Spectral Clustering of Synchronous Spike Trains

António R. C. Paiva, Sudhir Rao, Il Park and José C. Príncipe

Computational NeuroEngineering Laboratory (CNEL),
Department of Electrical and Computer Engineering,
University of Florida, USA

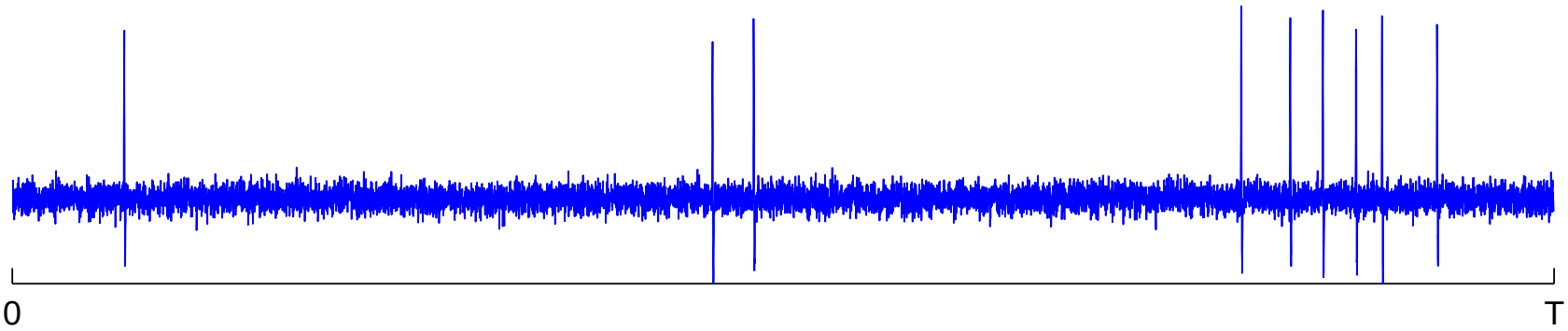
IJCNN 2007, August 2007

A.R.C. Paiva supported by FCT grant SFRH/BD/18217/2004.
Supported in part by NSF grants ECS-0601271, ECS-0422718 and CISE-0541241.

