Spike train metrics
Jonathan D Victor

Quantifying similarity and dissimilarity of spike trains is an important requisite for understanding neural codes. Spike metrics constitute a class of approaches to this problem. In contrast to most signal-processing methods, spike metrics operate on time series of all-or-none events, and are, thus, particularly appropriate for extracellularly recorded neural signals. The spike metric approach can be extended to multineuronal recordings, mitigating the ‘curse of dimensionality’ typically associated with analyses of multivariate data. Spike metrics have been usefully applied to the analysis of neural coding in a variety of systems, including vision, audition, olfaction, taste and electric sense.

Introduction

Variability is a prominent feature of neural activity and its sources and functional implications are the focus of much investigation. Variability places limits on the reliability of signals, but can also provide a rich language for neuronal populations and their interactions. To analyze variability, one must first quantify the extent to which two patterns of neural activity are dis-similar, that is, one needs a ‘metric’ for comparing patterns. Metrics that are specifically applicable to spike trains (see glossary) are the focus of the present review.

Although laboratory measurements never correspond to a mathematical ideal [1], the choice of an appropriate mathematical framework is a prerequisite for rigorous data analysis. This choice is particularly crucial to the study of neural coding, because neural coding is fundamentally an abstraction: the relationship between stimuli, actions, and/or behavioral states, and the activity of one or more neurons.

A voltage record, the starting point for most neurophysiologic analyses of neural activity, is typically acquired as a set of closely spaced digital measurements. It seems natural to conceptualize a voltage record as a continuous function of time. However, we might also choose to conceptualize a neurophysiologic voltage record as a sequence of stereotyped events. This viewpoint of a voltage record as a realization of a ‘point process’ (see glossary) is especially appropriate for extracellularly recorded neural signals, given the all-or-none nature of the action potential and its effects on postsynaptic neurons [2–4]. The point process viewpoint (discrete events in continuous time) is intermediate between the vector space viewpoint (continuous time, continuous events) and a third alternative, symbol sequences (discrete time, discrete events). The symbol sequence viewpoint is often used in application of information theory and nonlinear dynamics to neuronal data [5–8]. The choice of viewpoint can influence the results of data analysis, both quantitatively and qualitatively. These considerations are fundamental to the rationale for the spike metric approach.

As detailed in Box 1, the continuous-record viewpoint leads naturally to metrics that have a particular mathematical structure (Euclidean distances [see glossary] within a vector space). Given the prominence of nonlinearities at many stages of neural processing and the kinds of ‘spaces’ (see glossary) that neural activity ultimately must represent, this restriction might not be desirable. Metrics based on the point process viewpoint do not have these constraints.

This review describes several kinds of metrics that can be applied to point processes in general, but that are motivated by neurophysiologic considerations specific to series of action potentials (‘spike trains’). I consider metrics applicable to single-unit and multineuronal activity and recent examples of their use. Computational considerations are summarized in Box 2.

A simple example

To describe the nature of metrics and metric spaces, we use a simple example, consisting of a set of cities A, B, . . . linked by commercial airlines. We begin with a table of airfares that provides a list of costs for travel between various pairs of cities. We assume that fares between any two cities are independent of the direction of travel, that is, the costs are ‘symmetric’. The table of pairwise costs does not constitute a metric, because the cost c(A, B) of a flight from A to B might be greater than the sum of the cost of a flight from A to X, and from X to B. However, we can turn the table of costs into a metric by defining the
metric distance between two cities as the least total cost of transportation between them, making use of however many stopovers are necessary. More formally, we define the metric distance \( d(A, B) \) between two points \( A \) and \( B \) as

\[
d(A, B) = \min \left\{ \sum_{j=0}^{n-1} c(X_j, X_{j+1}) \right\}
\]

where \( \{X_0, X_1, \ldots, X_n\} \) is an itinerary from \( A \) to \( B \), with \( X_0 = A \) and \( X_n = B \). This definition guarantees that the metric distances between three points \( A, B, \) and \( C \) satisfy the triangle inequality \( d(A, C) \leq d(A, B) + d(B, C) \), because the cheapest itinerary from \( A \) to \( C \) cannot be more expensive than an itinerary that is constrained to stop at \( B \). It is easy to construct examples of metrics that violate the rules of Euclidean geometry.

