A comment on 'A fast $\mathcal{L}_p$ spike alignment metric' by A. J. Dubbs, B. A. Seiler and M. O. Magnasco [arXiv:0907.3137]

Conor Houghton[*]

School of Mathematics, Trinity College Dublin, Ireland

27 July 2009

Abstract

Measuring the transmitted information in metric-based clustering has become something of a standard test for the performance of a spike train metric. In this comment, the recently proposed $\mathcal{L}_p$ Victor-Purpura metric is used to cluster spiking responses to zebra finch songs, recorded from field L of anesthetized zebra finch. It is found that for these data the $\mathcal{L}_p$ metrics with $p > 1$ modestly outperform the standard, $p = 1$, Victor-Purpura metric. It is argued that this is because for larger values of p, the metric comes closer to performing windowed coincidence detection.

1 Introduction

There are a number of reasons why metrics are believed to be useful for studying spike trains. It can be argued that it is the most general, useful, mathematical framework for spike trains [Victor and Purpura 1997], it may be possible to find a manifold of spike trains using local linear methods and,

---

[*]houghton@maths.tcd.ie

of most relevance to the present discussion, it is possible that studying spike train metrics is a useful approach to understanding how content is coded in spike trains.

The Victor-Purpura metric is an edit distance metric on the space of spike trains. The distance between two spike trains is, in effect, calculated as the cost of changing one spike train into the other by adding, deleting, or moving spikes, with an individual cost for each type of elementary moves. In particular, a cost of one is associated with adding or deleting a spike and a cost of $q\delta t$ with moving a spike a time δt: the distance is

$$d(\mathbf{u}, \mathbf{v}; q) = \min_\gamma c_{\gamma;q}(\mathbf{u}, \mathbf{v}) \qquad (1)$$

where $c_\gamma(\mathbf{u}, \mathbf{v})$ is the cost of a sequence of elementary moves γ and is calculated by adding the cost of all the elementary moves in the sequence. The minimum is taken over all sequences γ changing $\mathbf{u}$ to $\mathbf{v}$. The parameter $2/q$ gives an important timescale for the metric: it is never worthwhile to move a spike more than $2/q$ since it would be cheaper to delete the spike from one temporal location and to add it to the other. This gives a timescale that, roughly speaking, separates jitter from unreliability in comparing spikes in the spike trains. However, the Victor-Purpura metric does more than this; it explicitly pairs up those spikes that can be thought of as being related by jitter.

The Victor-Purpura metric has an l^1 character: the cost c_γ is a simple linear sum of the individual costs of the individual moves. The generalization proposed in [Dubbs *et al.* 2009] changes this to an l^p-like sum. For notational convenience, from now on the set of sequences will be restricted: the minimum sequence will never involve moving a spike after it has been added, it would always be cheaper to add the spike in the correct location, similarly, spikes are never moved before they are deleted. It is also specified that spikes can only be moved once. With these restrictions, the sequence γ can be considered to be an unordered set made up of deletions, additions and moves. Let

$$\gamma = \alpha \cup \mu \qquad (2)$$

where α is the set of additions and deletions and μ is the set of moves. The elements of μ are pairs of spikes (u, v), with $u \in \mathbf{u}$ and $v \in \mathbf{v}$, that are related by jitter.

Now,

$$c_{\gamma;q}(\mathbf{u}, \mathbf{v}) = |\alpha| + \sum_{(u,v) \in \mu} q|u - v| \qquad (3)$$

2

This is generalized in [Dubbs *et al.* 2009] to

$$c_{\gamma;q,p}(\mathbf{u}, \mathbf{v}) = \left(|\alpha| + \sum_{(u,v) \in \mu} q^p |u - v|^p \right)^{1/p} \qquad (4)$$

with $p \geq 1$, to give the $\mathcal{L}_p$ Victor-Purpura metric

$$d(\mathbf{u}, \mathbf{v}; q, p) = \min_{\gamma} c_{\gamma;q,p}(\mathbf{u}, \mathbf{v}) \qquad (5)$$

where the minimum is taken over sequences γ changing $\mathbf{u}$ to $\mathbf{v}$ and satisfying the restrictions specified above. An algorithm for calculating this quantity is given in [Dubbs *et al.* 2009]; the existing algorithms used to calculate the Victor-Purpura metric, [Victor *et al.* 2009], for example, could also be adapted to $p > 1$.

