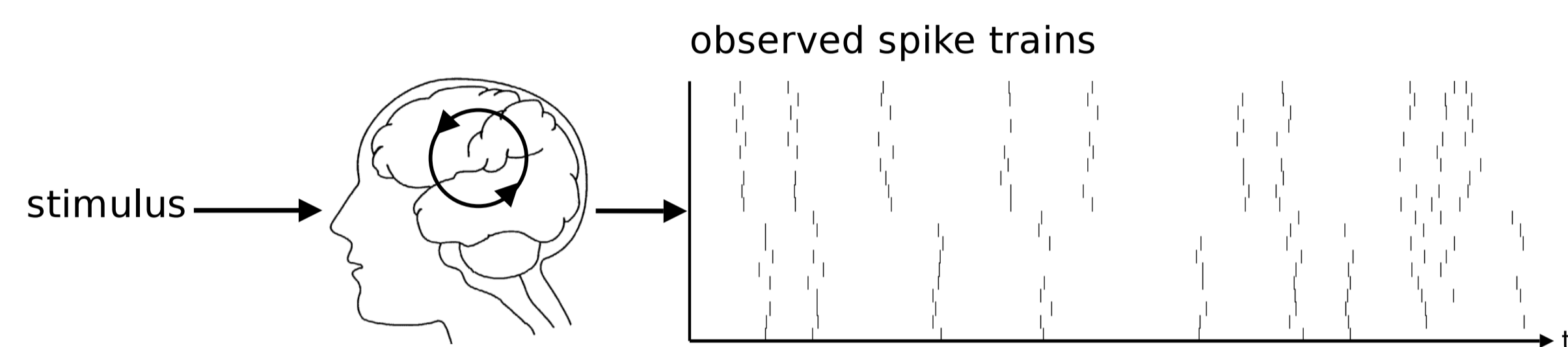


Which measure should we use for unsupervised spike train learning?

Motivation

In certain experimental paradigms, the dynamics of a neural system may not be fully determined by external stimuli because the neural activity depends on internal states from a wide range of possible causes. For example, bistable dynamics of a single neuron has been observed in vitro via frozen noise injection [1], local field potentials and EEG phase often correlates with response strength, and top-down control such as attention are known to affect responses.



In other words, in these cases, it is nearly impossible to control or observe all the internal variables. Still, we would like to infer these internal states by analyzing the observation variability.

The solution we propose is to use unsupervised learning methods, such as PCA and clustering [2, 3], to discover the internal states.

Spike train measures

Consider two spike trains $s_i, s_j \in \mathcal{S}(\mathcal{T})$ defined in the interval $\mathcal{T} = [0, T]$. We compare three inner product measures:

- The memoryless cross-intensity (mCI) inner product is defined as

$$I(s_i, s_j) = \int_{\mathcal{T}} \lambda_{s_i}(t) \lambda_{s_j}(t) dt.$$

This is the simplest of a family of cross-intensity inner products and is equivalent to a continuous-time cross-correlogram. It is also the inner product associated with van Rossum's distance.

- The inner product associated with the b-metric proposed by Houghton [4], henceforth denoted IPH, is defined as

$$H_{\mu}(s_i, s_j) = \int_{\mathcal{T}} f_{s_i}(t) f_{s_j}(t) dt,$$

where f characterizes the dynamics of synaptic model governed by equations:

$$\begin{aligned} \text{spike discontinuities : } f &\rightarrow (1 - \mu)f + 1 \\ \text{neurotransmitters unbinding : } \tau \frac{df}{dt} &= -f \end{aligned}$$

