Statistical Validation of Spike Patterns Revealed By Fuzzy Clustering Algorithms

984.22

V.J. Toups[1]; J.M. Fellous[2]; P.H.E. Tiesinga[1]

1. Physics Department, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

2. Department of Biomedical Engineering at Dook University, Durham, NC, USA

➲Spike patterns in the form of spike time correlations have been observed in in vivo and in vitro experiments [Fellous, et al 2004].

➲Patterns are computationally relevant:
A) spike timing dependent plasticity
B) short term synaptic plasticity [Tiesinga 2004]

➲Patterns are relevant for cross-correlation analysis because the common assumption of Poisson statistics may not be correct.

➲ Our goal is to develop a principled way of extracting and validating spike patterns.

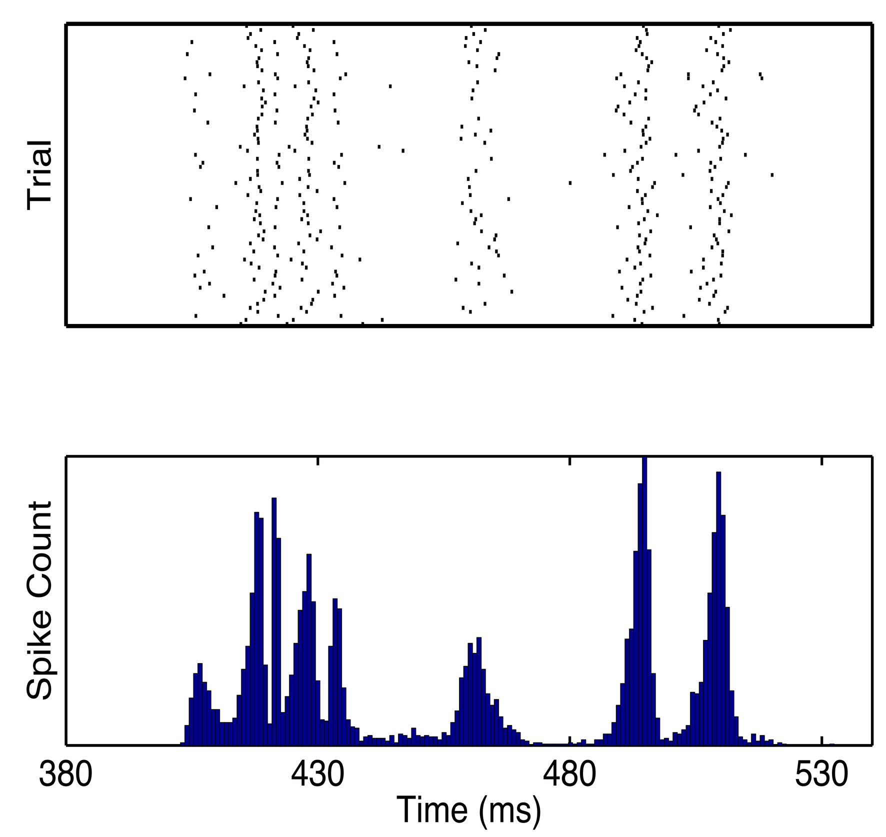

Figure 1: Spike patterns are impossible to detect in a histogram and difficult to find by eye in a rastergram. A clustering procedure must be used. Are there patterns in this data?