Outline

Outline

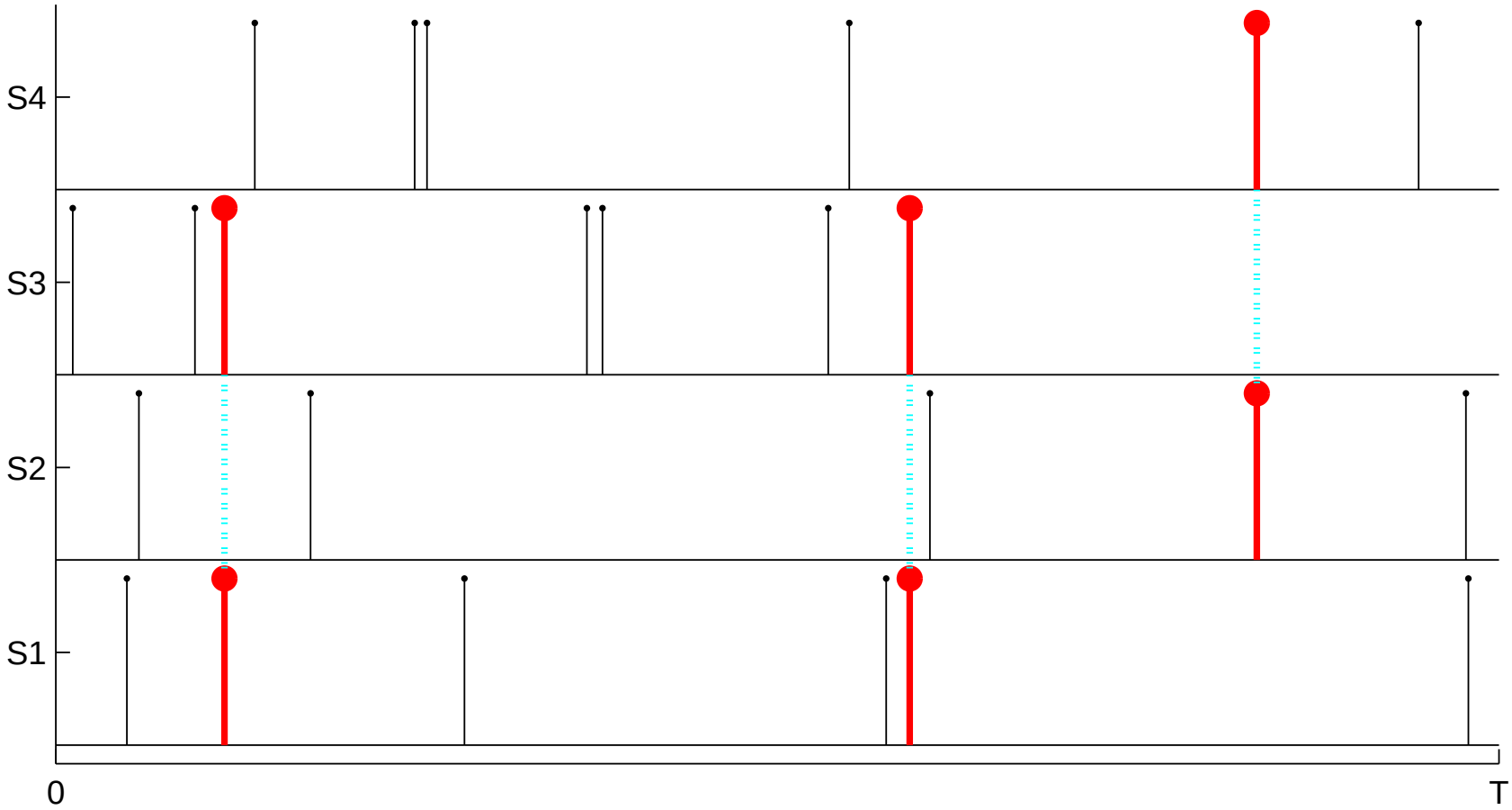
What are spike trains?



Definition

Spike trains are discrete representations of single-neuron activity as sets of spike times $\{t_m; m = 1, \ldots, N\}$. That is, the action potential shape and amplitude in the recordings is disregarded and only the instant it occurs is considered.

Clustering of synchronous spike trains?



Motivation

Clustering as an analysis tool

- In neuroscience:
  - Clustering “is” unsupervised classification.
    - Estimation of information.
  - Clustering groups neurons with “similar” spike trains allows to study:
    - Functional connectivity in the brain.
    - Organization in *neural assemblies*.
- In engineering:
  - Brain-Machine Interfaces (BMI)
  - Liquid-State Machine (LSM) computation
  - Spiking neural networks (SNN)
  - ...

Previous approaches

- Eggermont (2006) used the cross-correlation coefficient as a measure, and applied hierarchical clustering.
  - Computation of the cross-correlation coefficient assumes binning, which imposes time quantization.
- Fellous et al. (2004) utilized a “correlation-based measure” followed by fuzzy k-means.
  - Evaluation of the measure is computation intensive.
  - Dimensionality of clustering increases with the number of spike trains.
  - Determines one cluster at a time.