Cost-based metrics for spike trains

With some fine print, Equation 1 provides a way to turn any set of (symmetric) costs into a metric. In the present context, each spike train \( A, B, \ldots, \) is a point in the metric space (i.e. a ‘city’). The sequence of steps corresponding to each term on the right side of Equation 1 constitutes a path of elementary steps that transforms a spike train \( A \) into a spike train \( B \). In spike train metrics, the costs \( c(X_j, X_{j+1}) \) of the elementary steps are determined by a small number of rules [9,10] that are intended to capture basic biological features of how spike trains influence postsynaptic neurons.

**Spine time metrics**

The ‘spike time’ family of metrics is motivated by a caricature of a neuron as a coincidence detector [11–14]. For these metrics, deleting or inserting a spike has a cost of 1. This rule sets an overall scale for the metrics and ensures that any spike train can be transformed to any other spike train by a path: the path that successively deletes all spikes from train \( A \), and then successively inserts all spikes into train \( B \). The second rule, which confers sensitivity to spike timing, is that the cost of moving a single spike in time is proportional to the amount of time that it is moved. That is, if two spike trains \( A \) and \( B \) are identical except for a single spike that occurs at \( t_a \) in \( A \), and \( t_b \) in \( B \), then \( c(A, B) = q|t_a - t_b| \), where \( q \) (in units of 1/sec) is a parameter that determines the relative sensitivity of the metric to spike count and spike timing.
Box 2 Computational considerations.

Computation of the metrics that correspond to Euclidean distances [25,26] is straightforward: spike trains are convolved with a smoothing kernel, and the scalar product of the resulting time series is calculated. The computational burden is proportional to the duration of the spike trains. The cost-based metrics $D_{\text{time}}^q$ and $D_{\text{interval}}^q$ are similar to ‘alignments’, ‘edit-length distances’ and ‘string-matching procedures’ used for comparison of genetic sequences [58] and electroencephalogram (EEG) analysis [59], in which there is a cost associated with insertions, deletions and shifts. This kinship provides an efficient dynamic programming algorithm to calculate these metrics [9,10]. In particular, efficient calculation of the minimal cost sequence is enabled by the following observations concerning the elementary steps in a minimal-cost sequence. First, if the last spikes in the two trains are neither inserted nor deleted, then they must be connected by a shift. Second, the paths taken by two spikes cannot cross. Third, no spike needs to be shifted more than once. The computational burden of the algorithm is proportional to the product of the number of spikes in the trains to be compared.

For the multineuronal metric $D_{\text{multispike}}^q$, a clever algorithm in which the two spike trains are treated asymmetrically has been devised [60], which has a computational burden proportional to $N^{L+1}$, where $N$ is the number of spikes in a typical train, and $L$ is the number of labels (neurons).

Other cost-based metrics can be envisioned [9] (e.g. enabling both single-spike shifts as in $D_{\text{time}}^q$ and interval changes as in $D_{\text{interval}}^q$). However, no efficient computational algorithm for such metrics is available at present.

These rules, along with equation 1, suffice to provide a metric distance between all spike trains, denoted $D_{\text{spike}}^q$. Intuitively, the distance between two spike trains is the minimum total cost of a set of elementary steps that transforms one spike train into the other (Figure 1). When $q$ is very small, then the times of individual spikes have little influence on the calculated distance between spike trains. For $q = 0$, then spike timing is irrelevant, in that spikes can be shifted in time ‘for free’. Thus, for $q = 0$, a minimal path between spike trains $A$ and $B$ consists of deleting or inserting enough spikes into $A$ so that the total count matches that of $B$, and then shifting the spikes in time so that they match. That is, $D_{\text{spike}}^0[0](A,B) = |n(A) - n(B)|$, where $n(X)$ denotes the number of spikes in the train $X$. $D_{\text{spike}}^0[0]$ can, thus, be thought of as formalizing a ‘spike count’ code: spike trains are considered different only if they contain a different number of spikes.

