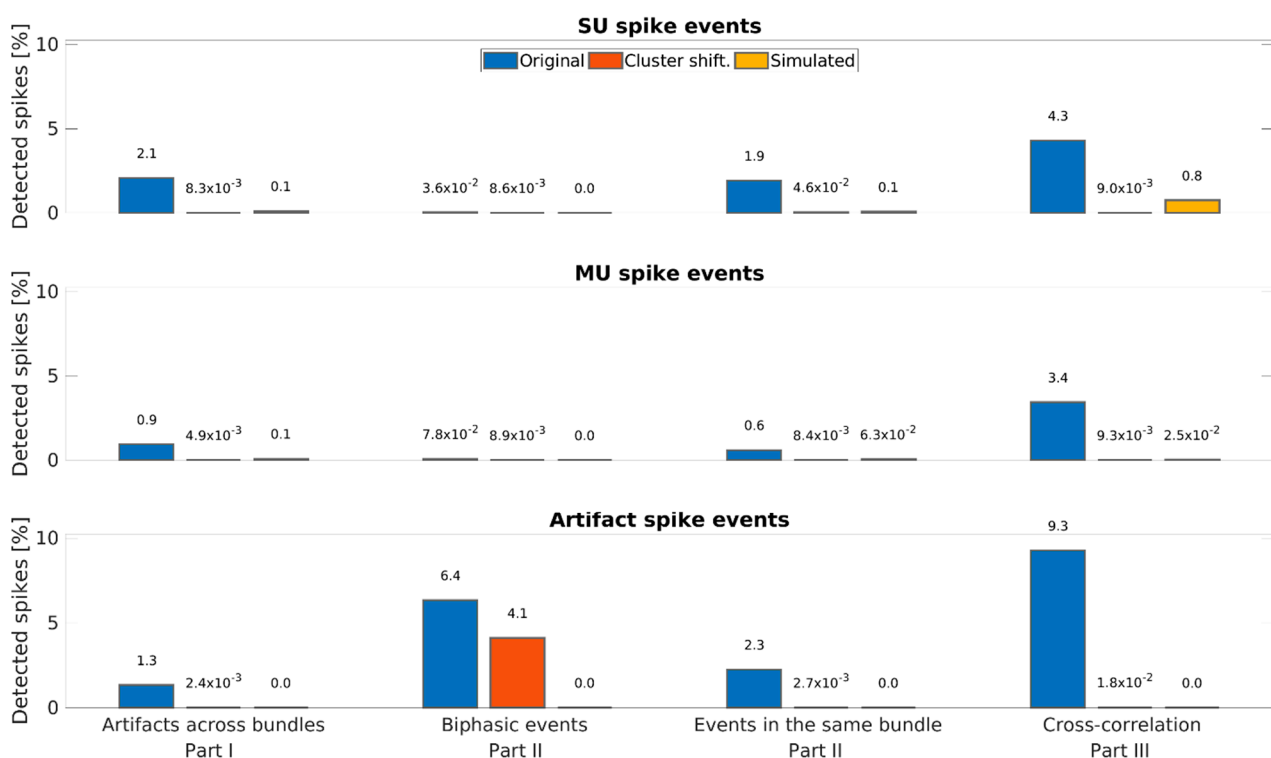
Figure 5. Estimation of the false-positive rate of the DER algorithm. Performance of the DER algorithm for three different datasets: original recorded data are shown in blue, cluster-wise time-shifted surrogate data are shown in red, and simulated data based on [58] are shown in yellow. The top panel shows the percentage of SU spike events found by the different parts of the DER algorithm. The middle and bottom panels show the same analyses for multi-unit and artifact clusters, respectively. The cluster shifted data as well as the simulated data are used to estimate the false-positive rate and thus specificity of the DER algorithm.

As an additional control analysis, we repeated our DER algorithm on several spike extraction thresholds of Combinato (see Figure A3). Combinato's standard extraction threshold of 5 σ seems to be optimal for our data, as the number of detected artifacts using the DER algorithm is strongly reduced compared to 4 σ, whereas only small changes appear when we increase the extraction threshold to 6 σ or 7 σ (while losing many low-amplitude spike events).

The overall results of our algorithm (cf. Figure 4) demonstrate that human single-unit recordings contain a substantial percentage of duplicate spike events. The DER algorithm is an effective way to clean these events and increase overall data quality.

4. Discussion

In this study, we have introduced a method to deal with the problem of coincident event detections in human single-unit recordings. We have demonstrated that coincident spike events appear considerably more often in actual recordings than expected based on a time-shifted surrogate distribution. In human single-unit studies, these duplicate spike events are typically not accounted for, although their removal represents a way to optimize data quality.