2   Evaluating the metric using clustering

In the data that will be considered here, spike trains were recorded from field L of the auditory fore-brain of anesthetized zebra finch during playback of 20 con-specific songs with each song repeated ten times, to give a total of 200 spike trains. These spike trains, and the experimental conditions used to produce them, are described in [Narayan *et al.* 2006, Wang *et al.* 2007] and they have previously been used to compare metrics in [Houghton 2009, Houghton and Victor 2009]. The key point, in this context, is that a good metric, one which depends on the content encoded in the spike train, should measure a smaller distance between spiking responses to the same song than between responses to different songs. One motivation for looking at spike train metrics is that scoring metrics in this way, and examining how good metrics measure distance, might reveal details about how content is encoded.

The usual method for evaluating how well a metric succeeds in this way is to calculate the transmitted information [Victor and Purpura 1996]. There are two ways to cluster the spike trains: a true clustering, based on the identity of the song that elicited the response and an estimated clustering, based on the metric distances between the responses. Roughly speaking, the transmitted information quantifies the amount of information about the true clustering given by the estimated clustering.

Calculating the transmitted information relies on the calculation of a confusion matrix N. This is described in detail in [Victor and Purpura 1996],

3

described specifically for the data considered here, N is a 20×20 matrix of integers where the entry N_{ij} counts how many of the spike trains from the ith true cluster is closest, on average, to the spike trains in the jth true cluster. As in [Victor and Purpura 1996] this averaging is carried out as a root-mean-square averaging to under-weigh outliers. If the metric clustering is close to the true clustering, N will be nearly diagonal. The transmitted information gives a measure of this, it is

$$\tilde{h} = \frac{1}{n \ln 20} \sum_{ij} N_{ij} \left(\ln N_{ij} - \ln \sum_k N_{kj} - \ln \sum_k N_{ik} + \ln n \right). \qquad (6)$$

where $n = \sum_i \sum_j N_{ij}$. All sums are from one to 20: the number of clusters. The factor of $\ln 20$ normalizes $\tilde{h}$ and so that it takes values between zero and one; a low value corresponds to poor clustering and a value near one corresponds to a near-diagonal confusion matrix.

The main purpose of this comment is to evaluate the $\mathcal{L}_p$ Victor-Purpura metric by calculating $\tilde{h}$ for the zebra finch data previous used to evaluate other spike train metrics [Houghton 2009, Houghton and Victor 2009]. The data set contains 24 sets of spike trains, corresponding to 24 recording sites from multiple birds. Although the songs themselves are of different lengths, all of the songs are at least a second long and, in each case, the spike train is truncated to one second, starting at song onset.