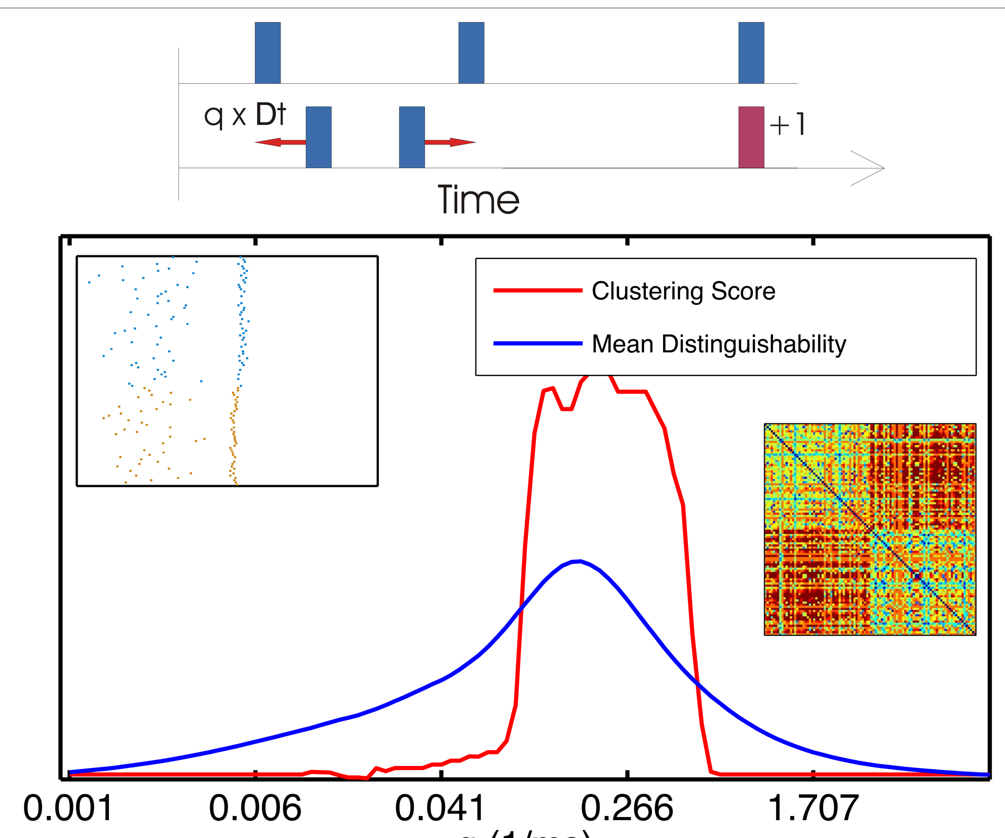

Figure 2: 'q' matters. The distance between spike trains has to be calculated in a controlled way. The q value in the Victor-Purpura metric parameterizes the trade-off between precision and reliability. For this simple set, only a small range of q values leads to correct clustering. The dataset contains two patterns each with a high precision and a low precision event.