The clinical environment in which single units are recorded contains numerous sources of environmental noise. Different approaches can reduce the influence of noise, such as carefully optimizing every aspect of the setup, or identifying and eliminating sources of noise before the recording. It has been analyzed how different implantation-, cutting-, and splicing-techniques may improve data quality [23]. Nevertheless, in a clinical setup it is impossible to eliminate all disruptive influences. Therefore, identifying artificial influences in the data is essential. Algorithms used for the extracting, sorting, and clustering of spike events from neuronal activity recorded by microwire bundles incorporate basic artifact detection. Nevertheless, most of these algorithms do not take into account that artificial events often appear simultaneously on different recording channels. Simulated data are widely used to test spike-sorting algorithms because they allow a comparison of the resulting unit classification to ground truth [27–29,41]. However, this is not feasible for the development of our method as the different noise sources in a clinical recording setup cannot be convincingly simulated. Therefore, we employed a data-driven approach based on a large and reliable dataset of recordings that were manually reclustered. Due to the lack of ground truth telling us whether a detected duplicate spike event represents an artifact or a physiological phenomenon, we employed bootstrap approaches to arrive at reasonable assumptions.

The proposed DER algorithm primarily detected events in clusters that had been labeled as artifacts in our dataset. This observation underlines that our method is in good accordance with human operators. Nevertheless, manual clustering leads to a more conservative dataset, as an operator tends to label an entire cluster as an artifact if it is contaminated by many artificial events even though it might contain some neural events as well. We showed that manual reclustering of automated sorted data is not essential for the detection of duplicate spike events. This entails the possibility to run the DER algorithm also before manual reclustering, minimizing such contaminations and complementing manual evaluation. Alternatively, executing the DER algorithm after manual evaluation allows the algorithm to use the labeling information for detecting artificial or duplicated spikes (see also Tables 2 and 3). Therefore, we recommend using the DER algorithm after a manual evaluation of clusters.

The evaluation of our algorithm on different surrogate datasets (cluster-wise time-shifted and simulated data) demonstrated a low false-positive rate and thus good specificity of the DER algorithm. The precise false-positive rate depends on many factors such as firing rate, number of channels and clusters, percentage of artifacts, etc. The best estimate for our setup is provided by the time-shifted dataset as it conserves these factors. The low percentages of resulting false-positive detections further encourages the use of our algorithm.

Despite the convincing results of the presented method, we recommend that users take the DER algorithm with a grain of salt. For studies specifically investigating coincident spikes (e.g., [56,59–61]), the manipulations of the DER algorithm could be counterproductive. Nevertheless, it is important to note that our algorithm only deletes events within a rather narrow time window (Parts I, II, and III use 0.05 ms, 0.65 ms, and 0.5 ms, respectively; see Table 4). This still facilitates to measure certain asymmetries in the firing of neurons caused by direct synaptic connections [56] as well as common synaptic inputs [62]. The default parameters of the DER algorithm (see Table 4) can be easily adjusted to different recording setups and research questions. It is currently compatible to the spike-sorting algorithms Combinato spike sorter [29] and Wave_clus [27,41] and can easily be adjusted

to others (for further information see https://github.com/Geaht/DER, accessed on 20 May 2021).

All data used in this study were recorded referentially against a low-impedance reference microwire that was stripped of insulation. Our spike-sorting program Combinato extracted spike events by independently applying both positive and negative detection thresholds. This led to biphasic spike events that are extracted twice if both amplitudes (positive and negative) exceeded the extraction threshold of our spike-sorting algorithm. In a recording setup that uses a bipolar montage, simultaneous biphasic events are likely to occur at a much higher frequency as a result of subtracting one channel's activity from another. As the search for simultaneous biphasic spike events within a channel of our DER algorithm (Part II) is calibrated for referential recordings, we recommend adjusting this part of the algorithm for bipolar montages or skipping it completely. The findings of our study encourage future systematic investigation into how duplicate events are affected by wire bundle splicing and cutting. Furthermore, possible influences of different referencing techniques (e.g., local- vs. bipolar-referencing) should be further analyzed. In order to further optimize data quality, it is desirable to combine datasets from different recording sites and identify individual as well as shared noise sources. Including recordings from different amplifier types may also yield additional insights into the origins of artificial spike events. Today, the gold standard for single-unit recordings are intracellular recordings. Combining these with extracellular recordings (e.g., [26,46,47]) and focusing on artificial events that are recorded only on extracellular electrodes would allow us to further improve our understanding of noise sources.