The performance of the $\mathcal{L}_p$ Victor-Purpura metric is illustrated in Fig. 1. A remarkable feature is that $\tilde{h}$ rarely decreases as p is increased. Other multi-parameter metrics show a mixed effect; although changing a parameter might increase the average $\tilde{h}$ across the sites, there will typically be many individual sites where $\tilde{h}$ decreases. In the case of p, while the average improvement as p increases is modest and a small number of sites show no change; there is no site where the clustering is negatively affected when p is changed from one to two, three do show a tiny decrease between $p = 1$ and $p = 1.5$. $\tilde{h}$ increases for 19 sites between $p = 1$ and $p = 2$ and for 15 between $p = 2$ and $p = 3$, the improvement gets less and less and appears to have plateaued by $p = 10$. The average performance is plotted against p in Fig. p. The optimal value of q does not vary significantly as p changes; however, the range of q that produces the optimal performance does get wider.
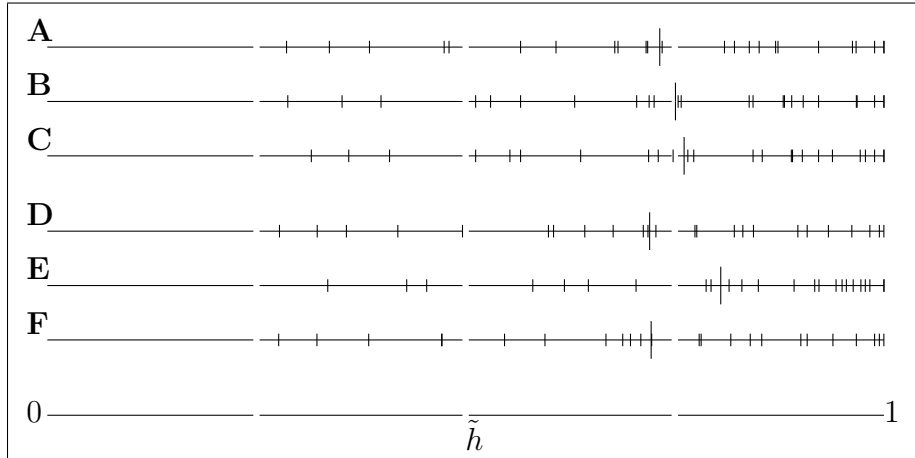
4

Figure 1: Performance as p changes. **A**, **B** and **C** correspond to $p = 1$, $p = 2$ and $p = 10$. Each horizontal line runs from zero to one. For clarity, there are tiny gaps at 0.25, 0.5 and 0.75. Each small vertical dash marks the $\tilde{h}$ value for one site where for each site the optimal value of q is used. The large vertical dash marks the average, 0.732, 0.751 and 0.761 respectively. To indicate the range over which $\tilde{h}$ values vary, two other metrics are plotted for comparison. **D** corresponds to the van Rossum metric [van Rossum 2001], with each site using its optimal value of τ. **E** corresponds to the synapse metric, with each site using its optimal value of τ and μ [Houghton 2009]. Finally, **F** corresponds to the L^1 van Rossum metric, with a boxcar filter; this metric measures a very similar distance to the $p = 1$ metric [Houghton and Sen 2008].
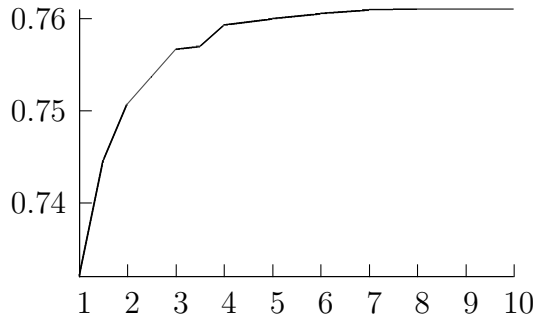
5

Figure 2: The average performance plotted against p. $\tilde{h}$ has been calculated for a range of p values using the optimal value of q for each site and the average value of $\tilde{h}$ is plotted against p.

3 Discussion

It is interesting to compare this metric to another generalization of the Victor-Purpura metric [Victor and Purpura 1996, Victor and Purpura 1997]. The linear distance function $q\delta t$ can be replaced by any convex function, $f(|\delta t|)$, of distance:

$$c_{\gamma;f(\cdot)}(\mathbf{u}, \mathbf{v}) = |\alpha| + \sum_{(u,v)\in\mu} f(|u - v|). \tag{7}$$

The distance function $f(|\delta t|)$ must be convex to ensure that the triangular inequality is satisfied. Now, the function $q^p \delta t^p$ in $c_{\gamma;q,p}$ above is concave for $p > 1$. However, because of the $1/p$ exponent in $d(\mathbf{u}, \mathbf{v}; q, p)$, this does not cause the same difficulty with the triangular inequality: an interesting property of the $\mathcal{L}_p$ metrics is that they satisfy the triangular inequality despite having a concave distance function.

