

A metric space approach to the information capacity of spike trains.

James Gillespie and Conor Houghton, Trinity College Dublin

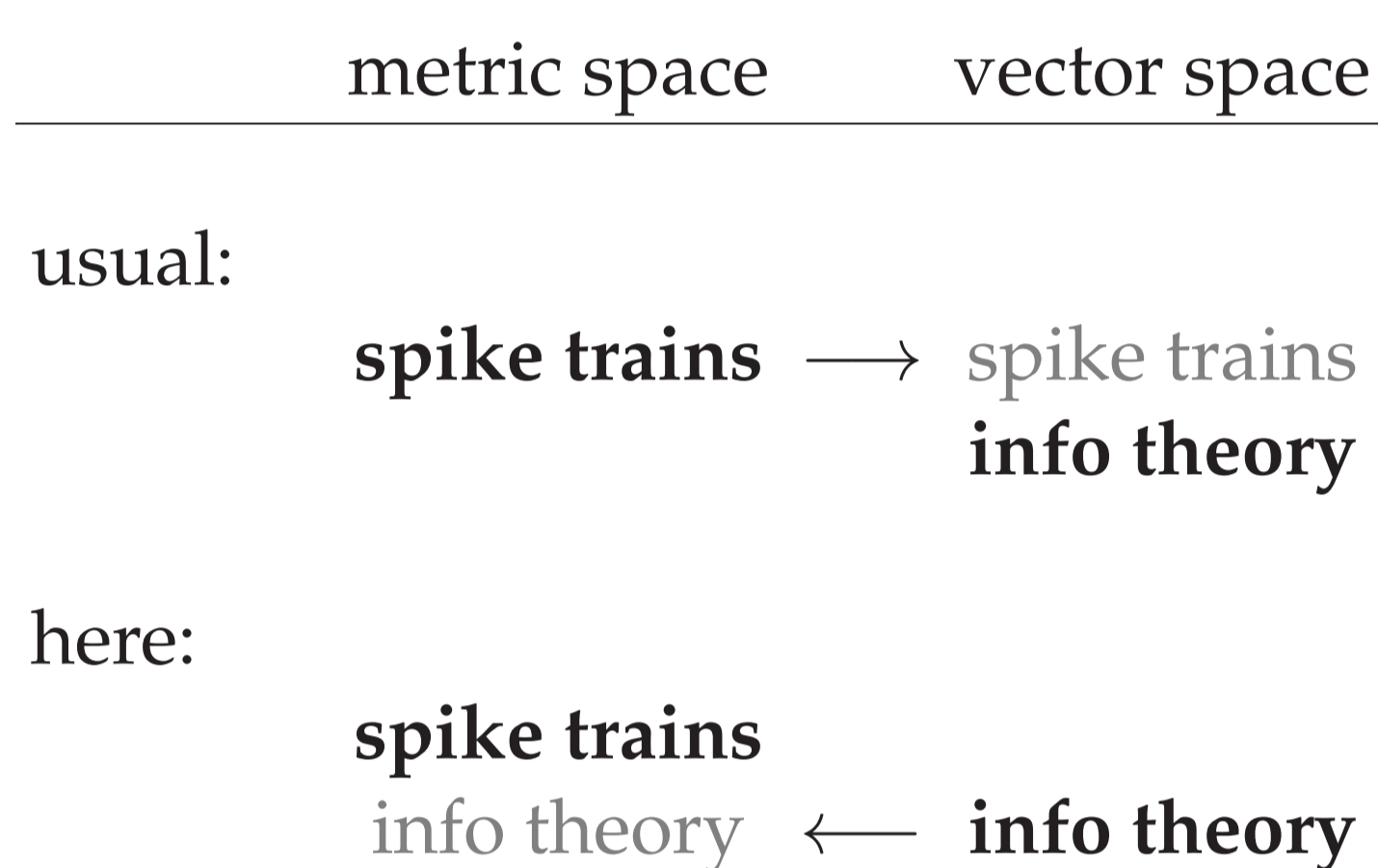
gillesjb@tcd.ie and houghton@maths.tcd.ie