As $q$ increases, the metric $D_{\text{spike}}^q$ becomes increasingly sensitive to spike timing. To see this, consider spike trains $A$ and $B$ that each contain only one spike, at $t_a$ in $A$, and $t_b$ in $B$. There are two paths to consider in Equation 1: first, deleting the spike in $A$ and inserting it into $B$; second, shifting the spike from $t_a$ to $t_b$. The first path has a cost of $2$; the second path has a cost of $|t_a - t_b|$. The distance $D_{\text{spike}}^q(A,B)$ is the minimum of these two numbers. The break-even point is for $|t_a - t_b| = 2/q$. That is, in the metric $D_{\text{spike}}^q$, two spikes are only considered as comparable if they occur within an interval of $2/q$ sec. Depending on whether the spike times in two trains $A$ and $B$ are similar, $D_{\text{spike}}^q(A,B)$ can range from $|n(A) - n(B)|$ (spike times in $A$ and $B$ match, no shifting of spike times needed) to $n(A) + n(B)$ (no spike times in $A$ and $B$ are within $2/q$ sec; spikes must be deleted from $A$ and then reinserted into $B$). Thus, $q$ explicitly represents
the relative importance of spike times and spike counts: a change in the time of a spike by $1/q$ sec influences the total cost of a path as much as deleting the spike altogether. For neurons that act as a coincidence detector with integration time (or temporal resolution) $1/q$, spike trains will have similar postsynaptic effects if they are similar in the sense quantified by $D_{\text{hyth}}[q]$. Often, temporal resolution is not known in advance — so, $q$ is retained as a parameter, with the goal of using the dependence on $q$ to analyze coding (see below and Box 3).

**Spike interval metrics**

For the ‘spike interval’ metrics [9,10], denoted $D_{\text{inter}}[q]$, the heuristic is that the postsynaptic effects of a spike might depend strongly on the recent activity at that synapse [4,15,16]. Correspondingly, the temporal dependence of $D_{\text{inter}}[q]$ is based on the intervals between spikes, rather than their absolute times. For $D_{\text{inter}}[q]$, the cost of insertion or deletion of an interspike interval is 1. The second rule is that shortening or extending an interspike interval by an amount $t$ has a cost $qt$. Note that changing the length of an interval differs from changing the time of a spike, in that when an interval length is changed, the time of all later spikes are also changed. For this reason, $D_{\text{inter}}[q]$ and $D_{\text{hyth}}[q]$ have fundamentally different topological characteristics [9].

**Multineuronal cost-based metrics**

Multineuronal recordings that enable the study of patterns of activity across neurons are increasingly available [17–19] and methods to analyze such data effectively are receiving increasing attention [20,21]. A multineuronal recording might be considered to be a sequence of labeled events. In this view, cost-based metrics are readily extended to the multineuronal context. The simplest approach [22] is to add a rule that sets the cost of changing the label associated with an event. This results in a two-parameter family of metrics, $D_{\text{hyth}}[q,k]$, where $k$ is a parameter that sets the cost of changing a label. $k = 0$ corresponds to the notion that the neuron of origin of a spike is irrelevant (because there is no cost associated with changing this label). $k = 2$ corresponds to a ‘labeled line’ interpretation, because changing the neuron of origin of a spike has the same cost as deleting a spike associated with one neuron, and inserting a spike associated with another neuron.

**Other metrics**

Within the cost-based framework, one can also construct metrics sensitive to motifs of spikes [23] (by having a rule associated with the cost of moving a set of spikes), and metrics that combine the rules of $D_{\text{hyth}}[q]$ and $D_{\text{inter}}[q]$. Metrics on spike trains can also be obtained by binning [24] or convolving them with a smoothing kernel (see glossary) [25,26], and then using standard vector-space distances between the derived temporal functions. These latter approaches necessarily lead to Euclidean distances.

**Applications**

Spike metrics have been applied to data obtained from a variety of neural systems, to quantify variability per se, [27,28,29,30,31], to characterize neural coding (at the level of temporal coding within single neurons [9,10,32–36] or neuronal pairs [22,37]), and to evaluate models [38,39].

**Electrosensory**

Kreiman and co-workers [27] studied the variability of the P-receptor afferents [40] in the weakly electric fish *Eigenmannia*, whose discharges are loosely phase-locked to the periodic (200–600 Hz) discharge of its electrosensory organ. They made extensive use of measures of variability based on $D_{\text{hyth}}$, because these, but not measures based on spike count or its variance, appeared to capture trial-to-trial variability in P-receptor activity. They found that the intrinsic variability of spike trains was not likely to degrade information transmission, but enabled improvement in coding by averaging across multiple afferent fibers.

