# Measures of spike train synchrony: From single neurons to populations

Conor Houghton, Thomas Kreuz

September 29, 2013

**Abstract**

With the increasing availability of multi-unit recordings the focus of attention is shifting from bivariate estimators of spike train synchrony towards methods that describe patterns of activity across many neurons. Such measures of population spike train synchrony are becoming indispensable tools for addressing issues such as spike timing reliability, network synchronization, and neuronal coding. This chapter will give an overview of different approaches designed to quantify multiple neuron synchrony. It will address both measures of synchrony among one group of neurons as well as measures that estimate the degree of synchronization between populations of neurons.

## 1 Introduction

A wide variety of approaches to quantifying the dissimilarity, or distance, between two spike trains has been suggested. Among these is the edit-distance metric introduced in [21], which evaluates the total 'cost' needed to transform one spike train into the other, using only certain elementary steps each with an individual cost. Another metric proposed in [20] first maps the spike trains to functions by convolving the spikes with an exponential function and then measures the Euclidean distances between the functions. Both methods involve one parameter that sets the time scale.

Recently, the ISI-distance [10, 11] and the SPIKE-distance [12, 14] have been proposed as parameter free and time scale adaptive alternatives. These new measures are complementary to the ones mentioned above; the van

Rossum metric and the ISI-distance quantifying dissimilarities in estimates of the neurons' local firing rate profiles whereas the Victor-Purpura metric and the SPIKE-distance track differences in spike times. The ISI- and the SPIKE-distance are defined using a time profile which means that they are useful for time-local monitoring of dissimilarity.

Like the Victor-Purpura metric and the van Rossum metric, the ISI-distance is known to be a metric [15]; however, we will refer to it as the ISI-distance since that has been common practise until now. Generally though, we will use 'distance' to mean a map from pairs of responses to a non-negative real number that is proposed for measuring dissimilarity and reserve the word 'metric' for distance measures that are metrics in the mathematical sense; that is, symmetric, non-degenerate distance measures satisfying the triangular inequality.

A distance measure is sensitive to the coding structure of spike trains if it measures short distances between responses to the same stimulus and longer distances between responses to different stimuli. Therefore, a distance measure can be evaluated by performing distance-based clustering and then calculating how accurately responses to the same stimulus are clustered together. In this chapter, the bivariate distance measures described above are compared using this approach. For the van Rossum metric we also illustrate a recently proposed trick that speeds up the computation considerably [7].

Advances in recording technology mean simultaneously recorded populations of neurons are increasingly common. In order to analyze these population recordings, the spike train distance measures have to be extended from the bivariate case to the population case. There are two sorts of population measures.

The first type quantifies how spread-out the spike trains in a population are. By summing pairwise bivariate distances the ISI- and the SPIKE-distance can be used to measure the dissimilarity within a population response: the time profile then gives a time-local indication of how the population is behaving, something that could be compared to the time course of a stimulus for example. A new, entropy based, measure of population dissimilarity is also examined.

The second type of population extension compares two different responses from a population of neurons, typically this would mean two responses from the same neurons corresponding to two different trials of a repeated sensory input. However, these extensions could be applied to any pair of responses composed of the same number of spike trains, for example, for comparing a

real set of responses to a simulated response from a model of the network, or for stereotypical neuronal networks like those found in insects, for responses from the same neurons in different animals.

Population extensions of this second type have been suggested for the Victor-Purpura [2] and the van Rossum metric [8]. They both introduce a second parameter that quantifies the importance of distinguishing spikes fired in different cells by interpolating between the two extremes of single neuron (labeled line, $LL$) and summed population ($SP$) coding. Here, these extensions are reviewed. For this kind of extension the above-mentioned trick for the van Rossum extension [7] is even more effective.

# 2 Measures of spike train distance

## 2.1 Notation

In the bivariate case we have two spike trains $x$ and $y$. We represent their spikes as $t_i^x$ and $t_j^y$ with $i = 1, \ldots, M_x$ and $j = 1, \ldots, M_y$ so $M_x$ and $M_y$ denote the numbers of spikes in $x$ and $y$, respectively. It is assumed the spikes are sorted in ascending order, so $t_i^x \leq t_{i+1}^x$ and $t_i^y \leq t_{i+1}^y$.

In the population case, we have two populations $X$ and $Y$ with $n = 1, \ldots, N$ spike trains each. Spikes are represented as $t_i^{x_n}$ and $t_j^{y_n}$ with $i = 1, \ldots, M_{x_n}$ and $j = 1, \ldots, M_{y_n}$; $M_{x_n}$ and $M_{y_n}$ denote the numbers of spikes in $x_n$ and $y_n$, the $n$th spike train of population $X$ and $Y$, respectively. Each population can also be represented by the pooled spike train which we denote as $t_i^X$ and $t_j^Y$ with $i = 1, \ldots, M_X$ and $j = 1, \ldots, M_Y$ as well as $M_X = \sum_n M_{x_n}$ and $M_Y = \sum_n M_{y_n}$.

## 2.2 The Victor-Purpura metric

The metric $D_V$ introduced in [21] defines the distance between two spike trains in terms of the minimum cost of transforming one spike train into the other using three basic operations: spike insertion, spike deletion and spike movement. Each is given a cost; one for inserting or deleting a spike and $c_V|\delta t|$ for moving a spike a temporal distance $\delta t$. The cost-per-time $c_V$ sets a time scale for the analysis.

In the minimum cost edit, a spike is never moved more than $2/c_V$ since the cost of doing that would be greater than deleting one of the spikes and

inserting another to match the spike in the other train. This means that for high $c_V$, the distance approaches the number of non-coincident spikes. In contrast, for small $c_V$, the distance approaches the difference in spike number because it is cheap to move spikes around and so most of the cost comes from adding spikes to the smaller spike train so that it matches the longer. Thus, by decreasing the cost, the Victor-Purpura metric is transformed from a timing distance to a rate distance.

This is a metric since it satisfies the three properties a distance must have to be a metric: symmetry, non-degeneracy and the triangular inequality. Although it is easiest to talk about transforming one spike train to match other, it makes no cost difference which spike train is being transformed and the distance is symmetric in the spike train order. Since each of the edit costs is positive, it is easy to see that the distance is only zero for two identical spike trains, this is non-degeneracy. The third condition, the triangular inequality, states that the distance between two spike trains is never greater than the distance taken via a third spike train; this follows from the definition of the distance as the minimum cost.

Although the Victor-Purpura metric is defined as a minimum cost, the calculation of the distance does not require a minimization: it can be calculated iteratively [21, 19]. This involves completing a $M_x \times M_y$ grid of distances between truncated spike trains; essentially the algorithm works by adding successive spikes at the ends of the spike trains.

## 2.3 The van Rossum metric

To describe the van Rossum metric [20] it is useful to first define a map from spike trains to functions: the spike train $x = \{t_1^x, t_2^x, \ldots, t_{M_x}^x\}$ is mapped to $f(t; x)$ by filtering it with a kernel $h(t)$:

$$x \mapsto f(t; x) = \sum_{i=1}^{M_x} h(t - t_i^x). \tag{1}$$

The kernel function has to be specified. In the original paper the causal exponential is used

$$h(t) = \begin{cases} 0 & t < 0 \\ e^{-t/\tau} & t \geq 0 \end{cases} \tag{2}$$

where the decay constant $\tau$ is the time scale which parameterizes the metric.

