Latent State Visualization of Neural Firing Rates

Austin J. Brockmeier, *Student Member, IEEE*, Evan G. Kriminger,
Justin C. Sanchez, *Member, IEEE* and José C. Príncipe, *Fellow, IEEE*

Abstract— **Visualizing the collective modulation of multiple neurons during a known behavioral task is useful for exploratory analysis, but handling the large dimensionality of neural recordings is challenging. We further investigate using static dimensionality reduction techniques on neural firing rate data during an arm movement task. This lower-dimensional representation of the data is able to capture the neural states corresponding to different portions of the behavior task. A simulation using a dynamical model lends credence to the ability of the technique to generate a representation that preserves underlying dynamics of the model. This technique is a straightforward way to extract a useful visualization for neural recordings during brain-machine interface tasks. Meaningful visualization confirms underlying structure in data, which can be captured with parametric modeling.**

I. INTRODUCTION

Natural movement reconstruction from neural recordings require well-trained decoding algorithms. Motor brain-machine interfaces have motivated algorithms to decode upper-limb movements from neural data [1], [2]. Invasive brain-machine interfaces use the firing rate of multiple individual neurons as the inputs and translate the signals into robotic or cursor movements. Often, natural movement reconstruction is a prerequisite to online operation, but independent of a given parametric model it is unclear how well the natural movement can be reconstructed from given data.

Without a well-trained parametric model it is difficult to quantify the relationship between the joint neural activity—a high-dimensional variable—and the movement—a 3-dimensional variable—especially on a trial-by-trial basis. Any movement-correlated modulation is often cloaked by both the sheer dimensionality of the data and the intrinsic noise of neural spiking. For the case of repeated time-locked trials, the average rate (or variance) of individual or pairs of neurons can be shown in the form of a peri-event time histogram, but the movement trajectories may vary widely between trials even for repeated tasks [3]. Without repeated trials, it is still possible to compute correlation between the individual movement dimensions and the firing rate over time. These methods are not defined for the joint

A. J. Brockmeier, E. G. Kriminger, and J. C. Príncipe are with the Department of Electrical and Computer Engineering, University of Florida, P.O. Box 116130 NEB 486, Bldg #33, University of Florida, Gainesville, FL 32611 USA (e-mail: {ajbrockmeier, kriminger, principe}@cnel.ufl.edu)

J. C. Sanchez is with the Department of Biomedical Engineering, University of Miami, Coral Gables, FL 33146 USA (e-mail: jcsanchez@miami.edu)

activity of more than 2 neurons. Probabilistic model-based approaches can be used to account for the noisiness and high dimensionality [4], [5], but for initial analysis we desire a simpler approach without estimation and explicit model formulation. A review of the problems and some suggested techniques are found in [3].

For exploratory analysis we propose using low-dimensional visualization of the multi-channel neural modulation during natural movement. As the underlying movement is in three-dimensions, it is a reasonable assumption that any relevant dynamics can be preserved in a visualization with equal or smaller dimension. A useful visualization would preserve any association between the neural modulation and the movement trajectories, i.e. whether nearby areas in the visualization correspond to similar portions of movement; thus, the location in the visualization can serve as a predictor of the position during movement.

We form the low-dimensional visualization by using dimensionality reduction on the original neural responses. We can consider the visualization to be a *latent space*, i.e. a low-dimensional space that preserves the underlying dynamics of the original space. Indeed, the static dimensionality reduction technique local linear embedding (LLE) [6] has proven useful for preserving neural trajectories in other neuroscience analyses [7], [8]. To form the visualization we use stochastic neighborhood embedding, *t-SNE* [9]. The *t-SNE* algorithm has proven useful for visualizing the aspects of high-dimension datasets such as clusters and manifolds.

In this work, we use *t-SNE* to produce a visualization of neural data and analyze how well the latent space preserves known external dynamics. Although visualization is inherently qualitative, we also use a quantitative measure of association to assess how well the visualization alone can be used to predict the movement. We test it on a simulated dynamical system and on real neural data recorded during a reach and grab task. In both cases, the method is able to preserve the association between portions of the latent space and the external variables while reducing the space to 2 dimensions. Thus, this method explores a method to visualize the joint neural dynamics during natural movement or brain-machine interface tasks, and can be used to gauge the neural modulation before explicit modeling.

II. METHOD

Dimensionality reduction techniques for visualization attempt to find a 2 or 3-dimensional embedding that preserves aspects of the data in its original, high-dimensional space,

such as clusters, manifolds, and pairwise dissimilarity. The specific characteristics preserved vary by method, e.g. local distances, local neighborhoods, pseudo-geodesic, or global distances. For the neural data we are interested in preserving the neighborhoods of similar neural modulation by having them nearby in the latent space, while at the same time separating dissimilar neural activity into different portions of latent space.

For qualitative analysis we use a 2-dimensional latent space mapped to a color-wheel. Thus, each point's location defines its color, and we color the corresponding point of the external dynamics with the same color for easy visual analysis.