We have demonstrated that, for recordings with microwire bundles in human patients, it is useful to examine possible interactions between different channels. Future single-unit studies should include similar algorithms to deal with this problem.
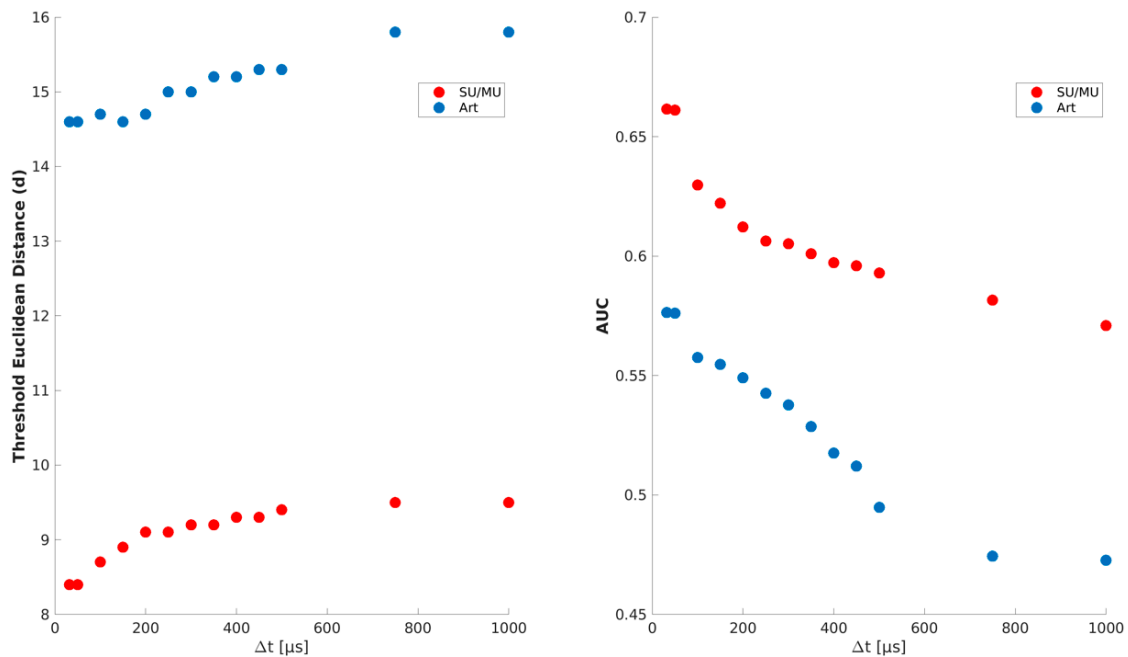
Appendix A

Figure A1. Thresholds of Euclidean distance (**left**) and AUC (**right**) for different time windows. Each point in the left subfigure represents the operating point in an ROC curve calculated from surrogate and real distributions of Euclidean distances for clusters manually labeled as units (MU/SU) or artifacts (Art). See Sections 2.3.2 and 2.4.2.
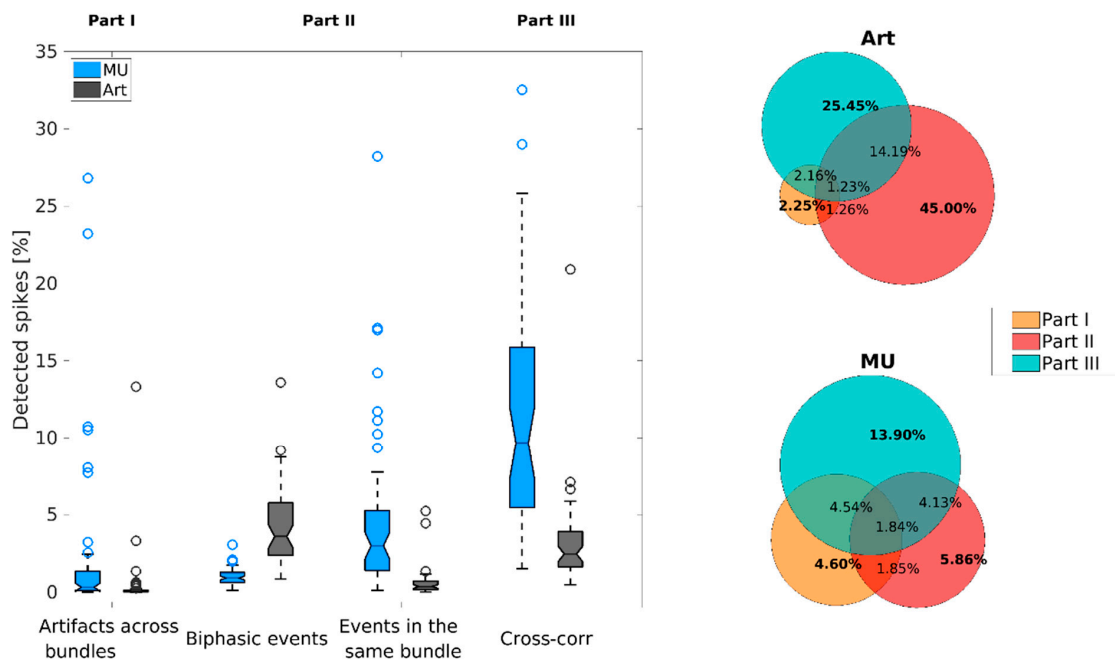