Outline

Binless distances between two spike trains

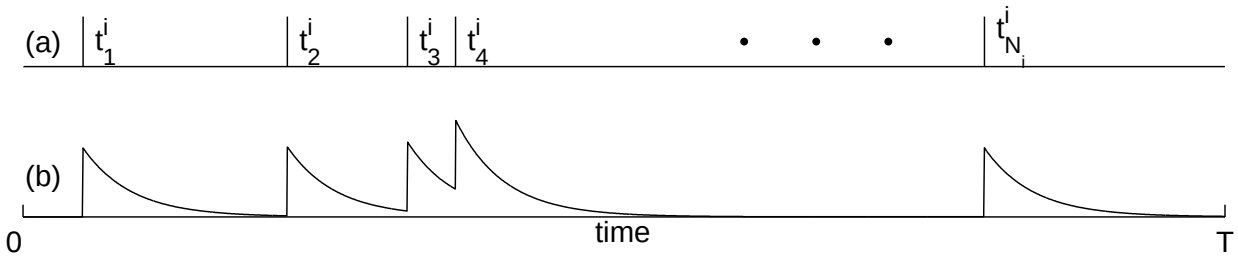
- Victor-Purpura's distance – $D^{spike}[q]$ (1995, 1997, 2005)

$$d_{ij} = \min_{\substack{\text{all spike pair} \\ \text{combinations, } k}} \sum_{k} \text{cost}(t^i_{c_i[k]} \rightarrow t^j_{c_j[k]})$$

  - Non-euclidean distance

- van Rossum's distance (2001)
  - Euclidean distance
  - Conceptually and computationally simple

van Rossum's distance (in more detail...)



- With filter impulse response $h(t) = \exp(-t/\tau)u(t)$, denote the i th filtered spike train as

$$f_i(t) = \sum_{m=1}^{N_i} h(t - t_m^i).$$

- van Rossum's distance is defined as

$$d_{ij} = \frac{1}{\tau} \int_0^\infty [f_i(t) - f_j(t)]^2 \, dt$$

Fast computation of van Rossum's distance

- Substituting the filter impulse response and the expression for the filtered spike trains into the definition of the distance yields

$$d_{ij} = \frac{1}{2}\left[\sum_{m=1}^{N_i}\sum_{n=1}^{N_i} L_\tau(t_m^i - t_n^i) + \sum_{m=1}^{N_j}\sum_{n=1}^{N_j} L_\tau(t_m^j - t_n^j)\right] + \sum_{m=1}^{N_i}\sum_{n=1}^{N_j} L_\tau(t_m^i - t_n^j),$$

with $L_\tau(\cdot) = \exp(-|\cdot|/\tau)$ is the Laplacian function.

Outline

Which clustering algorithm?

- Propose to use spectral clustering.
  - Clear distinction between measure similarity of data points and actual clustering procedure.
  - Good behavior with nonlinearly separable clusters.
  - It has a close relationship with information theoretic methods.
- The algorithm proposed by Ng et al. (2001) was used.

Affinity matrix

- Consider n spike trains $s_1, s_2, \ldots, s_n$.
- The affinity matrix A is an $n \times n$ matrix quantifying the *similarity* between any two spike trains.
- The ij th entry of the affinity matrix is

$$a_{ij} = \begin{cases} \exp\left(-\dfrac{d_{ij}^2}{2\sigma^2}\right), & \text{if } i \neq j \\ 0, & \text{otherwise} \end{cases}$$