A. Neural data representation

Multi-electrode neural recordings provide high-dimensional data in the form of series of neuron action potentials. Instead of using the exact time of the action potentials many models count only the spikes in contiguous non-overlapping fixed-width bins with a bin width in the tens to hundreds of milliseconds. At each time step a vector is composed of the spike counts of all neurons

$$x_i = [r_1(i), r_2(i), \ldots, r_n(i)] \in \mathbb{R}^n$$

where $r_j(i)$ is the count for the jth neuron at the ith time step, and n is the number of neurons. We wish to find an visualization of a set $\{x_i\}_{i=1}^N$ of these vectors that results in a set $\{y_i\}_{i=1}^N \in \mathbb{R}^2$ such that the location of the y_i's preserves the *relative* pairwise Euclidean distances between the x_i's.

One of the difficulties in rate counts is choosing a bin width that captures the dynamics while reducing the variability. For simplicity, we use a single choice of 100 ms bins and use a 3-tap moving-average filter on each individual neural channel before applying the static dimensionality reduction. In general, different choices of bin size and filter order need consideration depending on the data.

B. Stochastic Neighborhood Embedding

The algorithm t-distributed Stochastic Neighborhood Embedding, *t-SNE*, has been shown to produce useful visualizations of high-dimensional datasets [9]. The algorithm uses a probabilistic formulation with the Kullback-Leibler divergence as the cost function. Specifically, all of the pairwise Euclidean distances in both spaces, original and latent, are transformed to densities that represent the probability of points i and j being in the same neighborhood. In the original space, the joint density is formed as the symmetric combination of the conditional probability of finding j in the neighborhood of i and vice versa (2). The conditional density is considered a Gaussian density centered around point i, (1), where the scale parameter σ_i is chosen so that the conditional density has a user-defined perplexity— the logarithm of perplexity is the Shannon entropy. The perplexity corresponds to a smoothed estimate of number of neighbors each point has [9]. In the low-dimensional latent space, the density function centered at each point is the Student's t-distribution with one degree of freedom, i.e. the

Cauchy distribution (3). Unlike the original space, the choice of scale is arbitrary, but the Cauchy distribution provides a much larger tail than that of the Gaussian. This avoids a "crowding" problem [9]. In a high dimensional space, many points can exist at the same distance and density, but in the low dimensional space, the same number of points would have to crowd together to be at the same density with a Gaussian distribution. However, with the Cauchy distribution the density falls off much slower with increasing distance, thereby increasing the range of distances that are within a given density range. The increase in range of distances allows the points to spread out resulting in useful visualization.

Let $\{x_i\}_{i=1}^N$ be the points in the original space and $\{y_i\}_{i=1}^N$ be the points in the embedded space. The original conditional densities are represented by

$$p_{j|i} = \frac{\exp(-\|x_j - x_i\|^2/2\sigma_i^2)}{\sum_{k \neq i} \exp(-\|x_k - x_i\|^2/2\sigma_i^2)}. \tag{1}$$

Then the joint density is

$$p_{ij} = (p_{i|j} + p_{j|i})/2n. \tag{2}$$

The embedded space has joint density

$$q_{ij} = \frac{(1 + \|y_i - y_j\|^2)^{-1}}{\sum_{k \neq l}(1 + \|y_k - y_l\|^2)^{-1}}. \tag{3}$$

The cost function is the Kullback-Leibler Divergence,

$$C = D_{KL}(P\|Q) = \sum_i \sum_{j \neq i} p_{ij} \log(p_{ij}/q_{ij}). \tag{4}$$

The algorithm initializes the latent space with a PCA projection, and proceeds to minimize the KL divergence by gradient descent with a momentum term. See [9] for the details on the derivation; their code is publicly available[1].

C. Prediction Performance Measure

In section, we quantify how well the low-dimensional embedding preserves information that would be useful in a decoding task, versus using the high-dimensional data for the same task. As a data driven approach, we consider the performance of nearest neighbor prediction using a scaled version of the Γ_2 statistic proposed in [10]. In general, for a set of joint samples $\{(y_i, z_i)\}_{i=1}^N$, we find the nearest neighbor in the marginal space and find the corresponding neighborhood size of those point in the other marginal space, where the neighbors are defined using Euclidean distance. Let the index of the nearest neighbor of y_i be denoted $\nu_{\{y\}}(i)$, and let $\mathrm{KNN}_z(i, j)$ denote the function that returns which neighbor of z_i is z_j, e.g. 1st, 2nd, or kth- nearest neighbor. The measure is

$$\gamma_{\{(y,z)\}} = \frac{1}{N} \sum_i^N \frac{\mathrm{KNN}_{\{z\}}(i, \nu_{\{y\}}(i)) - 1}{N - 1}. \tag{5}$$

Values of γ close to 0 indicate good prediction, values close to 0.5 indicate chance prediction. The value of γ can be interpreted as the percentage of the predicted space that the nearest neighbor maps to: a one-to-one mapping corresponds to $\gamma = 0$.