Figure A2. Detected spike events in automatically clustered data (i.e., automated classification of clusters as multi-unit (MU) and artifact only). On the left side, a boxplot of detected spike events (percentage of all extracted spikes from a given cluster class) separated by cluster types and by the parts of the DER algorithm is shown. Part II is subdivided into the detection within the same channel and the same bundle. The *y*-axis is limited to 35% for visualization. On the right side, the percentage of detected spike events per cluster type is shown, separated into the three parts of the algorithm. Bold numbers indicate the percentage of spikes found in each individual part. Non-bold numbers in the intersections show the percentage of spike events detected by several parts. For the purpose of visualization, intersecting areas are not to scale.
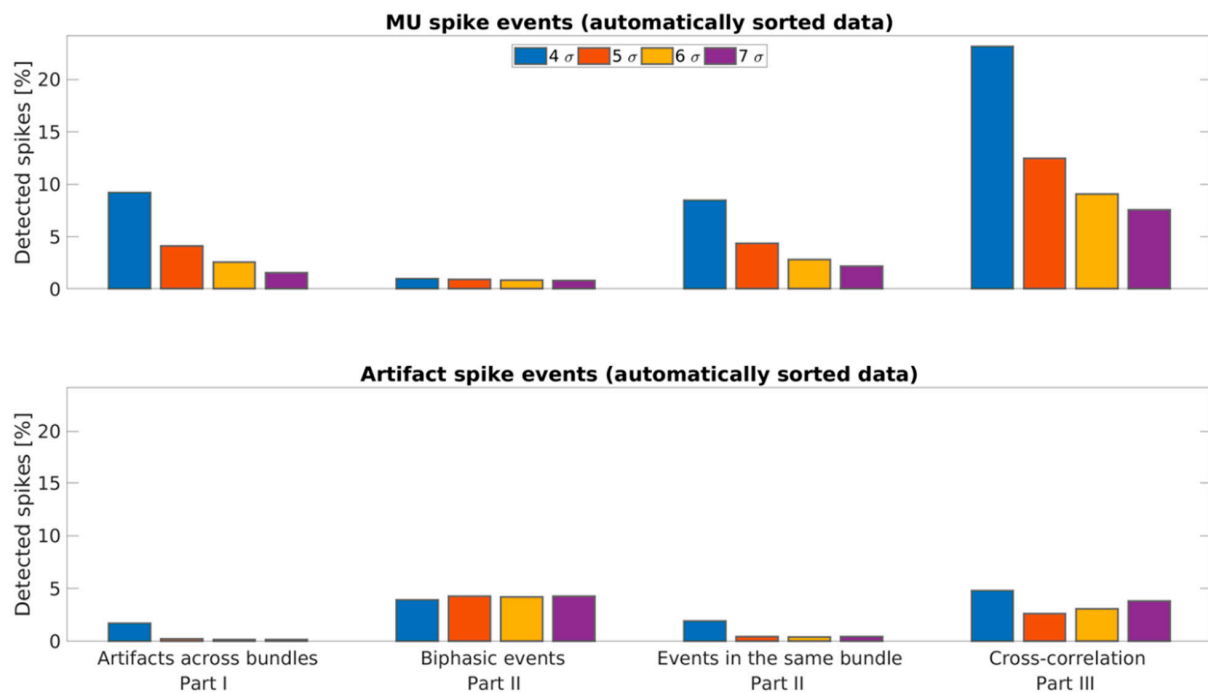
Figure A3. Performance of the DER algorithm for different spike-event extraction thresholds of Combinato on automatically clustered data. The percentage of detected spikes is shown for four different extraction thresholds in different colors (4 σ: blue, 50 million spike events; 5 σ: orange, 22 million spike events; 6 σ: yellow, 13 million spike events; 7 σ: purple, 9 million spike events), separately for multi-units and artifacts and for different parts of the DER algorithm.

Table A1. Resources.