The van Rossum metric is induced on the space of spike trains by the $L^2$ metric on the space of functions. In other words, the distance between the two spike trains $x$ and $y$ is given by

$$D_R = \sqrt{\frac{2}{\tau}\int_0^\infty dt[f(t;x) - f(t;y)]^2} \tag{3}$$

where the normalizing factor of $2/\tau$ is included so that there is a distance of one between a spike train with a single spike and one with no spikes. The metric properties of the van Rossum metric follow from those of the $L^2$ metric; it is symmetric in $x$ and $y$, zero only when $f(x,t) = f(y,t)$ for all $t$, which, in turn, implies $x = y$ and

$$\begin{aligned}[f(t;x) - f(t;y)]^2 &= [f(t;x) - f(t;z) + f(t;z) - f(t;y)]^2 \\ &\leq [f(t;x) - f(t;z)]^2 + [f(t;z) - f(t;y)]^2 \end{aligned} \tag{4}$$

for all $t$, establishing the triangular inequality.

For the causal exponential filter, the integral in the formula for $D_R$, Eq. 3, can be done explicitly to give a distance [18, 17]

$$D_R^2 = \sum_{i,j} e^{-|t_i^x - t_j^x|/\tau} + \sum_{i,j} e^{-|t_i^y - t_j^y|/\tau} - 2\sum_{i,j} e^{-|t_i^x - t_j^y|/\tau}. \tag{5}$$

Recently, in [7], a trick has been presented which reduces the computational cost for the regular van Rossum metric between two spike trains of similar length, $M_x \sim M_y$, to order $M = (M_x + M_y)/2$ from order $M^2$. The idea behind this trick is to imitate how things are calculated in biological systems where, for example, in developmental biology global order is established by local responses to chemical gradients; in our case this means replacing the numerous pairwise calculations with a running tally. Specifically, a new vector, referred to as a *markage vector* in [7], is defined: given the spike train $x$, the markage vector will have the same number, $M_x$, of element as there are spikes and the entries are defined recursively so that $m_1^x = 0$ and

$$m_i^x = (m_{i-1}^x + 1)e^{-(t_i^x - t_{i-1}^x)/\tau}. \tag{6}$$

This means

$$m_i^x = \sum_{j|i>j} e^{-(t_i^x - t_j^x)/\tau} \tag{7}$$

where we recall that the spikes are in ascending order, so the exponent is negative. This quantity is equal to the left limit of $f(t,x)$ at $t_i^x$, the value it would have at $t = t_i^x$ but for there being a spike there.

The markage vector is used to reduce the double sums in the expression for $D_R(x,y)$, Eq. 5, to single sums. For convenience this equation is first rewritten to avoid the use of the absolute value

$$\begin{aligned} D_R^2 &= \frac{M_x + M_y}{2} + \sum_i \sum_{j|i>j} e^{-(t_i^x - t_j^x)/\tau} + \sum_i \sum_{j|i>j} e^{-(t_i^y - t_j^y)/\tau} \\ &\quad - \sum_i \sum_{j|t_i^x > t_j^y} e^{-(t_i^x - t_j^y)/\tau} - \sum_i \sum_{j|t_i^y > t_j^x} e^{-(t_i^y - t_j^x)/\tau}, \end{aligned} \tag{8}$$

where $j|t_i^x > t_j^y$ indicates that the sum is restricted to values of $j$ where $t_j^y < t_i^x$. This yields

$$\sum_i \sum_{j|i>j} e^{-(t_i^x - t_j^y)/\tau} = \sum_i m_i^x. \tag{9}$$

The cross-like terms are trickier, let $t_{\rm P}^y(t_i^x)$ denote the last spike time in $y$ that is earlier than $t_i^x$, hence

$$t_{\rm P}^y(t_i^x) = \max_j \left(t_j^y | t_i^x > t_j^y\right), \tag{10}$$

which leads to

$$\begin{aligned} \sum_i \sum_{j|t_i^x > t_j^y} e^{-(t_i^x - t_j^y)/\tau} &= \sum_i e^{-(t_i^x - t_{\rm P}^y(t_i^x))/\tau} \sum_{j|t_{\rm P}^y(t_i^x) \geq t_j^y} e^{-(t_{\rm P}^y(t_i^x) - t_j^y)/\tau} \\ &= \sum_i e^{-(t_i^x - t_{\rm P}^y(t_i^x))/\tau} \left( 1 + \sum_{j|t_{\rm P}^y(t_i^x) > t_j^y} e^{-(t_{\rm P}^y(t_i^x) - t_j^y)/\tau} \right) \\ &= \sum_i e^{-[t_i^x - t_{\rm P}^y(t_i^x)]/\tau} \left[1 + m_P^y(t_i^x)\right]. \end{aligned} \tag{11}$$

where $m_{\rm P}^y(t_i^x)$ is the value of the $y$ markage vector corresponding to $t_{\rm P}^y(t_i^x)$, in other words, if $j$ is the index of $t_{\rm P}^y(t_i^x)$, so $t_j^y = t_{\rm P}^y(t_i^x)$, then $m_{\rm P}^y(t_i^x) = m_j^y$. Since the other two terms in the expression for $D_R$ are identical to the two above with $x$ and $y$ switched they can be calculated in the same way, so, for example,

$$\sum_i \sum_{j|t_i^y > t_j^x} e^{-(t_i^y - t_j^x)/\tau} = \sum_i e^{-[t_i^y - t_{\rm P}^x(t_i^y)]/\tau} \left[1 + m_{\rm P}^x(t_i^y)\right] \tag{12}$$

This trick reduces the calculation of all four terms from $M^2$ to $M$, however, using the markage vector does introduce extra calculations. Along with the calculation of the markage vector itself, there is the need to calculate $t_{\mathrm{P}}^y(t_i^x)$, this can be calculated iteratively by advancing it from its previous value when necessary. It is demonstrated in [7] that the constant prefactor to $M$ in the algorithm with markage is larger than the prefactor to $M^2$ in the traditional algorithm, but that it is worthwhile using the markage algorithm even for quite short spike trains.

## 2.4 The ISI- and the SPIKE-distance

For the van Rossum metric each spike train is initially transformed into a continuous function. Both the ISI- and the SPIKE-distance build on a similar first step, however, here the discrete spike times of a pair of spike trains are immediately transformed into a time profile, that is, a temporal sequence of instantaneous dissimilarity values. The overall distance is then the average of the respective time profile, so, for example, for the bivariate SPIKE-distance,

$$D_S = \frac{1}{T}\int_{t=0}^{T} S(t)dt \tag{13}$$

where $T$ denotes the overall length of the spike trains which would often be the duration of the recording in an experiment. In the following this equation is always omitted, and the discussion is restricted to showing how to derive the respective time profiles. In fact, the time profile are an important aspect of these distances since they allow for a time-local description of spike-train dissimilarity.