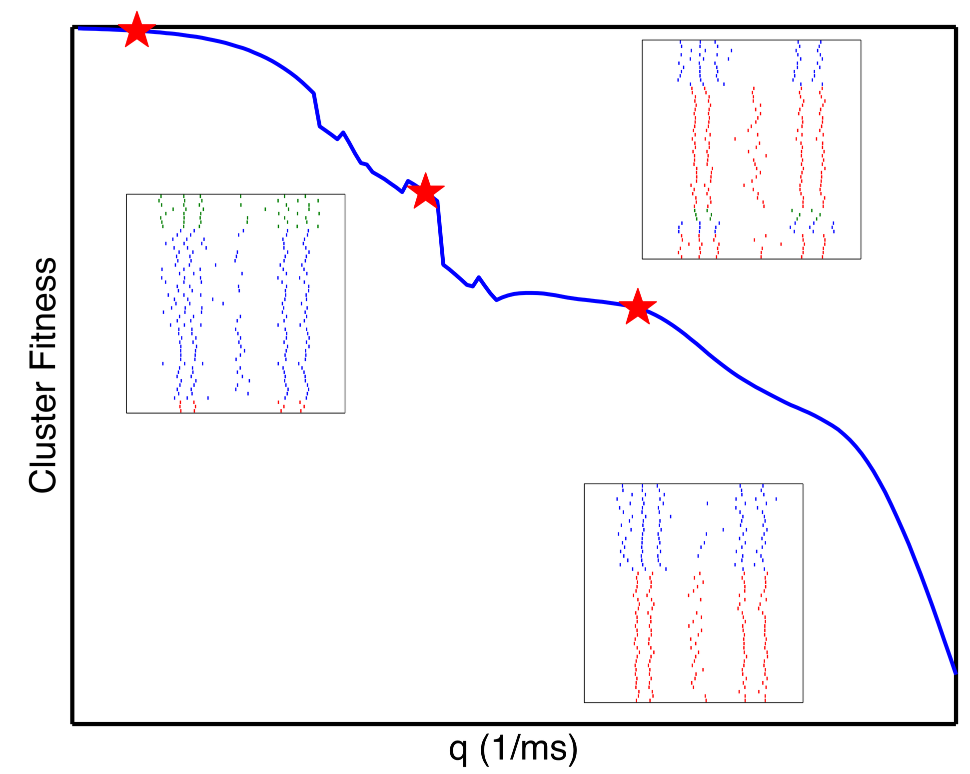

Figure 3: As q is varied a number of 'good' clusterings are observed. Which of these corresponds to a real pattern?

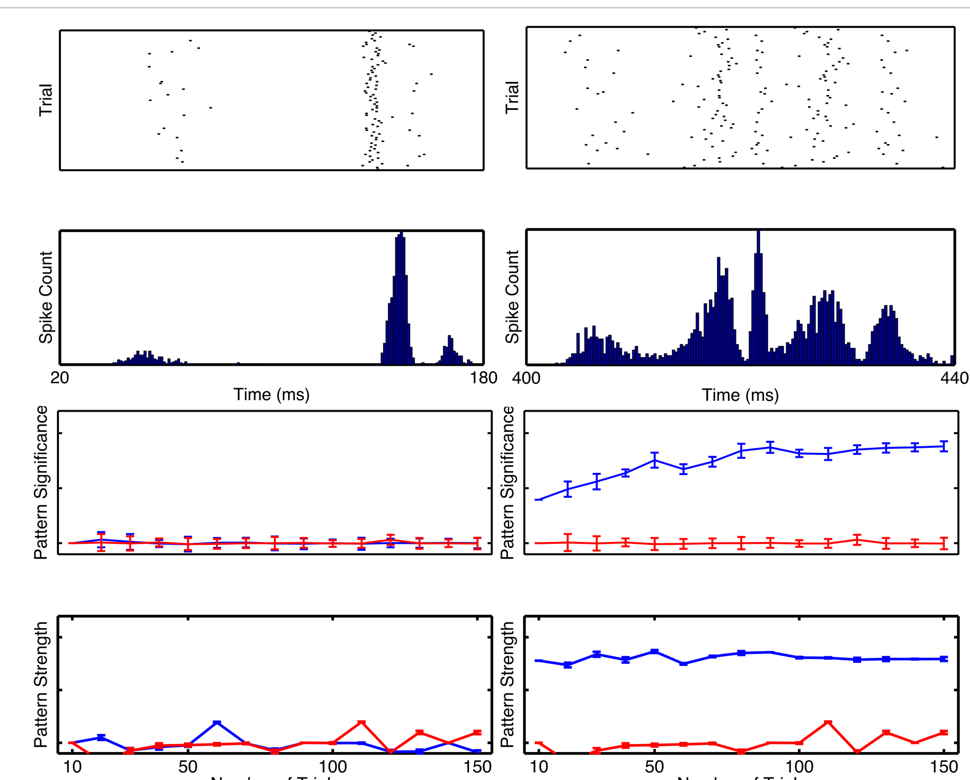

Figure 4: Pattern significance and strength distinguish between segments with real patterns and segments with fake patterns. On the left, the clustering has found two patterns, but there is only one pattern with unreliability in the first and second spikes. The pattern significance and strength of the fake pattern (blue) are statistically equal to those for a Poisson process with the same event structure (red). By contrast, the right set contains a real pattern, where the cluster strength and significance (red) are much higher than for a Poisson surrogate (blue). Overclustering is necessary to determine the event structure to sufficient accuracy for use in the validation procedure.