Abstract

- Classical information theory can be either discrete or continuous, corresponding to discrete or continuous random variables. However, although spike times in a spike train are described by continuous variables, the information content is usually calculated using discrete information theory. This is because the number of spikes, and hence, the number of variables, varies from spike train to spike train, making the continuous theory difficult to apply.
- It is possible to avoid this problem by using a metric space approach to spike trains. A metric gives a distance between different spike trains. The continuous version of information theory is then rephrased in terms of metric quantities and used to estimate the information capacity of spike trains.
- This method works by matching the distribution of distances between responses to the same stimulus to a χ -distribution: the χ -distribution is the length distribution for a vector of Gaussian variables. This defines a noise dimension for the spike train and gives a bound on the channel capacity.

Motivation

Typically the spike train is transformed in a quite unnatural way to make information theory applicable; here it is information theory which is adapted to the metric space of spike trains.



The space of spike trains

The geometry of the space of spike trains is derived from a distance function defined on the spike trains. In the **van Rossum metric** [1] the spike train, considered as a list of spike times, $\mathbf{t} = (t_1, t_2, \dots, t_n)$, is mapped to a real function, $f(t; \mathbf{t})$ using a kernel $h(t)$:

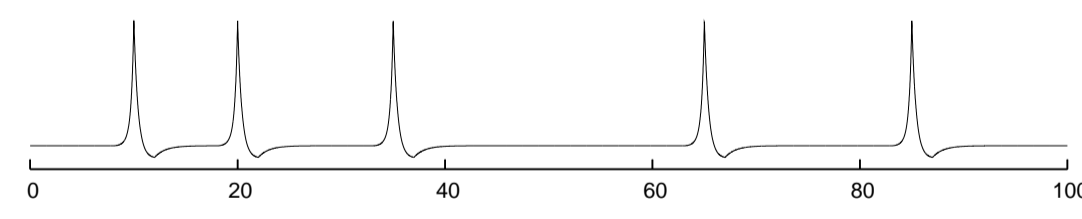
$$\mathbf{t} \mapsto f(t; \mathbf{t}) = \sum_{i=1}^n h(t - t_i).$$

The distance between two spike trains is taken to be the distance between the two corresponding functions, using the standard L^2 metric on the space of real functions. One common choice of kernel is the causal decaying exponential

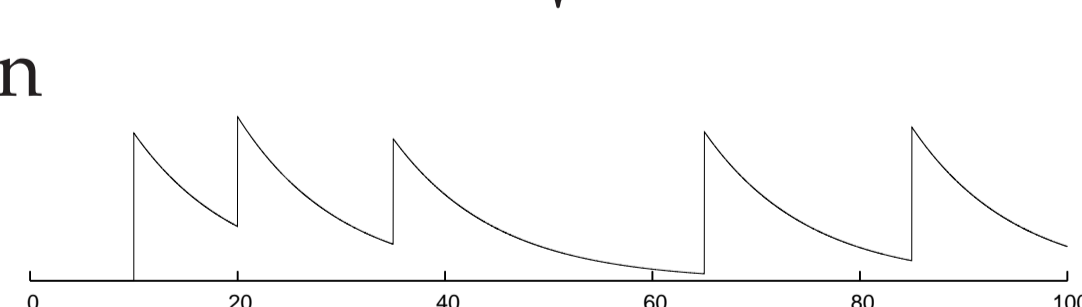
$$h(t) = \begin{cases} 0 & t < 0 \\ e^{-t/\tau} & t \geq 0 \end{cases}$$

where τ is a time-scale which parameterizes the metric.

Spikes



Function



For a review see [2].