Both time profiles rely on three piecewise constant quantities (Fig. 1) which are assigned to each time instant between zero and $T$. For the spike train $x$ these are the time of the preceding spikes

$$t_{\mathrm{P}}^x(t) = \max_i(t_i^x | t_i^x \leq t), \tag{14}$$

the time of the following spikes

$$t_{\mathrm{F}}^x(t) = \min_i(t_i^x | t_i^x > t), \tag{15}$$

as well as the instantaneous interspike interval

$$\lambda_{\mathrm{I}}^x(t) = t_{\mathrm{F}}^x(t) - t_{\mathrm{P}}^x(t). \tag{16}$$

The ambiguity regarding the definition of the very first and the very last interspike interval is resolved by adding an auxiliary leading spikes at time $t = 0$ and auxiliary trailing spikes at time $t = T$ to each spike train.

### 2.4.1 The ISI-distance

The time profile of the ISI-distance [10] is calculated as the instantaneous ratio between the interspike intervals $\lambda_{\mathrm{I}}^x$ and $\lambda_{\mathrm{I}}^y$ (Eq. 16) according to:

$$I(t) = \Lambda(1/\lambda_{\mathrm{I}}^x(t), 1/\lambda_{\mathrm{I}}^y(t)) \tag{17}$$

where

$$\Lambda(r_1, r_2) = \begin{cases} r_2/r_1 - 1 & \text{if } \ r_2 \le r_1 \\ 1 - r_1/r_2 & \text{otherwise.} \end{cases} \tag{18}$$

This ISI-ratio equals zero for identical ISI in the two spike trains, and approaches $-1$ and 1, respectively, if the first or the second spike train is much faster than the other. For the ISI-distance the temporal averaging in Eq. 13 is performed on the absolute value of the ISI-ratio and, therefore, treats both kinds of deviations equally.

The ISI-distance is a metric. Symmetry follows directly from the definition and the triangular inequality has been demonstrated in [15]. In fact, they show the integrand $|I(t)|$ satisfies the triangular inequality for each value of $t$. Without loss of generality take $r^x \le r^y$ so

$$|\Lambda(r^x, r^y)| = 1 - \frac{r^x}{r^y} \tag{19}$$

and consider $r^z$, the rate at $t$ for any other spike train $z$. If $r^z \le r^x$ then

$$0 \le (r^x - r^z)(r^x + r^y) = (r^x)^2 + r^x r^y - r^z r^y - r^z r^x. \tag{20}$$

Bringing the $(r^x)^2$ term to the left and dividing by $r^x r^y$ this gives

$$|\Lambda(r^x, r^y)| = 1 - \frac{r^x}{r^y} \le 1 - \frac{r^z}{r^x} + 1 - \frac{r^z}{r^y} = |\Lambda(r^x, r^z)| + |\Lambda(r^z, r^y)|. \tag{21}$$

The other two cases, $r^x \le r^z \le r^y$ and $r^x \le r^y \le r^z$ follow in the same way using, respectively,

$$\begin{aligned} 0 &\le (r^z - r^x)(r^y - r^z) \\ 0 &\le (r^z - r^y)(r^x + r^y). \end{aligned} \tag{22}$$

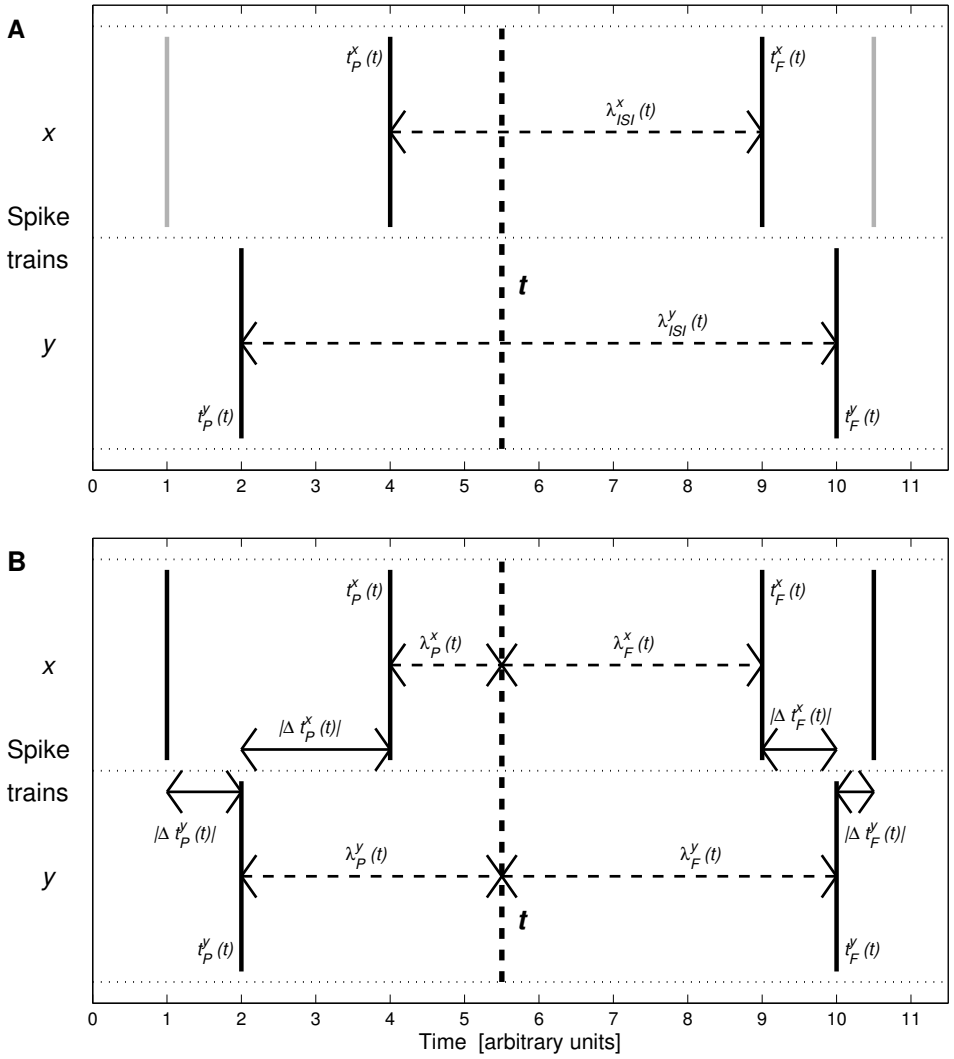


Figure 1: Illustration of the local quantities needed to define the time profiles of the two time-resolved distances for an arbitrary time instant $t$: **A**. ISI-distance. **B**. SPIKE-distance.

Non-degeneracy also holds, but in a subtle way. $I(t)$ is only zero if $r^x(t) = r^y(t)$. As described in [15], if $x$ and $y$ are periodic spike trains with the same, constant, firing rate they could differ in phase, but still have $I(t) = 0$. However, in the definition of the ISI-distance given here, with auxiliary spikes added at the beginning and end of the spike train, for a positive phase the boundary condition will break the periodicity for at least one of the spike trains and lift this degeneracy.