Overview of the clustering algorithm

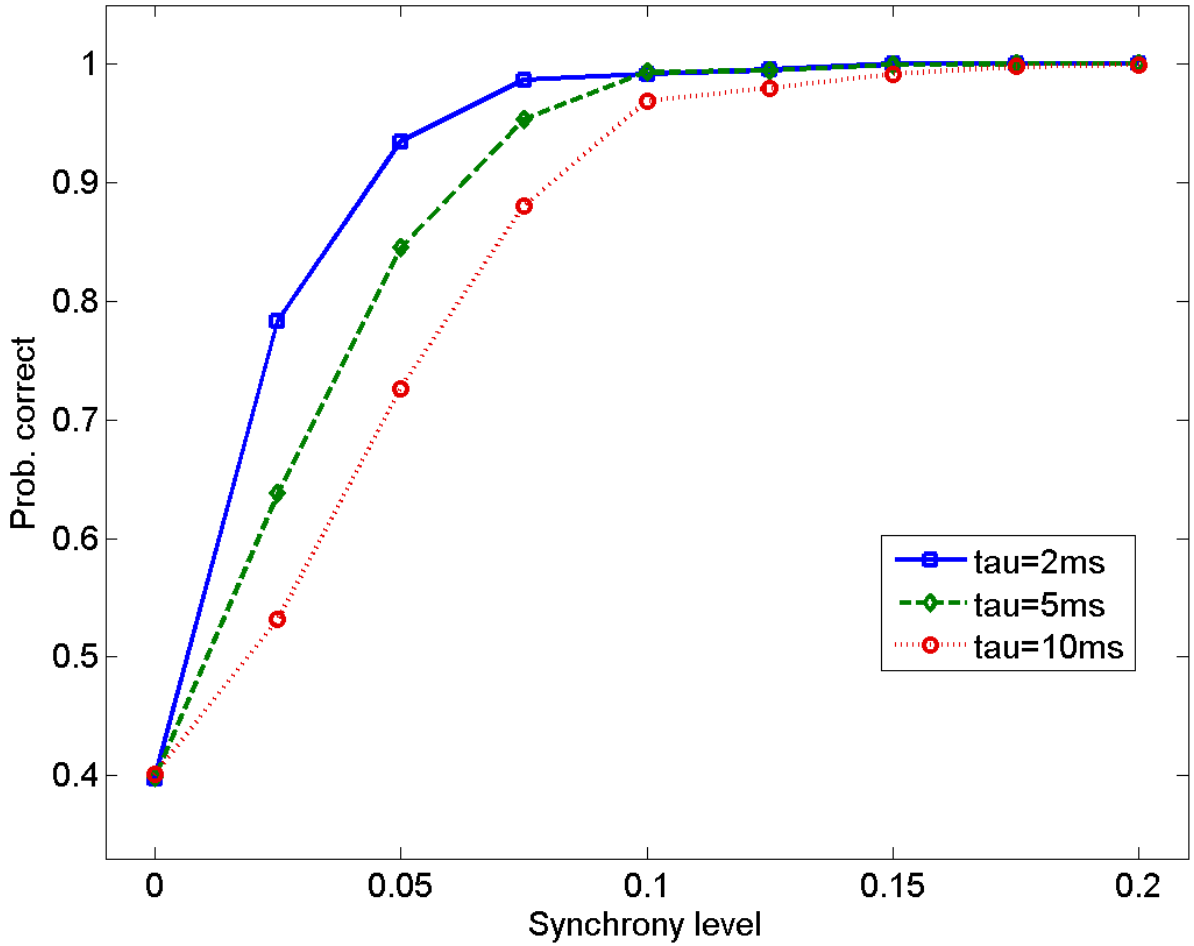
1. Compute the distance matrix of all spike train pairs.
2. Evaluate the affinity matrix.
3. Apply spectral clustering.