Software	Reference	Source
DER	This paper	https://github.com/Geaht/DER, accessed on 20 May 2021
MATLAB 9.8 Statistics and Machine Learning Toolbox 11.7 Wavelet Toolbox 5.4	MathWorks	https://www.mathworks.com, accessed on 08 April 2020
Octave	GNU	https://www.gnu.org/software/octave, accessed on 1 December 2014
Psychtoolbox	Brainard, 1997 [63]	http://psychtoolbox.org/, accessed on 11 June 2013
Cheetah software	Neuralynx Inc.	https://neuralynx.com/software/cheetah, accessed on 27 August 2012
Combinato Spike Sorting	Niediek et al., 2016 [29]	https://github.com/jniediek/combinato/, accessed on 11 May 2017
Wave_clus 3	Chaure et al., 2018 [41]	https://github.com/csn-le/wave_clus, accessed on 24 October 2020
tightfig.m	Kim Dohyun (2021) [64]	https://de.mathworks.com/matlabcentral/fileexchange/73644-tightfig, accessed on 22 January 2021
Venn.m	Darik (2021) [65]	https://de.mathworks.com/matlabcentral/fileexchange/22282-venn, accessed on 30 November 2020

References

1. Kreiman, G.; Koch, C.; Fried, I. Category-Specific Visual Responses of Single Neurons in the Human Medial Temporal Lobe. *Nat. Neurosci.* **2000**, *3*, 946–953. [CrossRef] [PubMed]
2. Quiroga, R.Q.; Reddy, L.; Kreiman, G.; Koch, C.; Fried, I. Invariant Visual Representation by Single Neurons in the Human Brain. *Nature* **2005**, *435*, 1102–1107. [CrossRef]
3. Mormann, F.; Dubois, J.; Kornblith, S.; Milosavljevic, M.; Cerf, M.; Ison, M.; Tsuchiya, N.; Kraskov, A.; Quiroga, R.Q.; Adolphs, R.; et al. A Category-Specific Response to Animals in the Right Human Amygdala. *Nat. Neurosci.* **2011**, *14*, 1247–1249. [CrossRef]
4. Rutishauser, U.; Ye, S.; Koroma, M.; Tudusciuc, O.; Ross, I.B.; Chung, J.M.; Mamelak, A.N. Representation of Retrieval Confidence by Single Neurons in the Human Medial Temporal Lobe. *Nat. Neurosci.* **2015**, *18*, 1041–1050. [CrossRef] [PubMed]
5. Mormann, F.; Niediek, J.; Tudusciuc, O.; Quesada, C.M.; Coenen, V.A.; Elger, C.E.; Adolphs, R. Neurons in the Human Amygdala Encode Face Identity, but Not Gaze Direction. *Nat. Neurosci.* **2015**, *18*, 1568–1570. [CrossRef]
6. Reber, T.P.; Faber, J.; Niediek, J.; Boström, J.; Elger, C.E.; Mormann, F. Single-Neuron Correlates of Conscious Perception in the Human Medial Temporal Lobe. *Curr. Biol.* **2017**, *27*, 2991–2998.e2. [CrossRef] [PubMed]
7. Mormann, F.; Kornblith, S.; Cerf, M.; Ison, M.J.; Kraskov, A.; Tran, M.; Knieling, S.; Quiroga, R.Q.; Koch, C.; Fried, I. Scene-Selective Coding by Single Neurons in the Human Parahippocampal Cortex. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 1153–1158. [CrossRef] [PubMed]
8. Reber, T.P.; Bausch, M.; Mackay, S.; Boström, J.; Elger, C.E.; Mormann, F. Representation of Abstract Semantic Knowledge in Populations of Human Single Neurons in the Medial Temporal Lobe. *PLoS Biol.* **2019**, *17*, e3000290. [CrossRef]
9. Rey, H.G.; Gori, B.; Chaure, F.J.; Collavini, S.; Blenkmann, A.O.; Seoane, P.; Seoane, E.; Kochen, S.; Quian Quiroga, R. Single Neuron Coding of Identity in the Human Hippocampal Formation. *Curr. Biol.* **2020**, *30*, 1152–1159.e3. [CrossRef]
10. Rutishauser, U.; Mamelak, A.N.; Schuman, E.M. Single-Trial Learning of Novel Stimuli by Individual Neurons of the Human Hippocampus-Amygdala Complex. *Neuron* **2006**, *49*, 805–813. [CrossRef]
11. Staresina, B.P.; Reber, T.P.; Niediek, J.; Boström, J.; Elger, C.E.; Mormann, F. Recollection in the Human Hippocampal-Entorhinal Cell Circuitry. *Nat. Commun.* **2019**, *10*, 1503. [CrossRef]
12. Vaz, A.P.; Wittig, J.H.; Inati, S.K.; Zaghloul, K.A. Replay of Cortical Spiking Sequences during Human Memory Retrieval. *Science* **2020**, *367*, 1131–1134. [CrossRef]