### 2.4.2 The SPIKE-distance

The ISI-distance relies on the relative length of simultaneous interspike intervals and is thus well-designed to quantify similarities in the neurons' firing-rate profiles. However, it is not ideally suited to track synchrony that is mediated by spike timing. The interspike interval is often larger than the changes in relative spike times between spikes in the two spikes trains and so a time profile which is based only on interspike intervals is often not useful for tracking changes in synchrony.

This kind of sensitivity can be very relevant since coincident spiking is found in many different neuronal circuits. It is important here, since a time profile graph is most likely to be useful if it plots instantaneous changes in synchrony. This issue is addressed by the SPIKE-distance which combines the properties of the ISI-distance with a specific focus on spike timing; see Kreuz et al. [12] for the original implementation and Kreuz et al. [14] for the improved version presented here.

The time profile of the SPIKE-distance relies on differences between the spike times in the two spike trains. It is calculated in two steps: First for each spike the distance to the nearest spike in the other spike train is calculated, then for each time instant the relevant spike time differences are selected, weighted, and normalized. Here 'relevant' means local; each time instant is uniquely surrounded by four *corner spikes*: the preceding spike of the first spike train $t^x_{\mathrm{P}}$, the following spike of the first spike train $t^x_{\mathrm{F}}$, the preceding spike of the second spike train $t^y_{\mathrm{P}}$, and, finally, the following spike of the second spike train $t^y_{\mathrm{F}}$. To each of these corner spikes can be identified with a spike time difference, for example, for the previous spike of the first spike train

$$\Delta t^x_{\mathrm{P}} = \min_i(|t^x_{\mathrm{P}} - t^y_i|). \tag{23}$$

and analogously for $t^x_{\mathrm{F}}$, $t^y_{\mathrm{P}}$, and $t^y_{\mathrm{F}}$.

For each spike train separately a locally weighted average is employed such that the differences for the closer spike dominate; the weighting factors depend on

$$\lambda_{\mathrm{P}}^{n}(t) = t - t_{\mathrm{P}}^{n}(t) \tag{24}$$

and

$$\lambda_{\mathrm{F}}^{n}(t) = t_{\mathrm{F}}^{n}(t) - t, \tag{25}$$

the intervals to the previous and the following spikes for each neuron $n = x, y$. The local weighting for the spike time differences of the first spike train reads

$$S_x(t) = \frac{\Delta t_{\mathrm{P}}^{x} \lambda_{\mathrm{F}}^{x} + \Delta t_{\mathrm{F}}^{x} \lambda_{\mathrm{P}}^{x}}{\lambda_{\mathrm{I}}^{x}}$$

and analogously $S_y(t)$ is obtained for the second spike train. Averaging over the two spike train contributions and normalizing by the mean interspike interval yields the SPIKE-distance

$$S(t) = \frac{S_x(t) + S_y(t)}{2\langle \lambda_{\mathrm{I}}^{n} \rangle_n}. \tag{26}$$

It seems likely that the SPIKE-distance is also a metric, although this has not yet been proved. It is certainly symmetric and non-degenerate. As for the triangular inequality, while it does not hold for all $t$, it might hold after the time profile has been integrated.

## 2.5 Entropy-based measure

Another formulation of a distance between spike trains is provided by the entropy. This does not lead to a metric but it is interesting to consider because it can be used to quantify the dissimilarity of the estimated firing rates for the different spike trains regardless of how the firing rates are estimated and because it has a natural generalization to populations.

If $r^x(t)$ and $r^y(t)$ are the estimated rates for the two spike trains, estimated either using the ISI or by filtering; then the conditional probability of a spike at time $t$ in spike train $n = x$ or $y$, conditioned on there being a spike in one of the two spike trains, is

$$p^n(t) = \frac{r^n(t)}{r^x(t) + r^y(t)}. \tag{27}$$

The entropy for this conditional probability reads

$$H_2(p^x(t), p^y(t)) = -p^x(t)\log_2 p^x(t) - p^y(t)\log p^y(t). \tag{28}$$

If the two rates are very different this is close to zero, if they are very similar, it is close to one; as such a distance measure given by

$$I_H(t) = 1 - H_2(p^x(t), p^y(t)). \tag{29}$$

measures the distance between the two trains.

# 3 Comparisons

## 3.1 The ISI- and the SPIKE-distance

The ISI-distance is based on interspike intervals and quantifies covariations in the local firing rate, while the SPIKE-distance tracks synchrony mediated by spike timing (Fig. 2). Note that this does not mean that the ISI-distance is sensitive to rate coding and the SPIKE-distance sensitive to temporal coding. It is the relative timing of interspike intervals and spikes, respectively, that matters.

## 3.2 The ISI-distance and the van Rossum metric

Although, as described in Appendix A of [11], there are subtleties in relating the firing rate to the expected inter-spike interval, $1/\lambda^x(t)$ and $1/\lambda^y(t)$ are, roughly speaking, instantaneous estimates of the firing rate at $t$; in fact it resembles a k-th nearest neighbor estimate with $k = 2$. In this way, the ISI-distance resembles the van Rossum metric and it is possible to mix-and-match: the $1/\lambda^x(t)$ and $1/\lambda^y(t)$ in the formula for $I(t)$ could be replaced by the filtered functions $f(t;x)$ and $f(t;y)$ used in the van Rossum metric and vice versa.

The functional form of $I(t)$ used in the ISI-distance is chosen in order to give a good representation of how the difference between the two spike trains evolves with $t$. The analogous functional form for the van Rossum metric is

$$R(t) = [r^x(t) - r^y(t)]^2 \tag{30}$$

where $r^x(t)$ and $r^y(t)$ are estimates of the firing rate. The ISI-function is invariant under a rescaling of the two firing rates by the same factor, the van