Evaluator:	Use:	Description:
Cluster Strength	Detects non-Poisson statistics in either a dataset or a cluster	The number of words in a given set of trials divided by the average number of words found in a Poisson re-sampling of the same data.
Cluster Significance	Detects non-Poisson Statistics in either a dataset or a cluster	The log-likelihood of drawing the given dataset minus the average log-likelihood of drawing the Poisson re-sampling of the same data.
Link Weakness	Detects if two clusters are merely a single, Poisson cluster	The ratio of the number of unique words found by shuffling spikes between two clusters linked by occupancy in at least one event and the sum of the number of words found by shuffling spikes within those clusters.

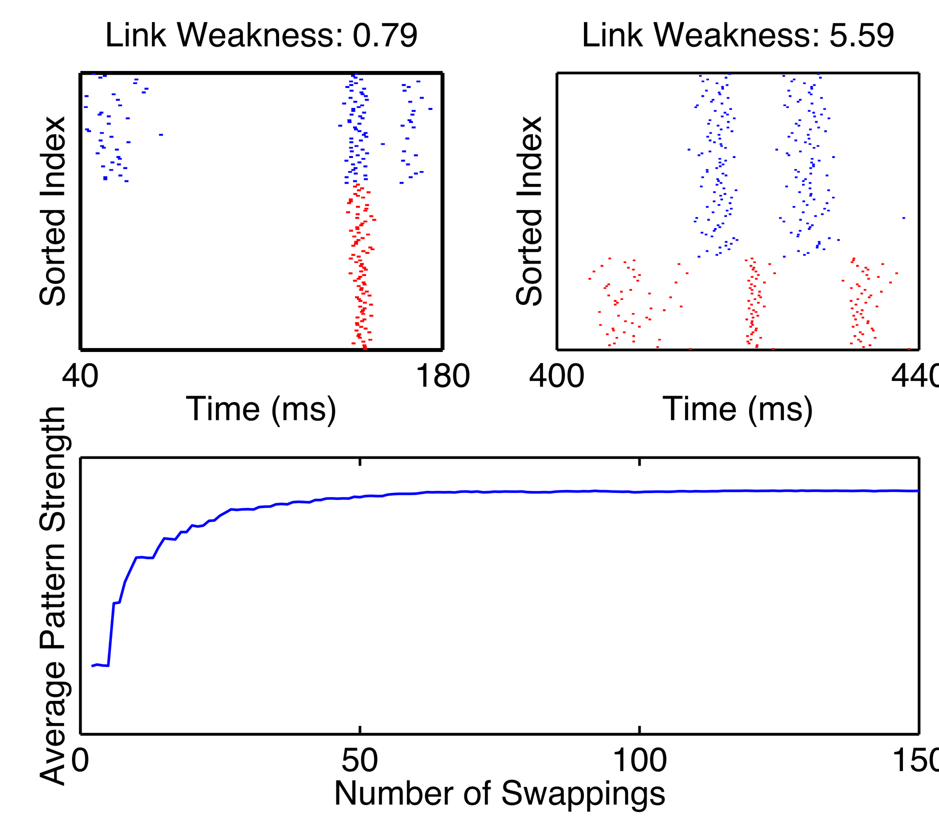

Figure 5: For the segment with the fake pattern, the Link Weakness indicates an over-clustering. The increase of within-cluster pattern strength as a function the number of swapped trials indicates that there are in fact two independent clusters.