**Vision**

*Blowfly H1*

Grewe et al. [28] used $D_{\text{hyth}}$, along with a Euclidean metric, to examine variability in the responses of the blowfly wide-field, motion-sensitive neuron H1, driven by motion stimuli with various levels of added noise. By determining the ‘equivalent noise’ (the maximum...
amount of added noise that enabled a criterion level of response classification), they deduced that internal noise, rather than photon noise, limited the performance of the H1 neuron.

**Mammalian retina**
Chichilnisky and Rieke [29] used $D^{\text{spike}}$ to analyze near-threshold signaling in rod photoreceptors and retinal ganglion cells of the tiger salamander. Although grouping of responses into signal versus no-signal clusters was no more accurate than could be achieved by standard methods such as the Fisher discriminant (see glossary) [41], the observation that optimal grouping was achieved for $D^{\text{spike}}[q]$ at $q = 0.1$ indicated a meaningful temporal resolution of 100 ms.

**Lateral geniculate nucleus**
In response to full-field random flicker, retinal and lateral geniculate neurons often fire in discrete ‘firing events’ consisting of several spikes, at times that are reproducible across trials [30,38]. Reinagel and Reid [30] used $D^{\text{spike}}$ to show that other than a possible small overall latency shift, these timing events were conserved not only across trials, but also across animals. Keat et al. [38] later used $D^{\text{spike}}$ to evaluate the ability of models to predict these firing events.

**Visual cortex**
Application of spike metrics to single-unit and multi-unit (single units not separated) recordings in primary visual cortex (V1) and early extrastriate cortex (V2 and V3) of the awake macaque [10] revealed that the temporal structure of spike trains contributed significantly to coding of visual information. Because information estimates based on $D^{\text{spike}}$ were generally larger than those based on $D^{\text{interval}}$, it was concluded that spike timing (relative to stimulus onset), rather than the interval structure, was the more crucial aspect of temporal structure. Moreover, because information estimates for the several stimulus attributes (contrast, spatial frequency, orientation, size, and texture) had distinctive dependences on $q$, it was concluded that typical neurons multiplexed visual information, representing contrast with high temporal precision (ca. 10–30 ms), and spatial aspects with lower temporal precision (ca. 100 ms).

In V1 of the anesthetized macaque, Reich et al. [32] used $D^{\text{spike}}$ to show that most of the information about contrast could be extracted from the latency of the first spike, although additional information could be extracted from the temporal structure of the response without regard to latency, and that temporal coding was particularly important at higher contrasts, at which the spike rate response neared saturation. Mechler et al. [33] used a variant of $D^{\text{spike}}$ appropriate for responses to periodic stimuli to demonstrate that temporal structure played a much larger role in the coding of edges (square-wave gratings) than of smooth variations (sine gratings). Cyclic variants of $D^{\text{spike}}$ were later used [34] to characterize the coding of edge-like, line-like, and intermediate one-dimensional features in V1, and showed that many neurons demonstrated feature opponency and/or feature selectivity for compound gratings.

Two studies focused on the role of the activity pattern within a local cluster of cortical neurons. Aronov et al. [22] used $D^{\text{spike}}[q,k]$ to characterize the coding of spatial phase across pairs of V1 neurons, isolated using tetrodes (four-element microelectrodes) [17]. Dependence of information estimates on $q$ indicated that spike times had an informative precision of ca. 30 ms; dependence of information estimates on $k$ indicated that the neuron of origin, and not just the total activity of the local population, contributed to coding of spatial phase. The geometry of the stimulus set (the circle of spatial phase) corresponded to the response similarities as determined by $D^{\text{spike}}[q,k]$. Greater fidelity of the representation was achieved for nonzero values of $k$, thus demonstrating that within a local cluster, the neuron of origin of a spike, in addition to its timing, carries information. This study also introduced a technique (designated ‘temporal profiles’) to identify the time course of temporal features that are crucial to stimulus representation. Samonds, Bonds and co-workers [37,42] examined signaling of orientation in cat primary visual cortex with $D^{\text{spike}}$ and $D^{\text{interval}}$. Although large angular differences were adequately represented by firing rate, type analysis [43] of responses of neuronal pairs suggested that cooperative signaling was present for small angular differences [42]. Metric-space analysis [37] demonstrated that orientation differences of 10 deg or less were signaled by the temporal fine structure (2–10 ms) of spike times and spike intervals.