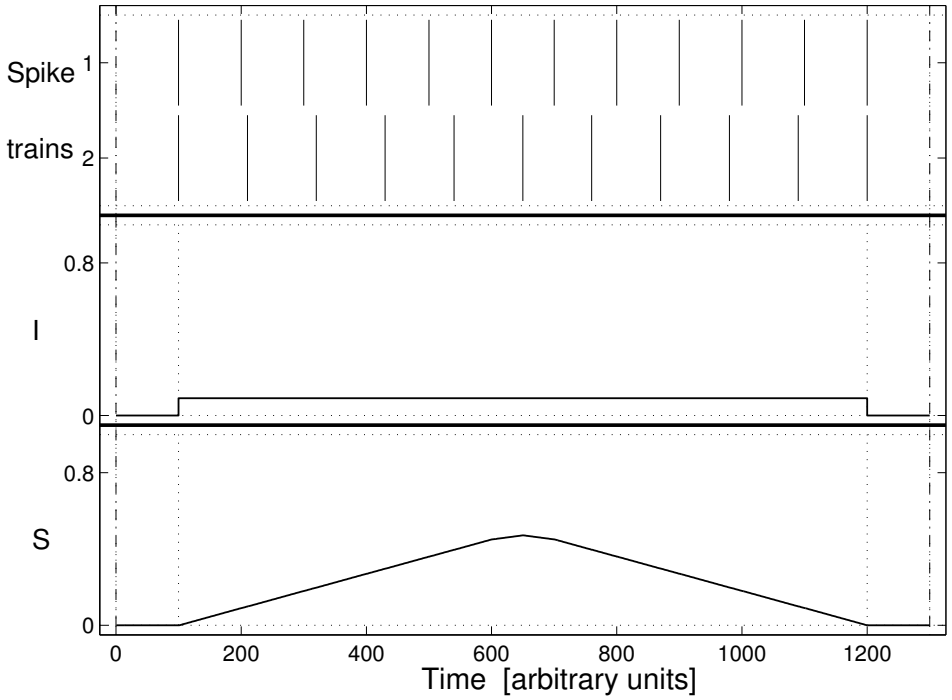


Figure 2: Comparison of the ISI- and the SPIKE-distance on constructed spike trains. Here $x$ and $y$ have constant, but different, firing rates. While $I(t)$ reflects covariations in the local firing rate and gives a constant time profile, $S(t)$ is sensitive to differences in spike timing and so it increases as the spike trains go out of phase.

Rossum function is not, but it does link the metric to an $L^2$ structure, which may prove useful in some mathematical applications.

## 3.3 The SPIKE-distance and the Victor-Purpura metric

Like the Victor-Purpura metric, the SPIKE-distance depends on gaps between spikes. However, the Victor-Purpura metric pairs up spikes whereas for the SPIKE-distance more than one spike in one spike train can be matched to a given spike in the other. Another difference is that the Victor-Purpura metric has a cut-off, spikes are not paired if they are more than $2/c_V$ apart; in the SPIKE-distance the gap between an isolated pair of spikes can contribute to $S(t)$ even if they are a large distance apart.

Of course, this could be changed by replacing the $\Delta t$s with saturating functions, however, this would introduce a scale into the SPIKE-distance. Moreover, the definition of $S(t)$ does have the interspike interval in the denominator, limiting the effect large gaps have on the distance. These differences are illustrated in Fig. 3.

## 3.4 Comparison of all distances on birdsong data

A common approach to evaluating how well a spike-train distance measure succeeds is to test it using clustering [21]. If a set of spike trains is made up of multiple responses to a set of different stimuli then they can be clustered so each cluster consists of different responses to a single stimulus. The distance measure is evaluated by quantifying how well it succeeds in measuring small distances between spike trains in the same cluster and longer distances between spike trains in different clusters.

One measure of clustering performance is the normalized transmitted information $\tilde{h}$ [21]. This is calculated from a confusion matrix, a $n_s \times n_s$ matrix, where $n_s$ is the number of stimuli. Starting with a matrix of zeros, one of the responses is chosen and the rest are clustered according to stimulus. If the chosen response is a response to stimulus $i$ and is closest to the cluster $j$, one is added to the $ij$th entry, $N_{ij}$. This is repeated with each of the responses used as the chosen response, so that at the end the entries in the confusion matrix add up to give the total number of responses.

A good distance function should measure shorter distances between a response and its own cluster, leading to more diagonal elements in the con-

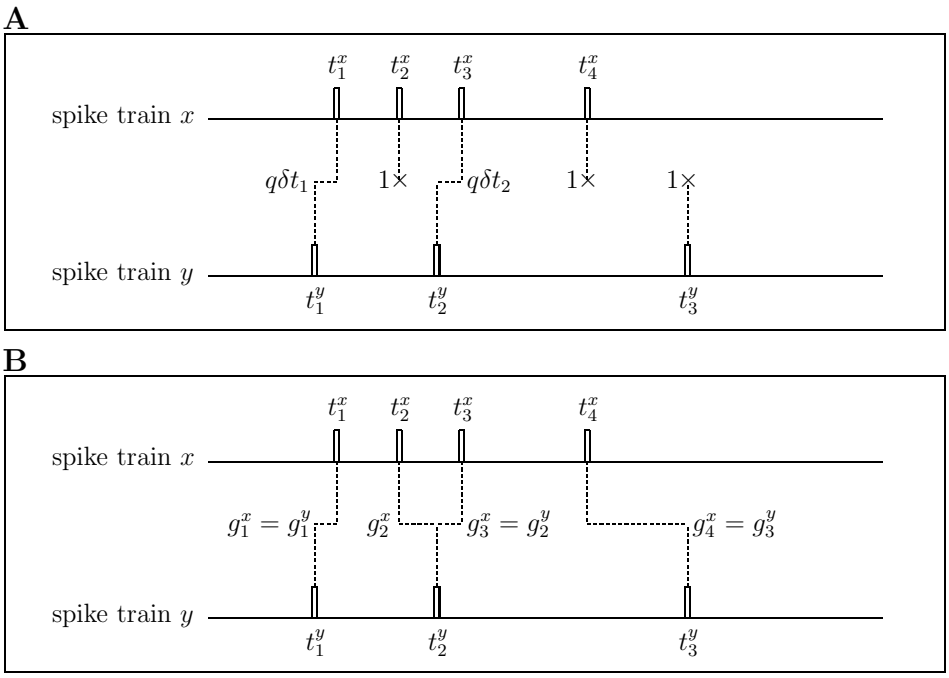


Figure 3: The difference between how the Victor-Purpura metric and SPIKE-distance pair spikes. In **A** the Victor-Purpura metric pairs the spikes in such a way as to produce the lowest cost edit, this means $t_2^x$, $t_4^x$ are deleted and $t_3^y$ added, the others are paired. In **B** the gaps are given by the nearest spike in the other train, so both $t_2^x$and $t_3^x$ are paired with $t_2^y$ and $t_4^x$ is paired with $t_3^y$, even though they are well separated from each other.

fusion matrix. The normalized transmitted information quantifies this, for equally likely stimuli, it is given by

$$\tilde{h} = \frac{1}{\sum_{ij} N_{ij}} \sum_{ij} N_{ij} \left( \log_{n_s} N_{ij} - \log_{n_s} \sum_k N_{kj} - \log_{n_s} \sum_k N_{ik} + \log_{n_s} \sum_{ij} N_{ij} \right). \tag{31}$$

where $\log_{n_s} N_{ij}$, for example, is the logarithm to the base $n_s$ of $N_i j$, so, $\log_{n_s} N_{ij} = \ln N_{ij} / \ln n_s$. Roughly, this measures how well clustering by distance transmits information about the stimulus-based clustering. A low value corresponds to low transmitted information and, therefore, a poor metric. Values close to one, the maximum, indicate that the metric performs well. One complication to this procedure is that a weighted average of the distances is often used to reduce the effect of outliers, this was described in [21] and is reviewed in [9].

This approach is used to evaluate the different distance measures considered here. The example test data used is a set of spiking responses recorded from the primary auditory neurons of zebra finch during repeated playback of songs from a standard repertoire. These electrophysiological data were original described in [16, 23] and these paper should be consulted for a detailed description of the experimental and primary data processing procedures. They are used as an example data set for evaluating metrics in [6, 9]. At each of the 24 sites ten responses to each of 20 zebra finch songs were recorded. The result of this test is shown in Fig. 4. It demonstrates a roughly comparable performance of the different measures. Although the van Rossum metric performs slightly better, the convenient ISI-distance does not come far behind, see (Fig. 4. The combination of entropy and the ISI rate estimate performs well, possibly reflecting the ability of k-th nearest neighbor methods to accurately estimate conditional probabilities. Finally, on this data set the SPIKE-distance performs better than the Victor-Purpura metric.