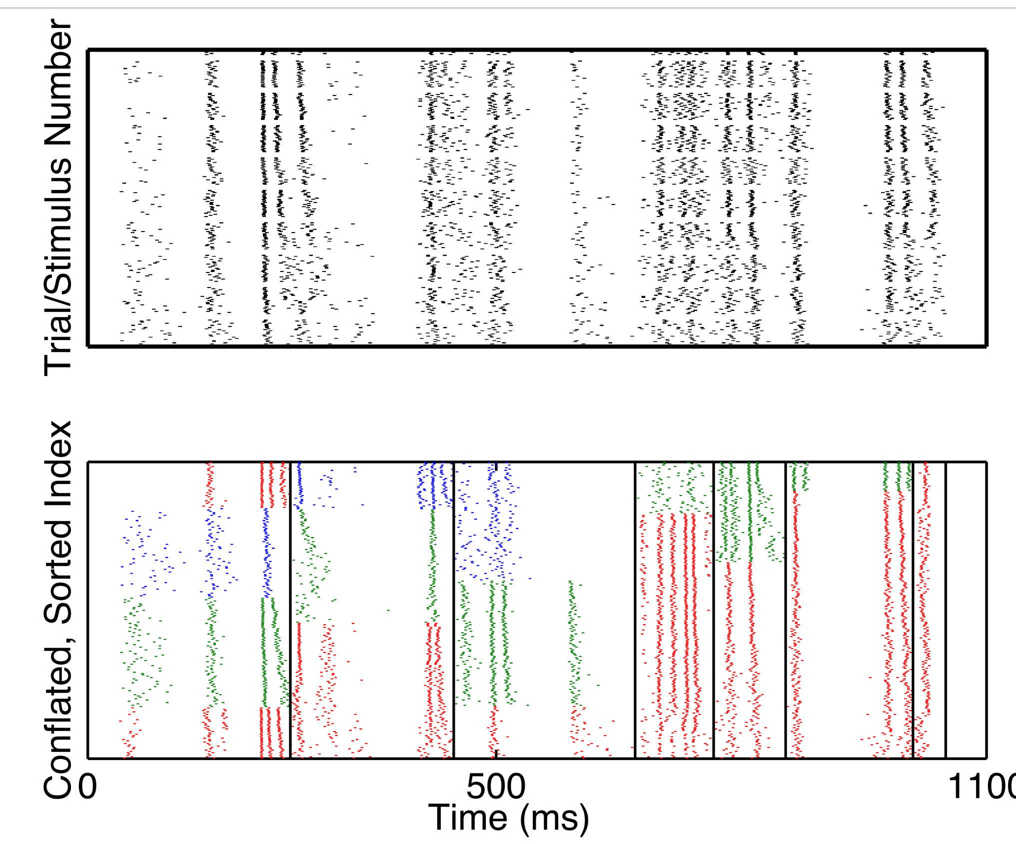

Figure 6: Spike patterns are conserved across different neurons driven by the same stimulus waveform. Patterns can be extracted using the conflated clustering method. Multiple patterns are obtained across trials and neurons. The differences in neuronal properties were modeled by changing the gain of the stimulus waveform.