As p is increased, $q^p|\delta t|^p$ becomes increasingly concave and spikes located at similar times contribute less and less to the total cost and this cost is predominately made up of the cost of adding and deleting spikes that are not related by jitter. This means that the metric performs more and more as a windowed coincidence detector as p increases. The fact that it is always beneficial to increase p appears to imply that, for these data, coincidence and coincidence detection has a role in the encoding and decoding of content in spike trains.
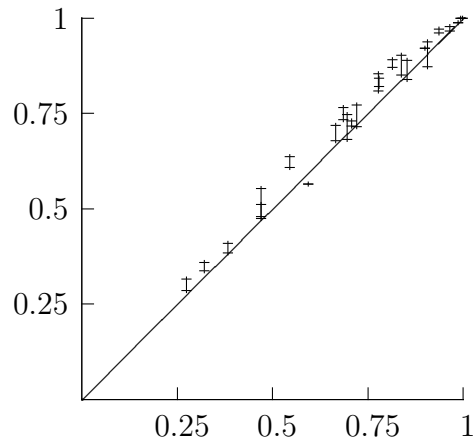
6

Figure 3: Comparing the L^1 boxcar van Rossum metric with the $p = 1$ and $p = 10$ $\mathcal{L}_p$ metrics. Here $\tilde{h}$ values for the $p = 1$ and $p = 10$ metrics are graphed against the values for the L^1 van Rossum metric with a boxcar filter with optimal values of the q or τ parameter has been used for each site. The $p = 1$ and $p = 10$ values are marked by a horizontal line, the two values for a given site are joined by a vertical line. Since increasing p never decreases $\tilde{h}$ between $p = 1$ and $p = 10$, the topmost of the two horizontal lines corresponds to $p = 10$. The 'x=y' line is also plotted for clarity.

The pairing of near coincident spikes is the key distinction between the Victor-Purpura metrics and the van Rossum metrics, a family of spike train metrics which are calculated by comparing reconstructed rate functions. In fact, computationally, an L^1 van Rossum metric with a boxcar filter measures a very similar distance to the Victor-Purpura metric [Houghton and Sen 2008] and, as seen in Fig. 3, has a very similar performance. The way in which $\mathcal{L}_p$ Victor-Purpura metric performance increased with p seems to indicate that the significant temporal structure of spikes is not fully accounted for by the temporal structure of a spike rate.

Acknowledgments

References

[Dubbs et al. 2009] A. J. Dubbs, B. A. Seiler, and M. O. Magnasco. A fast $\mathcal{L}_p$ spike alignment metric. (2009, preprint, [arXiv:0907.3137])

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

[Houghton and Sen 2008] C. Houghton and K. Sen. A new multi-neuron spike-train metric. *Neural Computation* 20:1495–1511, 2008.

[Houghton and Victor 2009] C. Houghton and J. D. Victor. Spike codes and spike metrics. (2009, book chapter, under review).

[Narayan et al. 2006] R. Narayan, G. Graña, and K. Sen. Distinct time scales in cortical discrimination of natural sounds in songbirds. *Journal of Neurophysiology*, 96:252–258, 2006.

[van Rossum 2001] M. van Rossum. A novel spike distance. *Neural Computation*, 13:751–763, 2001.

[Victor *et al.* 2009] J. D. Victor, D. Goldberg, and D. Gardner. Dynamic programming algorithms for comparing multineuronal spike trains via cost-based metrics and alignments. *Journal of Neuroscience Methods*, 161:351–360, 2007.

[Victor and Purpura 1996] J. D. Victor and K. P. Purpura. Nature and precision of temporal coding in visual cortex: a metric-space analysis. *Journal of Neurophysiology*, 76:1310–1326, 1996.

[Victor and Purpura 1997] J. D. Victor and K. P. Purpura. Metric-space analysis of spike trains: theory, algorithms and application. *Network: Computation in Neural Systems*, 8:127–164, 1997.

[Wang *et al.* 2007] L. Wang, R. Narayan, G. Graña, M. Shamir, and K. Sen. Cortical discrimination of complex natural stimuli: can single neurons match behavior? *Journal of Neuroscience*, 27:582–9, 2007.