# 4 Measuring the dissimilarity within a population

For all spike train distances there exists a straightforward extension to the case of more than two spike trains, the averaged bivariate distance. However, for the ISI- and the SPIKE-distance this average over all pairs of neurons commutes with the average over time, so in order to achieve the same

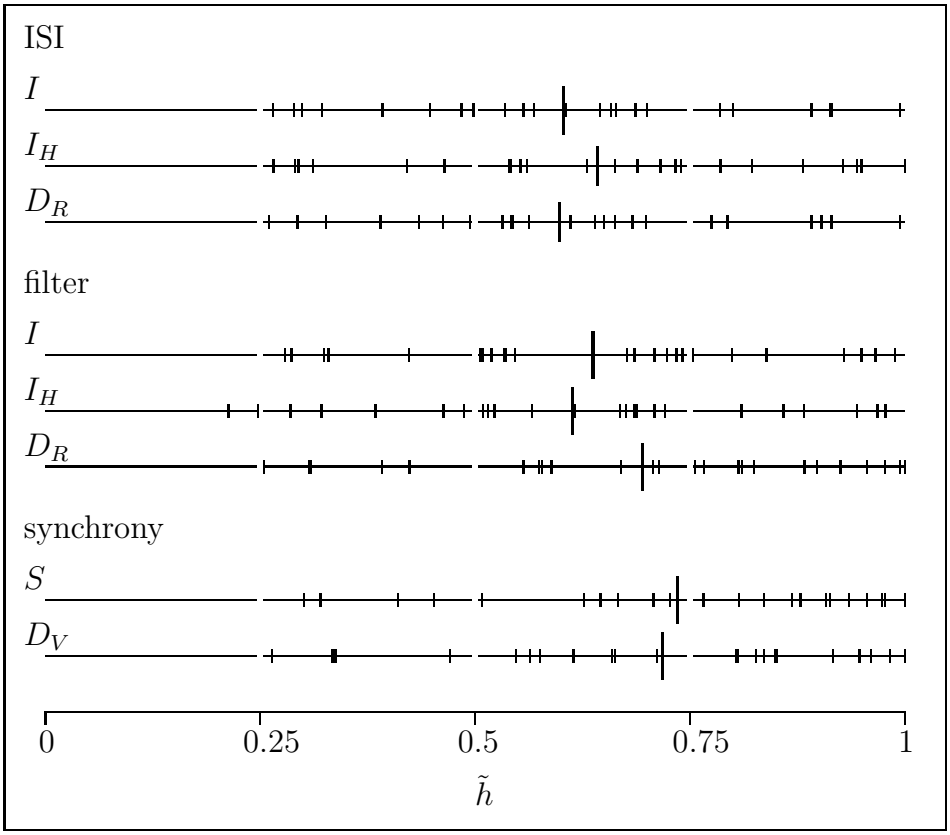


Figure 4: Comparing the various distance measures. In this figure the $\tilde{h}$ value has been plotted for each of the 24 sites in the zebra finch data. Each horizontal line corresponds to the performance of a single metric, the line runs from zero to one, as a visual aid a tiny gap is left at 0.25, 0.5 and 0.75. Along each line a small stroke corresponds to a single site, the long stroke corresponds to the average value. In three graphs marked 'ISI' ISI use the estimate the rate, the next three, marked 'filter' use the filter, with $\tau = 12.5$ms. For the two graphs marked $I(t)$ the function is $\Lambda(r_1, r_2)$ (Eq. 18), for the graphs marked $I_H$, the entropy-based function is used (Eq. 29) and for the graphs marked $D_H$ the $L^2$ function (Eq. 30) is used. Finally, in the section marked 'synchrony', $S$ gives the SPIKE-distance and $D_R$ the Victor-Purpura metric with $c_V = 71\text{s}^{-1}$. The values of $\tau$ and $c_V$ have been chosen to give a good overall performance, obviously choosing a different $\tau$ or $c_V$ for each site would improve performances, but would give a poorer comparison for the parameter-free methods.

kind of time-resolved visualization as in the bivariate case it is convenient to first calculate the instantaneous average, for example, $S^{\mathrm{a}}(t)$ over all pairwise instantaneous values $S^{mn}(t)$,

$$S^{\mathrm{a}}(t) = \frac{1}{N(N-1)/2} \sum_{n=1}^{N-1} \sum_{m=n+1}^{N} S^{mn}(t) \tag{32}$$

and then calculate the distance by averaging the resulting time profile using Eq. 13. All time profiles and thus all distances are bounded in the interval $[0, 1]$. The distance value zero is obtained for identical spike trains only.

An exemplary application of both distances to artificially generated multivariate data can be found in Fig. 5.

The entropy based measure (see Section 2.5) can be extended in a natural way to measure the similarity of a collection of spike trains. For the population $X$

$$I_H(t) = 1 - H_{M_x}(p^{x_1}(t), \ldots, p^{x_{M_x}}(t)) \tag{33}$$

where $H_{M_x}(p^{x_1}(t), \ldots, p^{x_{M_x}}(t)) = H_2(p^{x_1}(t), \ldots, p^{x_{M_x}}(t)) / \log_2(M_x)$ and

$$p^{x_i}(t) = \frac{r^{x_i}(t)}{\sum_j r^{x_j}(t)} \tag{34}$$

# 5 Measuring the dissimilarity between populations

## 5.1 The population extension of the Victor-Purpura metric

A population extension of the Victor-Purpura-Distance has been proposed [2] which can be used to uncover if and how populations of neurons cooperate to encode a sensory input. This extension adds one further edit type to the existing three allowed edits: a spike can be relabeled from one neuron to another at a cost of $k$. Thus, for simultaneous recordings from proximate neurons, spikes are labeled by the neuron that fired them, but this label can be changed at a cost of $k$, where $k$ is a second parameter. Hence, the population distance $D_V^p$ between two sets of labeled spike trains is the cheapest set of elementary moves transforming one set into the other, where, now, the