13. Rutishauser, U.; Reddy, L.; Mormann, F.; Sarnthein, J. The Architecture of Human Memory: Insights from Human Single-Neuron Recordings. *J. Neurosci.* **2020**, *41*, 883–890. [CrossRef] [PubMed]
14. Quian Quiroga, R. No Pattern Separation in the Human Hippocampus. *Trends Cogn. Sci.* **2020**, *24*, 994–1007. [CrossRef] [PubMed]
15. Kawasaki, H.; Adolphs, R.; Oya, H.; Kovach, C.; Damasio, H.; Kaufman, O.; Howard, M. Analysis of Single-Unit Responses to Emotional Scenes in Human Ventromedial Prefrontal Cortex. *J. Cogn. Neurosci.* **2005**, *17*, 1509–1518. [CrossRef] [PubMed]
16. Wang, S.; Tudusciuc, O.; Mamelak, A.N.; Ross, I.B.; Adolphs, R.; Rutishauser, U. Neurons in the Human Amygdala Selective for Perceived Emotion. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E3110–E3119. [CrossRef]
17. Mormann, F.; Bausch, M.; Knieling, S.; Fried, I. Neurons in the Human Left Amygdala Automatically Encode Subjective Value Irrespective of Task. *Cereb. Cortex* **2019**, *29*, 265–272. [CrossRef] [PubMed]
18. Unruh-Pinheiro, A.; Hill, M.R.; Weber, B.; Boström, J.; Elger, C.E.; Mormann, F. Single-Neuron Correlates of Decision Confidence in the Human Medial Temporal Lobe. *Curr. Biol.* **2020**, *30*, 4722–4732.e5. [CrossRef]
19. Crandall, P.H.; Walter, R.D.; Rand, R.W. Clinical Applications of Studies on Stereotactically Implanted Electrodes in Temporal-Lobe Epilepsy. *J. Neurosurg.* **1963**, *20*, 827–840. [CrossRef]
20. Bertrand, G.; Jasper, H. Microelectrode Recording of Unit Activity in the Human Thalamus. *Stereotact. Funct. Neurosurg.* **1965**, *26*, 205–208. [CrossRef]
21. Gillingham, J. Forty-Five Years of Stereotactic Surgery for Parkinson's Disease: A Review. *Stereotact. Funct. Neurosurg.* **2000**, *74*, 95–98. [CrossRef]
22. Holtzheimer, P.E.; Husain, M.M.; Lisanby, S.H.; Taylor, S.F.; Whitworth, L.A.; McClintock, S.; Slavin, K.V.; Berman, J.; McKhann, G.M.; Patil, P.G.; et al. Subcallosal Cingulate Deep Brain Stimulation for Treatment-Resistant Depression: A Multisite, Randomised, Sham-Controlled Trial. *Lancet Psychiatry* **2017**, *4*, 839–849. [CrossRef]
23. Misra, A.; Burke, J.F.; Ramayya, A.G.; Jacobs, J.; Sperling, M.R.; Moxon, K.A.; Kahana, M.J.; Evans, J.J.; Sharan, A.D. Methods for Implantation of Micro-Wire Bundles and Optimization of Single/Multi-Unit Recordings from Human Mesial Temporal Lobe. *J. Neural Eng.* **2014**, *11*, 026013. [CrossRef] [PubMed]
24. McNaughton, B.L.; O'Keefe, J.; Barnes, C.A. The Stereotrode: A New Technique for Simultaneous Isolation of Several Single Units in the Central Nervous System from Multiple Unit Records. *J. Neurosci. Methods* **1983**, *8*, 391–397. [CrossRef]
25. Gray, C.M.; Maldonado, P.E.; Wilson, M.; McNaughton, B. Tetrodes Markedly Improve the Reliability and Yield of Multiple Single-Unit Isolation from Multi-Unit Recordings in Cat Striate Cortex. *J. Neurosci. Methods* **1995**, *63*, 43–54. [CrossRef]
26. Harris, K.D.; Henze, D.A.; Csicsvari, J.; Hirase, H.; Buzsáki, G. Accuracy of Tetrode Spike Separation as Determined by Simultaneous Intracellular and Extracellular Measurements. *J. Neurophysiol.* **2000**, *84*, 401–414. [CrossRef] [PubMed]
27. Quiroga, R.Q.; Nadasdy, Z.; Ben-Shaul, Y. Unsupervised Spike Detection and Sorting with Wavelets and Superparamagnetic Clustering. *Neural Comput.* **2004**, *16*, 1661–1687. [CrossRef] [PubMed]
28. Rutishauser, U.; Schuman, E.M.; Mamelak, A.N. Online Detection and Sorting of Extracellularly Recorded Action Potentials in Human Medial Temporal Lobe Recordings, in Vivo. *J. Neurosci. Methods* **2006**, *154*, 204–224. [CrossRef]

29. Niediek, J.; Boström, J.; Elger, C.E.; Mormann, F. Reliable Analysis of Single-Unit Recordings from the Human Brain under Noisy Conditions: Tracking Neurons over Hours. *PLoS ONE* **2016**, *11*, e0166598. [CrossRef] [PubMed]