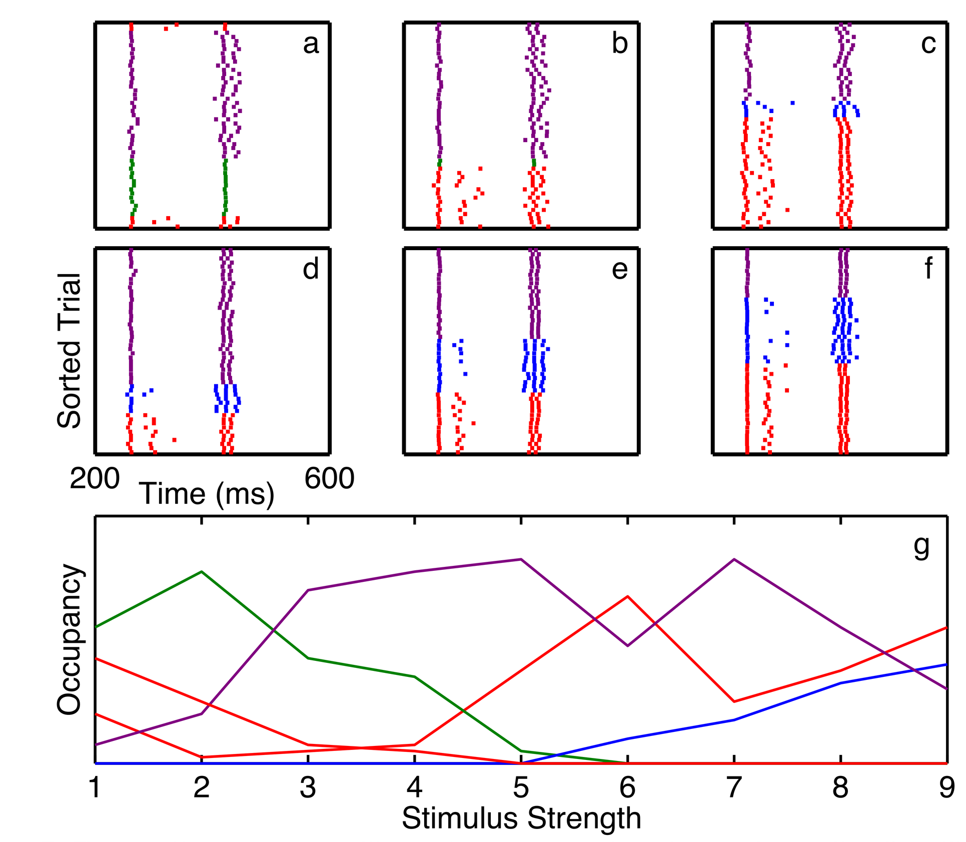

Figure 7: The pattern composition of a segment varies systematically with the gain of the stimulus waveform.

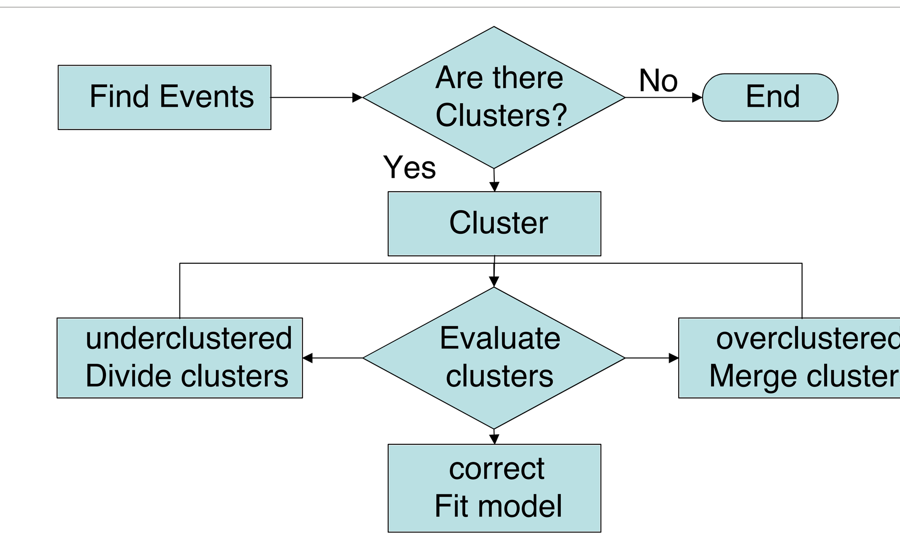

Figure 8: These techniques provide a systematic framework for determining whether patterns are present in a segment and how many there are.

Conclusions:

➲Fuzzy clustering on the Victor-Purpura distances makes it possible to extract patterns from spike-train data.

➲The presence and number of patterns can be assessed using the significance and strength.

➲This method can also be used on datasets which combine different experiments.

➲This framework is being extended for use with multi-unit data.

Work supported by NIH (R01 Mh06848), we thank Terry Sejnowski and Peter Thomas for support and discussion.

References:

1.Fellous, J.-M., et al (2004). Journal of Neuroscience *24, 2989-3001*.
2.Tiesinga, P. H. E. (2004). Physical Review E *69, 031912*.
3.Tiesinga, P. H. E., et al (2002a). Neural Computation *14*, 1629-1650.
4.Victor J.D., Purpura K.P. (1996) J Neurophysiol 76,1310-26.