Outline

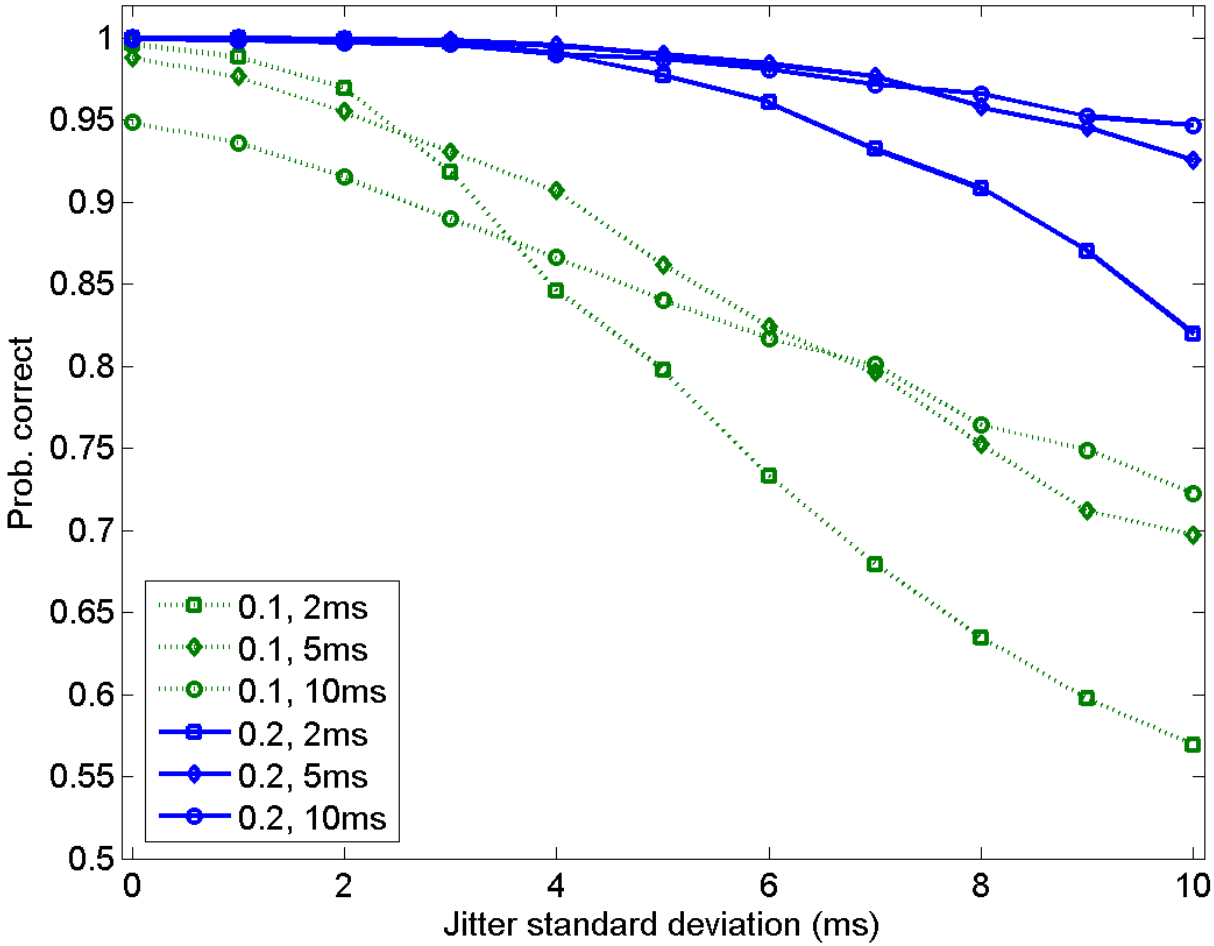
Simulation paradigm

- Spike trains were modeled as homogeneous Poisson processes, with firing rate 20 spk/s, and where generated according to a Multiple Interaction Process (see Kuhn et al., 2001) with synchrony parameter ε.
- For set of parameters, we generated 100 spike trains distributed through 3 clusters.
- Results were averaged over 100 and 500 Monte Carlo runs (without and with jitter, respectively).
- Gaussian kernel size was fixed at $\sigma = 10$.