- The nonlinear cross-intensity (nCI) inner product is defined as

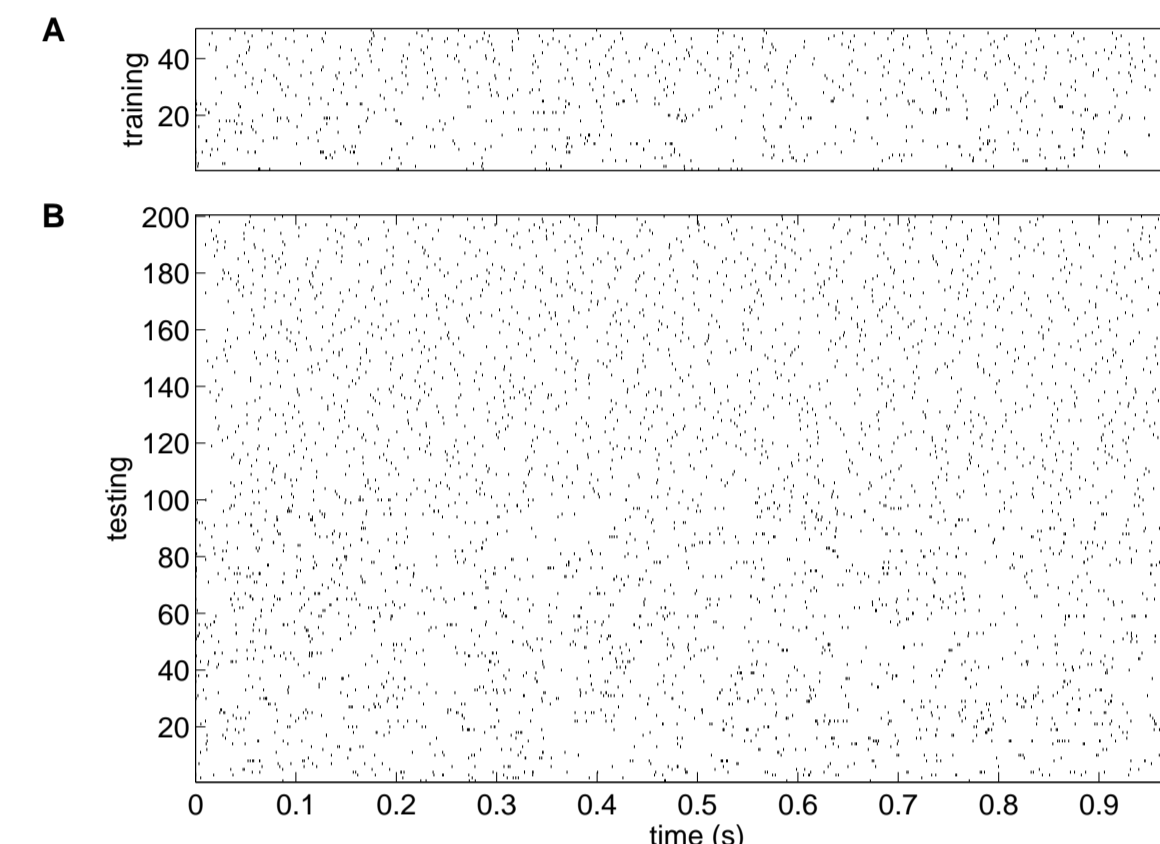
$$\mathcal{I}_{\sigma}(s_i, s_j) = \int_{\mathcal{T}} \mathcal{K}_{\sigma}(\lambda_{s_i}(t), \lambda_{s_j}(t)) dt,$$

where \mathcal{K}_{σ} is a symmetric positive definite kernel with parameter σ .