30. Rossant, C.; Kadir, S.N.; Goodman, D.F.M.; Schulman, J.; Hunter, M.L.D.; Saleem, A.B.; Grosmark, A.; Belluscio, M.; Denfield, G.H.; Ecker, A.S.; et al. Spike Sorting for Large, Dense Electrode Arrays. *Nat. Neurosci.* **2016**, *19*, 634–641. [CrossRef]

31. 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.e6. [CrossRef]

32. Keshtkaran, M.R.; Yang, Z. Noise-Robust Unsupervised Spike Sorting Based on Discriminative Subspace Learning with Outlier Handling. *J. Neural Eng.* **2017**, *14*, 036003. [CrossRef] [PubMed]

33. Jun, J.J.; Mitelut, C.; Lai, C.; Gratiy, S.L.; Anastassiou, C.A.; Harris, T.D. Real-Time Spike Sorting Platform for High-Density Extracellular Probes with Ground-Truth Validation and Drift Correction. *BioRxiv* **2017**, 101030.

34. Saif-ur-Rehman, M.; Lienkämper, R.; Parpaley, Y.; Wellmer, J.; Liu, C.; Lee, B.; Kellis, S.; Andersen, R.; Iossifidis, I.; Glasmachers, T.; et al. SpikeDeeptector: A Deep-Learning Based Method for Detection of Neural Spiking Activity. *J. Neural Eng.* **2019**, *16*, 056003. [CrossRef]

35. Pachitariu, M.; Steinmetz, N.A.; Kadir, S.N.; Carandini, M.; Harris, K.D. Fast and Accurate Spike Sorting of High-Channel Count Probes with KiloSort. In *Advances in Neural Information Processing Systems 29 (NIPS 2016)*; NIPS Proceedings: Barcelona, Spain, 2016.

36. Pouzat, C.; Mazor, O.; Laurent, G. Using Noise Signature to Optimize Spike-Sorting and to Assess Neuronal Classification Quality. *J. Neurosci. Methods* **2002**, *122*, 43–57. [CrossRef]

37. Schmitzer-Torbert, N.; Jackson, J.; Henze, D.; Harris, K.; Redish, A.D. Quantitative Measures of Cluster Quality for Use in Extracellular Recordings. *Neuroscience* **2005**, *131*, 1–11. [CrossRef]

38. Knieling, S.; Sridharan, K.S.; Belardinelli, P.; Naros, G.; Weiss, D.; Mormann, F.; Gharabaghi, A. An Unsupervised Online Spike-Sorting Framework. *Int. J. Neural Syst.* **2016**, *26*, 1550042. [CrossRef]

39. Hilgen, G.; Sorbaro, M.; Pirmoradian, S.; Muthmann, J.-O.; Kepiro, I.E.; Ullo, S.; Ramirez, C.J.; Puente Encinas, A.; Maccione, A.; Berdondini, L.; et al. Unsupervised Spike Sorting for Large-Scale, High-Density Multielectrode Arrays. *Cell Rep.* **2017**, *18*, 2521–2532. [CrossRef]

40. Lee, J.; Carlson, D.; Shokri, H.; Yao, W.; Goetz, G.; Hagen, E.; Batty, E.; Chichilnisky, E.; Einevoll, G.; Paninski, L. YASS: Yet Another Spike Sorter. *BioRxiv* **2017**, 151928. [CrossRef]

41. Chaure, F.J.; Rey, H.G.; Quian Quiroga, R. A Novel and Fully Automatic Spike-Sorting Implementation with Variable Number of Features. *J. Neurophysiol.* **2018**, *120*, 1859–1871. [CrossRef]

42. 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**, *7*, e34518. [CrossRef]

43. Rácz, M.; Liber, C.; Németh, E.; Fiáth, R.; Rokai, J.; Harmati, I.; Ulbert, I.; Márton, G. Spike Detection and Sorting with Deep Learning. *J. Neural Eng.* **2020**, *17*, 016038. [CrossRef]

44. Park, I.Y.; Eom, J.; Jang, H.; Kim, S.; Park, S.; Huh, Y.; Hwang, D. Deep Learning-Based Template Matching Spike Classification for Extracellular Recordings. *Appl. Sci.* **2020**, *10*, 301. [CrossRef]

45. Chapeton, J.I.; Haque, R.; Wittig, J.H.; Inati, S.K.; Zaghloul, K.A. Large-Scale Communication in the Human Brain Is Rhythmically Modulated through Alpha Coherence. *Curr. Biol.* **2019**, *29*, 2801–2811.e5. [CrossRef] [PubMed]

46. Henze, D.A.; Borhegyi, Z.; Csicsvari, J.; Mamiya, A.; Harris, K.D.; Buzsáki, G. Intracellular Features Predicted by Extracellular Recordings in the Hippocampus In Vivo. *J. Neurophysiol.* **2000**, *84*, 390–400. [CrossRef]