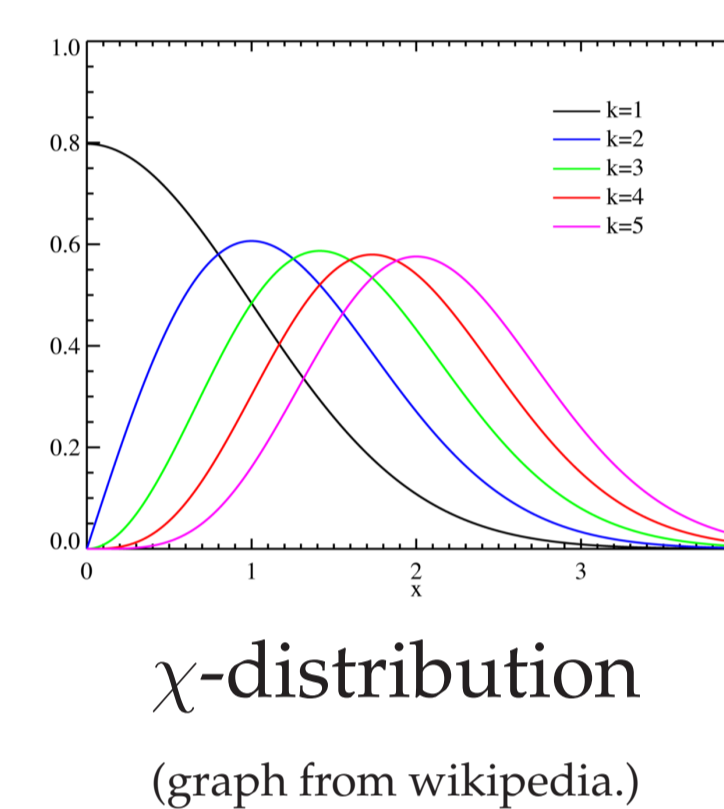
Noise and distances

The distance between points with Gaussian distributed coordinates have a χ -distribution. Consider a vector

$$\mathbf{x} = \mathbf{X} + \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_k \end{pmatrix} \sim \begin{matrix} N(0, \sigma) \\ N(0, \sigma) \\ \vdots \\ N(0, \sigma) \end{matrix}$$

where k is the dimension, \mathbf{X} is a constant standing for the signal and $\xi_i \sim N(0, \sigma)$ denotes Gaussian distributed noise. The distance between two such vectors is χ -distributed:

$$|\mathbf{x}_1 - \mathbf{x}_2| \sim \chi(k, \sqrt{2}\sigma)$$



Noise on the space of spike trains

In the case of spike trains there is no vector space, there is no equivalent of the \mathbf{x} above. However, the space of spike trains is a metric space, so there is an equivalent of $|\mathbf{x}_1 - \mathbf{x}_2|$.

Proposal

- Model noise in the metric space of spike trains using the χ -distribution.
- Apply channel capacity theory on the metric space to estimate information.