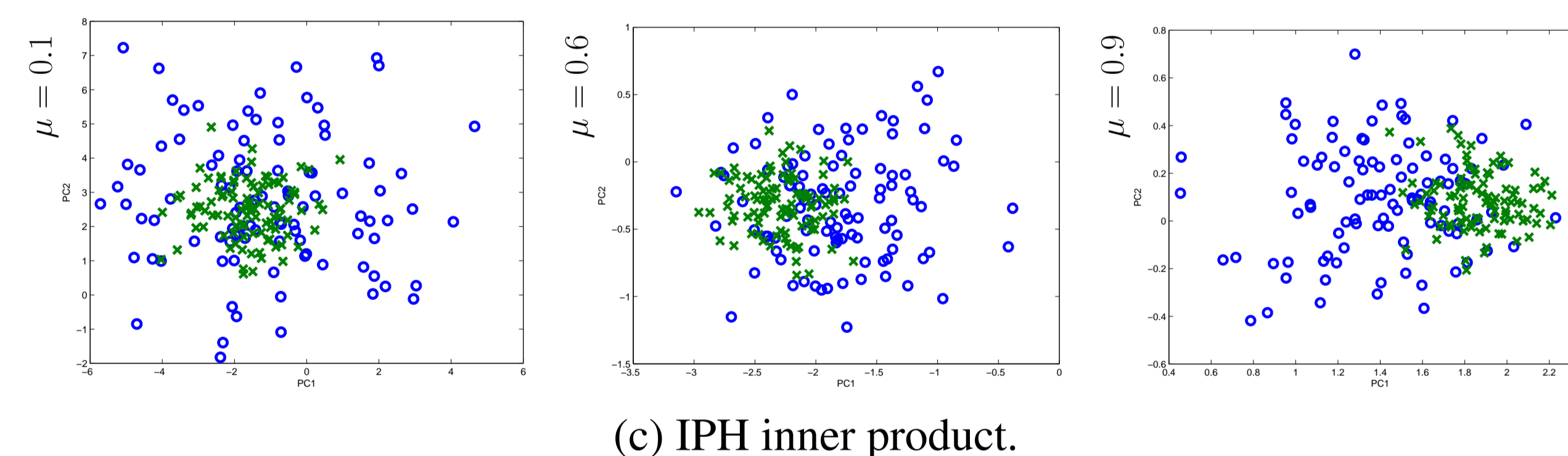
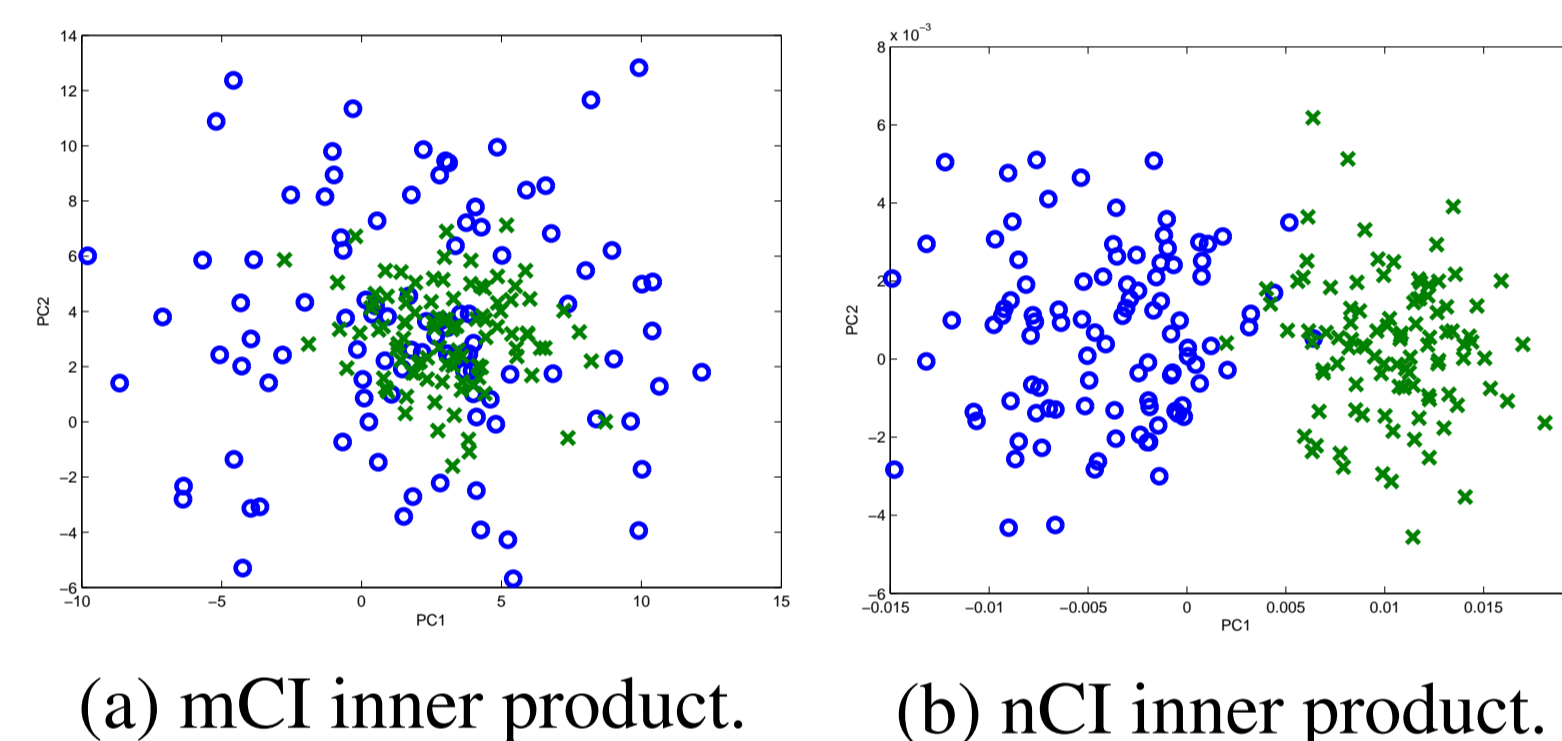
Empirical comparison