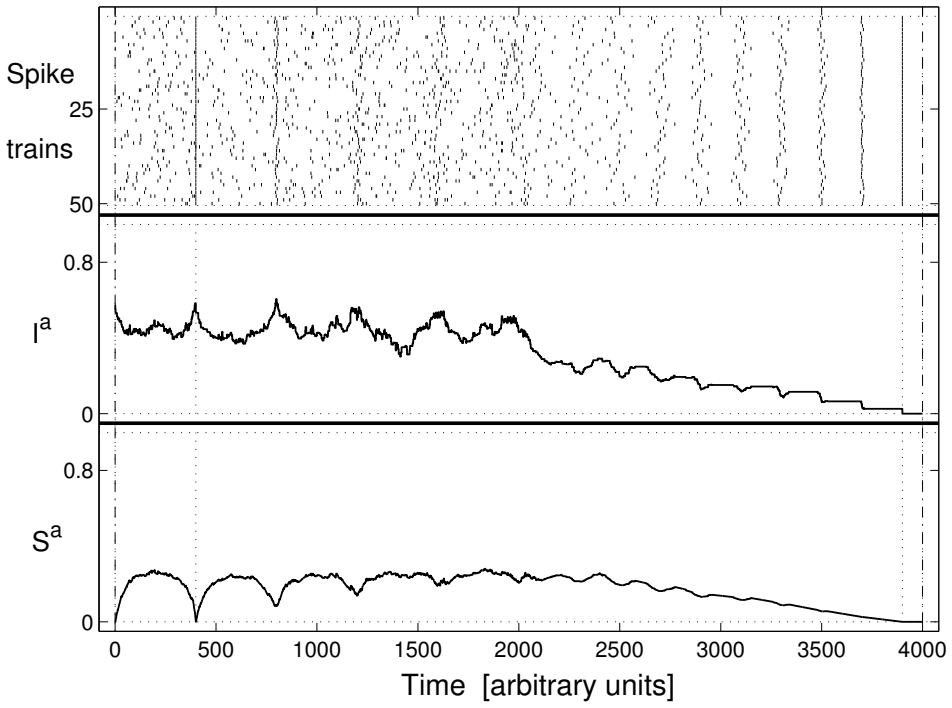


Figure 5: Comparing the population distance measures. In the first half within the noisy background there are four regularly spaced spiking events with increasing jitter. The second half consists of ten spiking events with decreasing jitter but now without any background noise. Both $I^a$ and $S^a$ distinguish between the noisy and noiseless period and chart the decline in jitter in the noiseless period. $I^a$ has peeks for the spiking event with noise, since the noise causes wide fluctuations in the interspike intervals. In contrast, $S^a$ has troughs, since the spiking events give local synchrony.

elementary moves are adding or deleting a spike at a cost of one, moving a spike by an interval $\delta t$ at a cost of $q|\delta t|$ and relabeling a spike at a cost of $k$.

In [2] two different coding strategies for neuron populations are distinguished: a 'summed population code' ($SP$) metric where the two spike trains from the two neurons are super-imposed before the distance is calculated, and a 'labeled line code' ($LL$) metric where the distance is measured for each neuron separately and then added. These two possibilities correspond to the metrics at either end of the one-parameter family of $D_V^p$ metrics obtained by varying $k$.

This dimensionless parameter quantifies the importance of distinguishing spikes fired by different neurons. When $k = 0$, there is no cost for reassigning a spike's label, and the entire population discharge is viewed as a sequence of spikes fired by a single-unit, corresponding to a $SP$ metric. When $k \geq 2$, spikes fired by different neurons are never considered similar, since deleting two spikes with different labels, for a cost of two, is not more expensive than changing their labels to match. For values of $k$ in $[0, 2]$, spikes fired by different neurons $\delta t$ apart can be matched in a transformation if the cost of this transformation step, $c_V|\delta t| + k$, is less than two. Thus, for these values of $k$, spikes fired by different neurons can be considered similar if they occur within $(2 - k)/c_V$ of each other. For $k \geq 2$ the population distance $D_V^p$ between the two sets of spike trains is the same as the sum of the individual Victor-Purpura metrics. Therefore, this is a $LL$ metric.

Details of algorithms for calculating this distance can be found in [1] and [22].

## 5.2 The population extension of the van Rossum metric

Also the van Rossum metric has been extended to the population case [8] allowing a spike train distance to be measured between the two populations $X$ and $Y$. This works by mapping the $N$ spike trains of the two population responses to a vector of functions using a different direction in the vector space for each neuron. The interpolation between the labeled line ($LL$) coding and the summed population ($SP$) coding, is given by varying the angle $\theta$ between these directions from zero ($SP$) to $\pi/2$ ($LL$) or, equivalently,

the parameter $c = \cos\theta$ from one to zero:

$$D_R = \sqrt{\frac{2}{\tau}\sum_n \left( \int_0^\infty |\delta_n|^2 dt + c \sum_{m \neq n} \int_0^\infty \delta_m \delta_n dt \right)} \tag{35}$$

with $\delta_n = f(t, x_n) - f(t, y_n)$.

Here, computational expense is an even greater difficulty but the same markage algorithm as in Section 2.3 can be used [7]. First, similar to the bivariate case above, the population van Rossum distance can also be estimated in a more efficient way by integrating analytically:

$$D_R = \sqrt{\sum_n \left( R_n + \; c \sum_{m \neq n} R_{nm} \right)} \tag{36}$$

with

$$R_n = \sum_{i,j} e^{-|t_i^{x_n} - t_j^{x_n}|/\tau} + \sum_{i,j} e^{-|t_i^{y_n} - y_j^{y_n}|/\tau} - 2\sum_{i,j} e^{-|t_i^{x_n} - t_j^{y_n}|/\tau} \tag{37}$$

being the bivariate single neuron distance, as in Eq. 5, between the $p$-th spike trains of $u$ and $v$ and

$$\begin{aligned} R_{nm} \;\; = \;\; & \sum_{i,j} e^{-|t_i^{x_n} - t_j^{x_m}|/\tau} + \sum_{i,j} e^{-|t_i^{y_n} - t_j^{y_m}|/\tau} \\ & - \sum_{i,j} e^{-|t_i^{x_n} - t_j^{y_m}|/\tau} - \sum_{i,j} e^{-|t_i^{y_n} - t_j^{x_n}|/\tau} \end{aligned} \tag{38}$$

representing the cross-neuron terms that are needed to fully capture the activity of the pooled population. As in the estimate in Eq. 35 the variable $c$ interpolates between the labeled line and the summed population distance.

These are the two equations for which the markage trick can be applied since all of these sums of exponentials can be rewritten using markage vectors in the same way as it was done in Eqs. 8, 9, and 11.

# 6 Discussion

In this chapter we have reviewed measures of spike train dissimilarity; we have described four different distance measures, the Victor-Purpura and the

van Rossum metric as well as the ISI- and the SPIKE-distance and we have indicated some of the ways they differ and some of the ways they resemble each other. While these spike train distances are applied to quantify the dissimilarity between just two individual spike trains, electrophysiological population data is becoming increasingly important and the spike trains distance measures are being extended so that they can be applied to populations. We review the progress that has been made on this problem. There are two kinds of population extensions aimed at measuring the dissimilarity either within one population of spike trains or between two populations of spike trains.

The measures for quantifying the dissimilarity of pairs of spike trains can be grouped in different ways. One basic division groups the ISI-distance and van Rossum metric together; in each of these two cases a time function is associated with each of the two spike trains and these functions are compared to give the distance measure. A third member of this group, based on entropy, is suggested. In the other group, the SPIKE-distance and Victor-Purpura metric compare spikes in the two spike trains directly and in each case the dissimilarity measure is based on the local differences in individual spike times. However, the two distance measures do differ in how they describe these local timing differences, the Victor-Purpura metric is an edit distance metric whereas the SPIKE-distance quantifies dyssynchrony.