47. Buzsáki, G. Large-Scale Recording of Neuronal Ensembles. *Nat. Neurosci.* **2004**, *7*, 446–451. [CrossRef]

48. Singer, W. Neuronal Synchrony: A Versatile Code for the Definition of Relations? *Neuron* **1999**, *24*, 49–65. [CrossRef]

49. Gold, C.; Henze, D.A.; Koch, C.; Buzsáki, G. On the Origin of the Extracellular Action Potential Waveform: A Modeling Study. *J. Neurophysiol.* **2006**, *95*, 3113–3128. [CrossRef]

50. Eom, J.; Kim, S.; Jang, H.; Shin, H.; Hwang, J.H.; Park, S.; Huh, Y.; Choi, H.J.; Hwang, D. Neural Spike Classification via Deep Neural Network. *IBRO Rep.* **2019**, *6*, S139–S140. [CrossRef]

51. Grün, S. Data-Driven Significance Estimation for Precise Spike Correlation. *J. Neurophysiol.* **2009**, *101*, 1126–1140. [CrossRef] [PubMed]

52. Song, B.; Zhang, G.; Zhu, W.; Liang, Z. ROC Operating Point Selection for Classification of Imbalanced Data with Application to Computer-Aided Polyp Detection in CT Colonography. *Int. J. Comput. Assist. Radiol. Surg.* **2014**, *9*, 79–89. [CrossRef]

53. Ison, M.J.; Mormann, F.; Cerf, M.; Koch, C.; Fried, I.; Quiroga, R.Q. Selectivity of Pyramidal Cells and Interneurons in the Human Medial Temporal Lobe. *J. Neurophysiol.* **2011**, *106*, 1713–1721. [CrossRef]

54. Gast, H.; Niediek, J.; Schindler, K.; Boström, J.; Coenen, V.A.; Beck, H.; Elger, C.E.; Mormann, F. Burst Firing of Single Neurons in the Human Medial Temporal Lobe Changes before Epileptic Seizures. *Clin. Neurophysiol.* **2016**, *127*, 3329–3334. [CrossRef] [PubMed]

55. Perkel, D.H.; Gerstein, G.L.; Moore, G.P. Neuronal Spike Trains and Stochastic Point Processes: II. Simultaneous Spike Trains. *Biophys. J.* **1967**, *7*, 419–440. [CrossRef]

56. Csicsvari, J.; Hirase, H.; Czurko, A.; Buzsáki, G. Reliability and State Dependence of Pyramidal Cell–Interneuron Synapses in the Hippocampus. *Neuron* **1998**, *21*, 179–189. [CrossRef]

57. Maldonado, P.E.; Friedman-Hill, S.; Gray, C.M. Dynamics of Striate Cortical Activity in the Alert Macaque: II. Fast Time Scale Synchronization. *Cereb. Cortex* **2000**, *10*, 1117–1131. [CrossRef]

58. Pedreira, C.; Martinez, J.; Ison, M.J.; Quian Quiroga, R. How Many Neurons Can We See with Current Spike Sorting Algorithms? *J. Neurosci. Methods* **2012**, *211*, 58–65. [CrossRef]

59. Bair, W.; Zohary, E.; Newsome, W.T. Correlated Firing in Macaque Visual Area MT: Time Scales and Relationship to Behavior. *J. Neurosci.* **2001**, *21*, 1676–1697. [CrossRef] [PubMed]

60. Fujisawa, S.; Amarasingham, A.; Harrison, M.T.; Buzsáki, G. Behavior-Dependent Short-Term Assembly Dynamics in the Medial Prefrontal Cortex. *Nat. Neurosci.* **2008**, *11*, 823–833. [CrossRef] [PubMed]

61. Harris, K.D.; Hirase, H.; Leinekugel, X.; Henze, D.A.; Buzsáki, G. Temporal Interaction between Single Spikes and Complex Spike Bursts in Hippocampal Pyramidal Cells. *Neuron* **2001**, *32*, 141–149. [CrossRef]

62. Ostojic, S.; Brunel, N.; Hakim, V. How Connectivity, Background Activity, and Synaptic Properties Shape the Cross-Correlation between Spike Trains. *J. Neurosci.* **2009**, *29*, 10234–10253. [CrossRef] [PubMed]

63. Brainard, D.H. The psychophysics toolbox. *Spat. Vis.* **1997**, *10*, 433–436. [CrossRef] [PubMed]

64. Available online: https://de.mathworks.com/matlabcen-tral/fileexchange/73644-tightfig (accessed on 22 January 2021).

65. Available online: https://de.mathworks.com/matlabcen-tral/fileexchange/22282-venn (accessed on 30 November 2020).