Principal component analysis of renewal point processes

The measures were used to compute principal component analysis (PCA) of spike trains generated from two homogenous renewal point processes with gamma distributed inter-spike intervals (shape parameter $\theta = 0.5$ and $\theta = 3$). The algorithm follows the same steps as Kernel PCA using one of the spike train inner product measures. The smoothing parameter was $\tau = 50$ ms.



Using the nCI and IPH (with $\mu = 0.9$) inner products the spike train projections clearly distinguish the different processes, revealing that these inner products are sensitive to the time history needed to discriminate between renewal point processes.



Clustering of spike trains from an in-vitro neural culture

The measures were compared on a clustering task [2] using multi-channel spike trains recorded from an in-vitro neural culture that was electrically stimulated at ten different channels in a random order.

Smoothing parameter	Inner product measure		
	mCI	IPH	nCI
1ms	60.5%	71.2%	99.3%
3ms	56.1%	59.8%	98.6%
5ms	45.0%	47.8%	97.4%

The performance was evaluated in terms of the number of spike trains correctly grouped with regards to stimuli. For the IPH and nCI inner products, the best results are shown, corresponding to using $\mu = 0.8$ and $\sigma = [0.1, 10]$. Note that the performance using the nCI was the same across a large range of the nonlinearity parameter σ , and much better than using the other two inner products.

Why does the nCI perform better?

Theorem 1. Given a strictly positive definite function $\mathcal{K}_{\sigma} : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$, the nCI kernel $\mathcal{I} : \mathcal{S}(\mathcal{T}) \times \mathcal{S}(\mathcal{T}) \rightarrow \mathbb{R}$ is strictly positive definite, and therefore a characteristic kernel.

Proof. The theorem follows immediately from noticing that

$$\sum_i \sum_j a_i a_j \mathcal{I}_{\sigma}(s_i, s_j) = \int \left[\sum_i \sum_j a_i a_j \mathcal{K}_{\sigma}(\lambda_{s_i}(t), \lambda_{s_j}(t)) \right] dt \geq 0,$$

and, because \mathcal{K}_{σ} is strictly positive definite, the equality holds if and only if $\lambda_{s_i}(t) = \lambda_{s_j}(t) \Leftrightarrow s_i = s_j$. \square

Since the nCI is a characteristic kernel, the following theorem follows.

Theorem 2. The nCI-based measure between two point processes P and Q ,

$$\mathcal{D}_K(P, Q) = \iint \mathcal{I}(x, y) d\mu(x) d\mu(y),$$

with probability measure $\mu = P - Q$, is a divergence.

This means that two point processes map to different points in the reproducing kernel Hilbert space (RKHS) induced by the kernel. In other words, this kernel ensures that spike trains corresponding to different point processes have distinct means in the RKHS, which allows statistical inference to distinguish between them.

Conclusion

We show that the nCI inner product is the most powerful measure. This was shown here explicitly for unsupervised learning, but similar results have been obtained for supervised learning. This is because the nCI is a characteristic kernel, and, intuitively, sensitive to memory dependences of the point processes.

The disadvantage of the nCI is that the explicit form of a congruent inner product in the spike train or intensity function space is unknown. This prevents the results from being related back to those spaces, where the analysis can be more easily understood in terms of synaptic and statistical models.

References

- [1] J.-M. Fellous, P. H. E. Tiesinga, P. J. Thomas, and T. J. Sejnowski Discovering Spike Patterns in Neuronal Responses J. of Neuroscience, 24(12):2989-3001, March 2004.
- [2] A. R. C. Paiva, I. Park, and J. C. Principe A reproducing kernel Hilbert space framework for spike train signal processing Neural Comp., 21(2):424449, Feb. 2009.
- [3] A. R. C. Paiva, I. Park, and J. C. Principe Inner products for representation and learning in the spike train domain In Karim G. Oweiss, editor, Statistical Signal Processing for Neuroscience, Academic Press, 2010. In-press.
- [4] C. Houghton Studying spike trains using a van Rossum metric with a synapse-like filter J. Comp. Neurosci., 26:149155, 2009.