Noise dimension

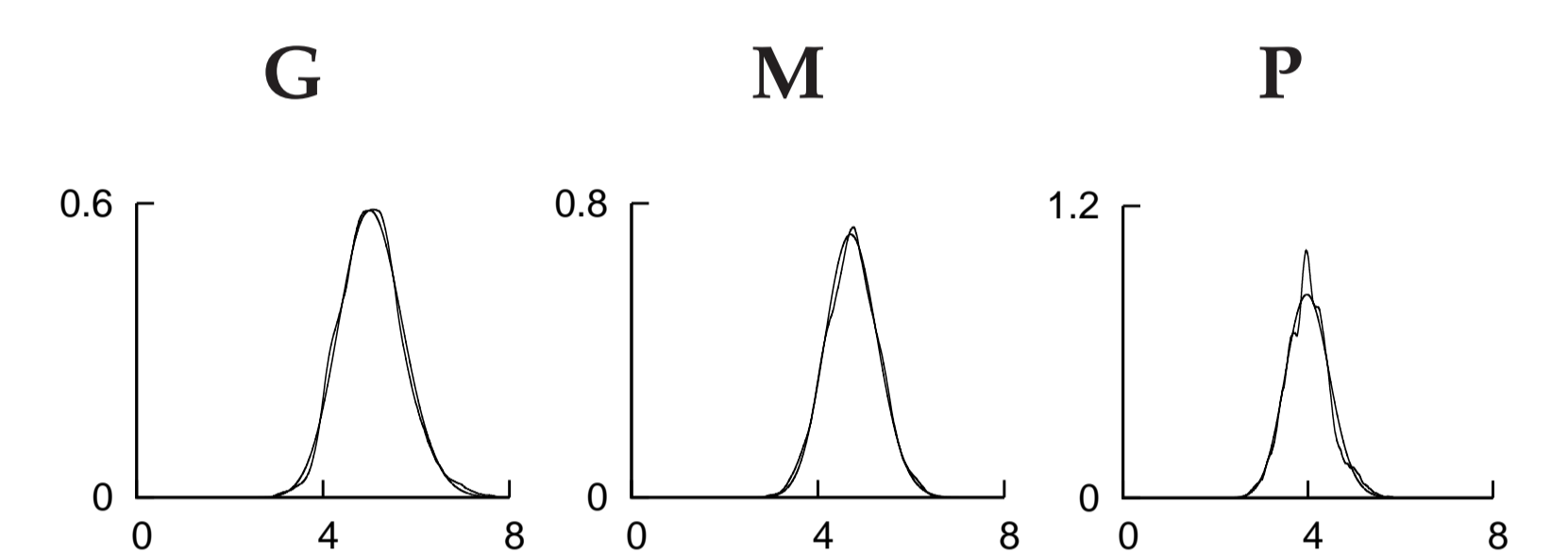
One interesting consequence is that the noise defines a dimension per length for spike trains corresponding to k . This is, in general, non-integer. It is the average dimension of the Gaussian vector which would produce the same χ -distribution.

References

- [1] van Rossum M. *Neural Computation*, 13:751–763, 2001.
- [2] Houghton CJ and Victor JD. Measuring representational distances – the spike-train metrics approach. (Invited book chapter, in press, 2009).
- [3] Narayan R, Graña G, Sen K. *Journal of Neurophysiology*, 96:252–258, 2006.
- [4] Gillespie J and Houghton CJ. A metric space approach to the information capacity of spike trains. (Preprint tcd-mnl-14, 2010).

Results - noise

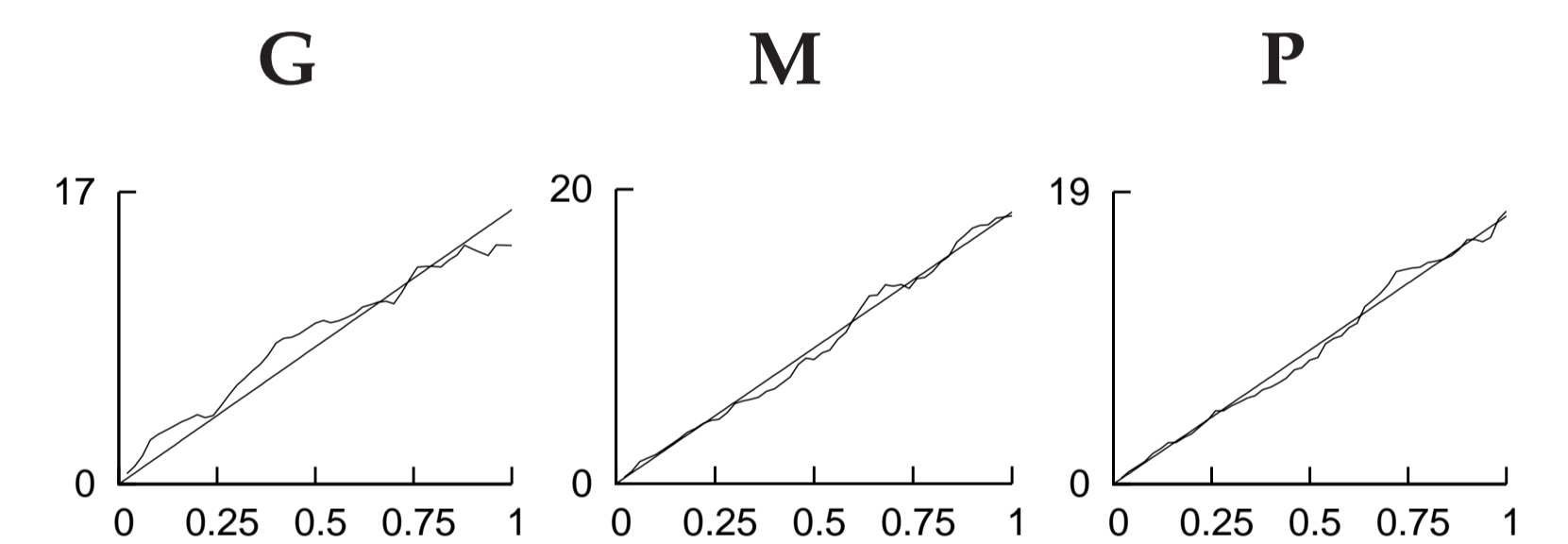
These ideas were applied to the in vivo spike trains described below; the distance distribution of noise was calculated by comparing spiking responses to the same sensory stimulus. This distribution was modelled using the χ -distribution:



A comparison between the distance distributions for the noise responses and the corresponding χ -distributions. Kernel density estimation is used to generate the probability density distribution of the distances, for an usually good site G, a typical site M and an unusually poor site P. In each case, this is compared to the χ -distribution, where the parameters of the distribution are estimated from the moments. The **Anderson-Darling test** was used to evaluate the hypothesis that the noise distribution follows a χ -distribution: general the measured statistic lies within half a sigma or closer to the statistic for the corresponding χ -distribution.

Results - noise dimension

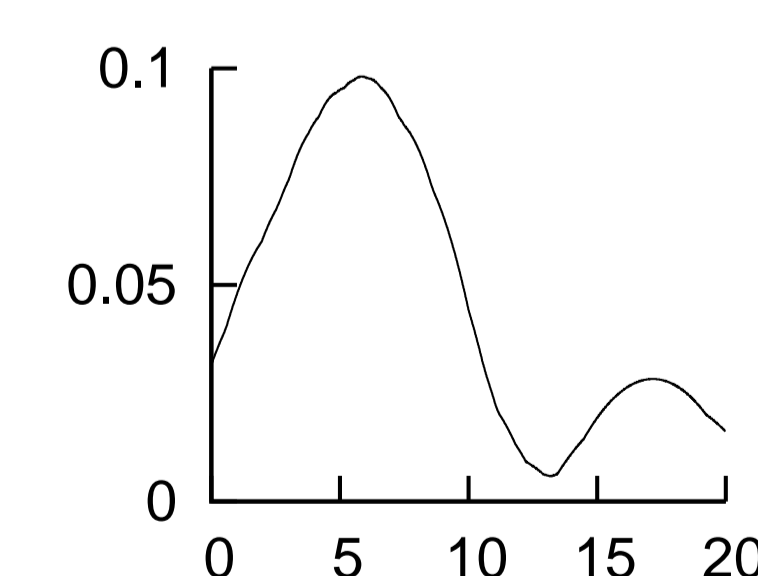
Since the noise dimension is a dimension per unit length, it should be linear:



Plots of $k(L)/2$. For G, M and P the moments of the distance distributions have been used to calculate $k(L)/2$ against fragment lengths, L , up to one second. The straight line represents the predicted least squares fit of the data.