Clustering perfectly synchronized spike trains



Clustering spike trains in the presence of jitter



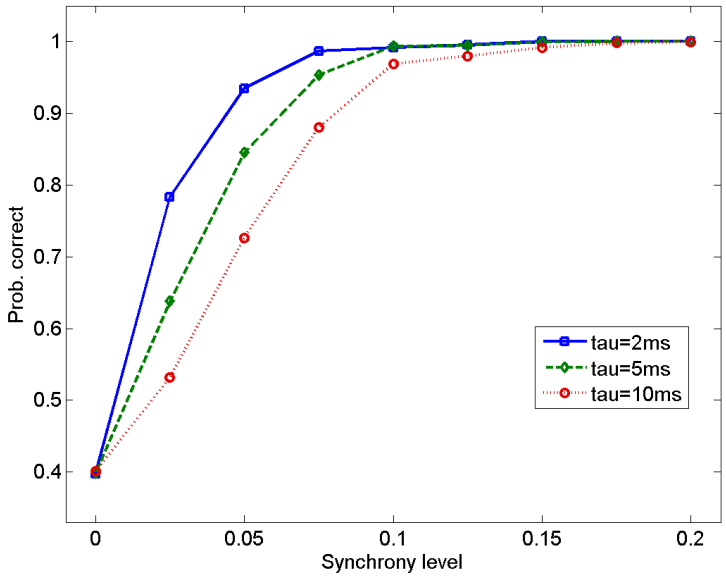
Summary

- Proposed simple method for clustering of synchronized spike trains.
- Presented computationally efficient method for evaluation of van Rossum's distance.
- Use of spectral clustering is a good choice.
  - Clear distinction between measure similarity of data points and actual clustering procedure.

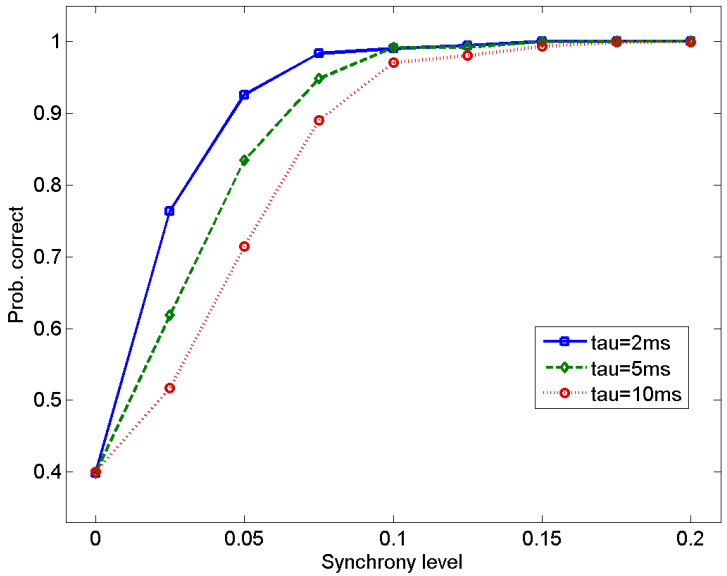
van Rossum vs. VP distance

Clustering perfectly synchronized spike trains

van Rossum's distance



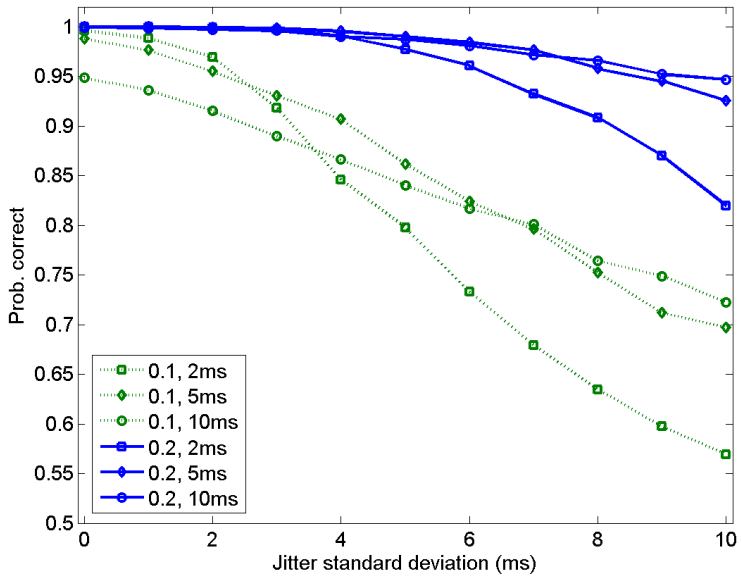
VP distance



van Rossum vs. VP distance

Clustering spike trains in the presence of jitter

van Rossum's distance



VP distance