The distance measures can also be grouped according to whether they require a time scale parameter or not. The Victor-Purpura metric and the van Rossum metric both depend on a time scale parameter whereas the ISI- and the SPIKE-distance do not. Often, in applying distance measures to experimental situations, the need to specify a time scale is awkward since it is unclear how it should be calculated, particularly in situations like maze exploration by rats where, without identical trials, it is difficult to optimize the metric by calculating the transmitted information. This certainly makes the ISI- and the SPIKE- distance more convenient.

Of course, the time scale parameter in the Victor Purpura and the van Rossum metric may allow them to probe aspect of the spike train structure that are invisible to the ISI- and the SPIKE-distance. Here, the performance of the metrics is evaluated using clustering and this does not show any strong difference in performance across the four distance measures and so fails to demonstrate an advantage of the distance measure with a parameter. It should be noted though that the evaluation only uses one data set and, hopefully, as these distance measures are applied to other data sets it will

become clearer how the properties of the different measures are related to properties of spike trains.

For example, a prominent assumption is that the optimal values of the time scale parameter in the Victor-Purpura and the van Rossum metric might be related to temporal properties of coding in the spike trains [21]. However, how this can be described precisely is unclear. In fact, in Chicharro et al. [3] and Chicharro et al. [4] the optimal value of the time scale parameter was studied for transient constant and time-varying stimuli and it was shown that the optimal time scale obtained from a spike train discrimination analysis is far from being conclusive. Rather it results in a non-trivial way from the interplay of many different factors such as the distribution of the information contained in different parts of the response or the degree of redundancy between them.

The time scale parameter in the Victor-Purpura metric and the van Rossum metric does allow them to test for neuronal codes ranging from a rate code to a coincidence detector. It might turn out that in some neuronal systems the type of neuronal coding means the data favors one spike train analysis tool over another and, if a particular coding scheme is assumed, specific measures are needed for a confirmatory analysis, otherwise more general measures are very well suited for an exploratory analysis [12].

The population codes are not yet fully developed, for example, no distance measures are given here for using the ISI- and the SPIKE-distance to compare population responses. It is likely that multi-neuron distance measures generalizing the ISI- and the SPIKE-distance could be defined using the population Victor-Purpura and van Rossum metrics as guides.

As with the time scale parameter, the population Victor-Purpura and van Rossum metrics have a parameter which interpolates between coding schemes, in this case between the summed population code and the labeled line code. It would also be interesting to understand this parameter in a more principled way and to extend the parametrization to other population coding schemes. More importantly, although early applications of the population metrics have been promising, with, for example, an interesting investigation of population auditory coding in grasshoppers using the population van Rossum metric [5], there have not been a substantial number of investigations using these methods. That, however, seems likely to change in the future.

We close by noting that links to source codes (mostly Matlab but also some C++ and Python) for all methods dealt with in this chapter can be

found under http://www.fi.isc.cnr.it/users/thomas.kreuz/sourcecode.html.

## Acknowledgements

We gratefully acknowledge Daniel Chicharro for useful discussions and a careful reading of the manuscript. CJH is grateful to the James S McDonnell Foundation for financial support through a Scholar Award in Human Cognition. We thank Kamal Sen for providing the zebra finch data analyzed here.

## References

[1] Aronov D. Fast algorithm for the metric-space analysis of simultaneous responses of multiple single neurons. J Neurosci Methods, 2003;124:175-9.

[2] Aronov D, Reich DS, Mechler F, Victor JD. Neural coding of spatial phase in V1 of the macaque monkey. J Neurophysiol, 2003;89:3304-27.

[3] Chicharro D, Kreuz T, Andrzejak RG. What can spike train distances tell us about the neural code? J Neurosci Methods, 2011;199:146-65.

[4] Chicharro D, Caporello E, Gentner TQ, Kreuz T, Andrzejak RG. Disambiguating natural stimulus encoding and discrimination of spike trains. submitted, 2012.

[5] Clemens J, Kutzki O, Ronacher B, Schreiber S, Wohlgemuth S. Efficient transformation of an auditory population code in a small sensory system, Proceedings of the National Academy of Sciences, 2011;108:13812-13817.

[6] Houghton C. Studying spike trains using a van Rossum metric with a synapses-like filter. Journal of Computational Neuroscience, 2009;26:149–155.

[7] Houghton C, Kreuz T. On the efficient calculation of van Rossum distances. Network: Computation in Neural Systems, 2012;23:48-58.

[8] Houghton C, Sen K. A new multineuron spike train metric. Neural Comput, 2008;20:1495-511.

[9] Houghton C and Victor P. *Spike rates and spike metrics* in Visual Population Codes: Toward a Common Multivariate Framework for Cell Recording and Functional Imaging, ed. Kriegeskorte N and Kreiman G. (MIT Press, Cambridge MA, 2012).

[10] Kreuz T, Haas JS, Morelli A, Abarbanel HDI, Politi A. Measuring spike train synchrony. J Neurosci Methods, 2007;165:151-61.

[11] Kreuz T, Chicharro D, Andrzejak RG, Haas JS, Abarbanel HDI. Measuring multiple spike train synchrony. J Neurosci Methods, 2009;183:287-99.

[12] Kreuz T, Chicharro D, Greschner M, Andrzejak RG. Time-resolved and time-scale adaptive measures of spike train synchrony. J Neurosci Methods, 2011;195:92-106.

[13] Kreuz T. Measures of spike train synchrony. Scholarpedia, 2011;6:11934.

[14] Kreuz T, Chicharro D, Houghton C, Andrzejak RG, Mormann F. Monitoring spike train synchrony. J Neurosci Methods (in preperation XXXXX), 2012.

[15] Lyttle D, Fellous JM. A new similarity measure for spike trains: Sensitivity to bursts and periods of inhibition. J Neurosci Methods, 2011;199:296-309.

[16] Narayan R, Graña GD, Sen K (2006) Distinct time-scales in cortical discrimination of natural sounds in songbirds. J Neurophysiol 96(1):252–258.

[17] Paiva ARC, Park I, Príncipe JC. 2009. A reproducing kernel Hilbert space framework for spike train signal processing. *Neural Computation* 21:424–449.

[18] Schrauwen B, Van Campenhout J. 2007 Linking non-binned spike train kernels to several existing spike train metrics. *Neurocomputing* 70:1247–1253.

[19] Sellers, PH. 1974 On the theory and computation of evolutionary distances. *SIAM Journal on Applied Mathematics* 26:787–793.

[20] van Rossum MCW. A novel spike distance. Neural Comput, 2001;13:751-63.

[21] Victor JD, Purpura KP. Nature and precision of temporal coding in visual cortex: A metric-space analysis. J Neurophysiol, 1996;76:1310-26.

[22] Victor JD, Purpura KP, Gardner D. Dynamic programming algorithms for comparing multineuronal spike trains via cost-based metrics and alignments. J Neurosci Methods, 2007;161:351-60.

[23] Wang L, Narayan R, Graña G, Shamir M, Sen K (2007) Cortical discrimination of complex natural stimuli: Can single neurons match behavior? J Neurosci 27(3):582–589.