Results - information capacity

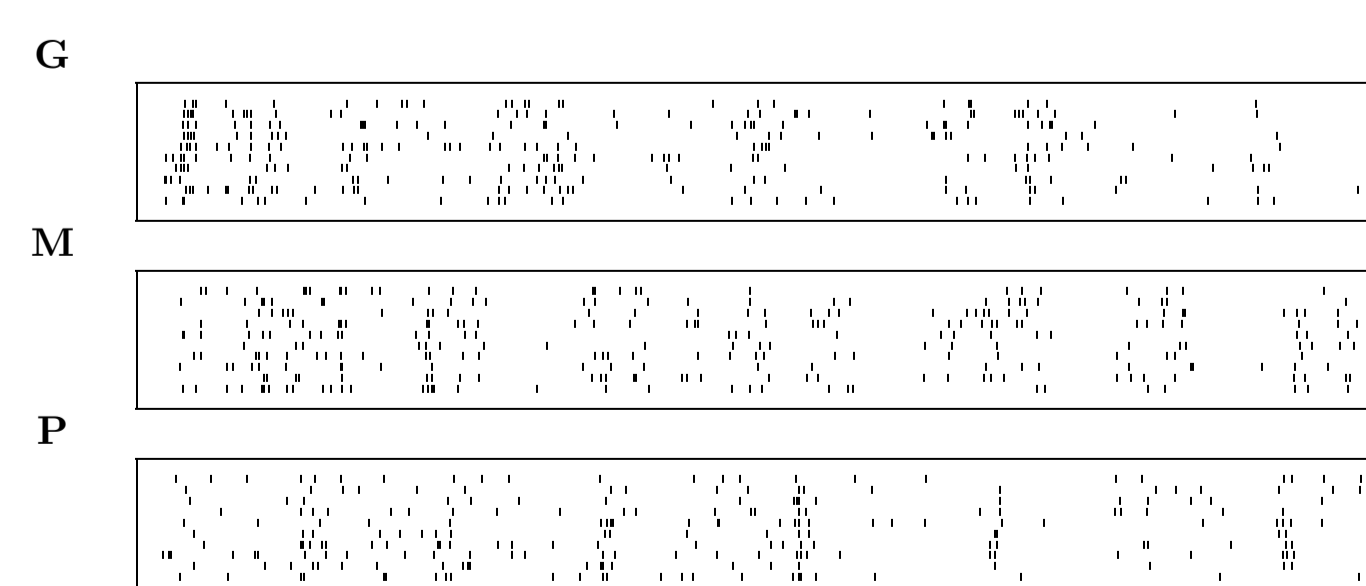
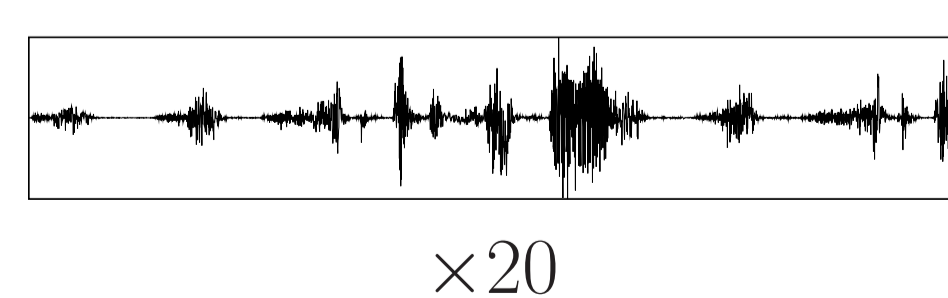
Distances are also calculated for responses to different stimuli; this gives a measure of the signal variance, using this along with the noise variance and the noise dimension gives an estimate of the channel capacity. Distribution of channel capacities in bits per second:



A plot of the probability density for capacity estimated using kernel density estimation on 24 sites. Kernel density estimation has been used to estimate a smooth distribution for the channel capacity.

Data

The results are here are for electrophysiological data recorded from the primary auditory area of zebra finch during playback of conspecific songs [3].



Ten responses are recorded to each song; responses for three different cells are shown here.

The recordings were taken from field L of anesthetized adult male zebra finch and data was collected from sites which showed enhanced activity during song playback. In the ascending auditory pathway, area field L is afferent to the song system and is considered the oscine analogue of the primary auditory cortex. 24 sites are considered here; of these, six are classified as single-unit sites and the rest as consisting of two to five units. The average spike rate during song playback is 15.1 Hz with a range of across sites of 10.5-33 Hz.