**Audition**
Middlebrooks and co-workers [44] showed that responses of single neurons in cat ectosylvian gyrus, when analyzed with a neural-network approach, represent the azimuth of a sound stimulus in a panoramic (360 deg) fashion. A reanalysis of these data using $D^{\text{spike}}$ and $D^{\text{interval}}$ showed that the representation relied on spike timing, at a resolution of ca. 4 ms. A comparable conclusion was also reached in a neural-network analysis of A2 neurons [45] using surrogate-data methods.

Machens et al. [36] showed that 100 ms samples of natural songs from up to eight different conspecific grasshoppers could be distinguished by spike trains of single auditory neurons via $D^{\text{spike}}[q]$ at $q = 100$, corresponding to an informative precision of 10 ms.

**Chemical senses**
Laurent and co-workers [31] identified a population of neurons in the olfactory system of the locust, the β-lobe neurons, that are crucial to reading out the output of the
mushroom body, a structure involved in odor learning. Desynchronization of projection neurons (PNs), two synapses upstream of these β-lobe cells, could be achieved by blocking fast GABA inhibition with picROTOxin. This leads to loss of behavioral discrimination of similar odors, although coarse odor discrimination remains intact [46]. $D^{\text{pike}}$ was used [31] to show that information about odors contained in the spike trains of β-lobe cells is lost when the PNs are desynchronized, even though no such loss is observed within spike trains of individual PNs. These elegant experiments demonstrate the functional relevance of neuronal synchronization.

Di Lorenzo and Victor [35] applied $D^{\text{pike}}$ and $D^{\text{interval}}$ to the analysis of gustatory coding in single-neuron responses in the nucleus of the solitary tract of the rat. In 10 of 19 neurons, the temporal structure of the initial 2 s of the response contributed to the coding of the four basic taste qualities. The informative precision of spike timing was much lower than in the other studies reviewed here, typically about 300 ms, and response dynamics contributed the most to coding in neurons that were the most variable, in terms of their overall firing rate. As in the visual cortex studies described above [10], analysis of surrogate datasets obtained by shuffling spikes across trials served to demonstrate that the informative aspects of spike timing went beyond that of a Poisson sampling of a firing rate envelope.

**Motor**

In preliminary work [47] $D^{\text{pike}}$ was used to identify aspects of single-neuron activity in parietal cortex of the macaque that were correlated with arm approach and grasp style. Of note, this re-analysis made use of a public database and analysis toolkit.

**Conclusions**

Spike metrics are applicable to data that can be viewed as discrete events in continuous time. Cost-based spike metrics are a general strategy for formalizing biologically motivated notions of distance, and, thus, constitute a principled approach for the analysis of variability of single- and multineuronal extracellular recordings. They have been fruitfully applied in a variety of neural systems, to characterize and consequently help to understand neuronal variability and coding. Analyses in several sensory systems have shown that spike count might suffice for signaling gross sensory differences, but spike timing is important for signaling subtle differences [31,32,37**].

By parameterizing biologically motivated notions of distances between time series of events, metric-space methods complement non-parametric approaches to estimate information-theoretic quantities from limited data. Spike metrics are related to edit-length distances (see glossary) used for genetic analysis, and present related computational challenges.

**Acknowledgements**

This work is supported in part by National Eye Institute Grant 1RO1 EY9314 to J Victor and a National Institute of Mental Health Grant 1RO1 MH68012 to D Gardner. The author thanks K Purpura for helpful comments on the manuscript.

**References and recommended reading**

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- **of outstanding interest**

29. Chichilnisky EJ, Rieke F: Detection sensitivity and temporal resolution of visual signals near absolute threshold in the salamander retina. J Neurosci 2005, 25:318-330. The authors used D\textsubscript{now} to decode ganglion cell responses, and to deduce that spike times, at a resolution of 100 ms, contributed to the detection of near-threshold signals.
37. Samonds JM, Bonds AB: From another angle: differences in cortical coding between fine and coarse discrimination of orientation. J Neurophysiol 2004, 91:1193-1202. The authors used metric-space analysis to show that neurons in primary visual cortex use firing rate to signal large orientation differences, whereas the fine temporal structure of their responses signal orientation differences of 10 deg or less.


