Unsupervised spike sorting for large scale, high density multielectrode arrays
Supplementary material

Hilgen et al.

Shape-based event filtering

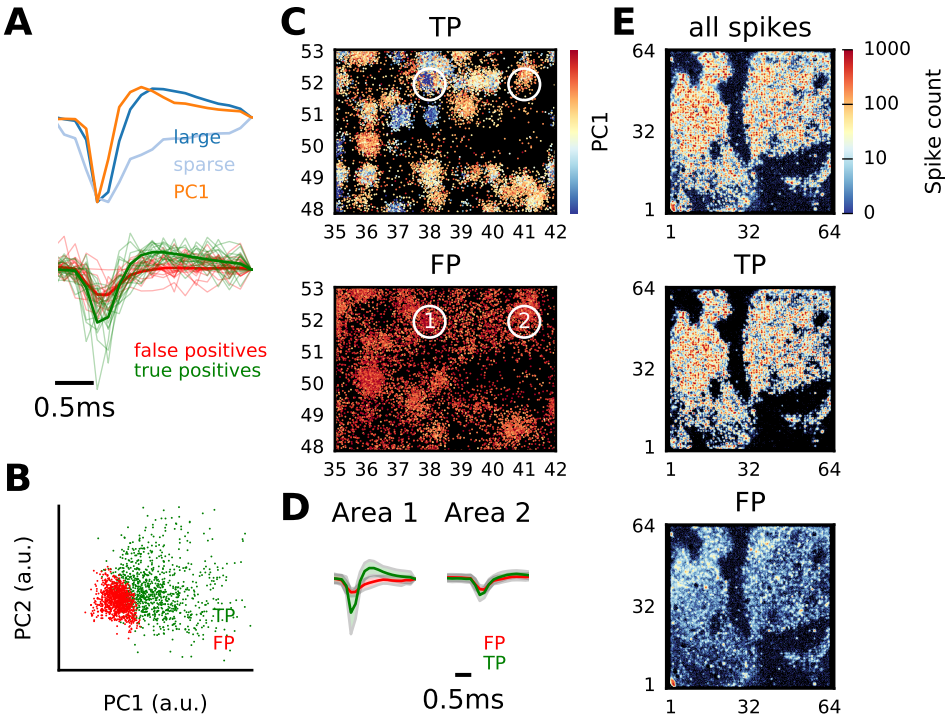
Spike detection was generally performed with a low threshold to reduce false negatives. In particular in recordings with low sampling rates (7 kHz), this could yield events that were clearly not spikes, but threshold crossings due to noise. For instance, a small number of events are usually detected in areas of the chip not covered by tissue (see Figure 1E, top). This could be easily avoided by increasing the detection threshold, or by introducing additional shape criteria to remove such events during the detection phase. We however found that this typically leads to an unreasonable increase in the fraction of false negatives (Muthmann et al., 2015), and it is generally difficult to determine appropriate detection parameters a priori to avoid false positives consistently. Therefore, a method for post hoc event selection was developed, which also helped to compensate for variations between preparations.

To distinguish between true and false positives, examples of events with either high amplitudes or from areas with very low spike density were randomly chosen from the data and used to train a radial basis function Support Vector Machine (SVM) classifier. High amplitude events showed the typical biphasic spike waveform, which resembled the first principal component (PC) estimated from all events, while low density events lacked the repolarization (Figure 1A). This confirmed the premise that regions with very low spike density contained almost exclusively noise. A classifier trained on these examples separated events roughly, but not exactly along projections along the first PC (Figure 1B).

The comparison in Figure 1C showed that events classified as true positives were typically part of localized spatial density peaks, while false positives were more homogeneously distributed with low density (Figure 1C). True positives show clear, biphasic waveforms (see area 1 in Figure 1C,D). Yet the separation would still remain ambiguous for small events with amplitudes closer to the noise level, where events classified as false positives still showed spatial clustering (area 2 in Figure 1C,D), suggesting the presence of poorly detected current sources as well as other events related to neural activity such as strong synaptic currents (Muthmann et al., 2015). As shown in Figure 1E for a recording from 64x64 channels at 7kHz, most of the spatial structure was retained in the map of spikes classified as true positives, while events in areas where no spikes were expected (e.g. optic disk, incisions) were correctly removed. On the other hand, the map of false positives showed weak spatial clustering in areas with high activity. This indicated that the spike record of some neurons with weak signals was most likely incomplete, and a further selection of events had to be performed after spike sorting.

References

Muthmann, J.-O., Amin, H., Sernagor, E., Maccione, A., Panas, D., Berdondini, L., Bhalla, U. S., & Hennig, M. H. (2015). Spike Detection for Large Neural Populations Using High Density Multielectrode Arrays. Frontiers in Neuroinformatics, 9, 1–21.



Supplementary Figure S1 Filtering of detected spikes through spike shape classification.
(A) Average waveforms of events sampled from areas with low event density ("sparse") and with high amplitude ("large") samples, which were used to train the classifier (top). Example waveforms of events classified as true and false positives are shown below.
(B) Projections along the first two principal components (PCs) of waveforms classified as true and false positives (TP, green, and FP, red, respectively).
(C) Events classified as true (top) and false positives (bottom), at their estimated locations. Color indicates the projection along the first PC.
(D) Average waveforms of all TP and FP in the two circled areas in panel C.
(E) Spatial event density maps for a complete recording. Shown are all spikes (top), true (middle) and false positives (bottom). Data in this figure are from the same retina as in Figure 1 of the main text, but recorded at 7 kHz.