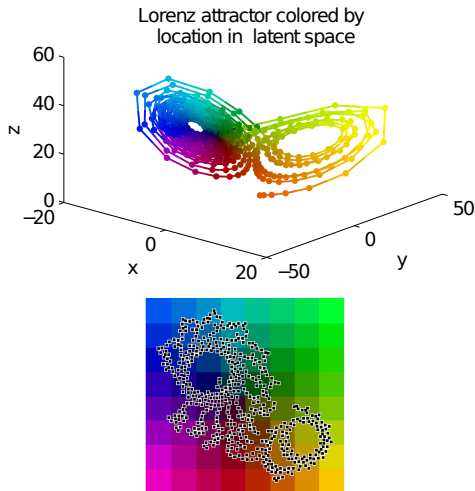[1] http://homepage.tudelft.nl/19j49/t-SNE.html

Fig. 1. (Top) The Lorenz attractor colored by its latent space embedding. (Bottom) The points in latent space define the color coding. The latent space clearly captures the two orbits of the attractor.
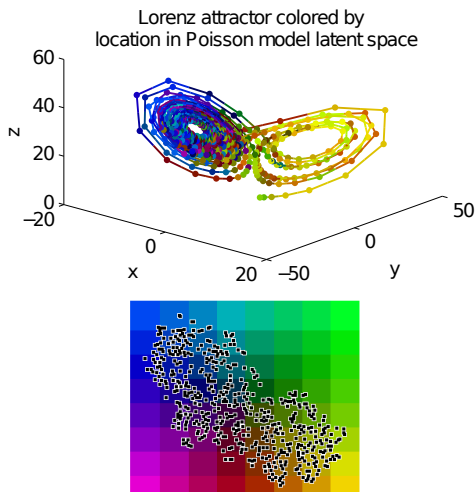

Fig. 2. (Top) The Lorenz attractor colored by the Poisson model's latent space embedding. (Bottom) The points in latent space define the color coding. The original orbits are not visible as in Fig. 1, but based on the color-coding their relative locations are still preserved.

III. RESULTS

A. Simulated Data

The technique is tested on data with known dynamics, the Lorenz attractor, a dynamical system (6) that produces quasi-periodic cycles in $\mathbb{R}^3$. A simulated neural spike count model is driven by the Lorenz attractor's state space. The underlying probabilistic model is Poisson distributed spike counts, a common model for the variability of spiking data. The location in state space (x_i, y_i, z_i) at time step i controls the rate of three inhomogeneous Poisson processes, $(\lambda_{1,i}, \lambda_{2,i}, \lambda_{3,i})$ through non-linear equations using the normal cumulative

distribution function on the empirically normalized locations (7). In addition, three additional dimensions, $(\lambda_4, \lambda_5, \lambda_6)$, are added that correspond to uninformative 'neurons' whose counts are Poisson distributed with rate of 3.

$$\frac{dx}{dt} = 10(y - x), \frac{dy}{dt} = x(28 - z) - y, \frac{dz}{dt} = xy - \frac{8}{3}z \quad (6)$$

$$
\begin{aligned}
\lambda_{1,i} &= 1 + 5\Phi((x_i - \bar{x})/s_x) \\
\lambda_{2,i} &= 10 - 5\Phi((y_i - \bar{y})/s_y) \\
\lambda_{3,i} &= 1 + 5\Phi((z_i - \bar{z})/s_z) \quad (7) \\
\Phi(t) &= 1/\sqrt{2\pi} \int_{-\infty}^{t} \exp(-\tau^2/2)d\tau
\end{aligned}
$$

The Lorenz attractor was sampled 500 times during 25 s of dynamics (the initial point was $(1, 1, 1.5)$). The sampled locations drive the Poisson model and the resulting count series is filtered with a 3-tap moving average filter. We use *t-SNE* with a perplexity of 25 on both the original samples, Fig. 1, and the filtered count data, Fig. 2. From Fig. 2, it appears that the variability of the count data destroys the distinctive trajectories as in Fig. 1, but the location in the latent space corresponds to location in the original state space as seen by the color-coding in Fig. 2.

For quantitative analysis, 200 random realizations are generated for the same dynamics, and the ability to predict locations in the original 3-dimensional state-space is assessed by (5). Table I has the mean and standard deviation of γ when using both the original count and the filtered count data, *t-SNE*, and the first two principle components (PCA). From the table it is clear that filtering is necessary with the count data, *t-SNE* preserves the filtered structure, and PCA is not able to capture the same structure.

B. Reach and Grab Task

In this section, we test goal of this method in producing a visualization of high-dimensional neural space that preserves the association to known behavioral data. The neural data we use was collected in Dr. Nicolelis's primate laboratory at Duke University. Specifics can be found in [1]. The data is recorded from an owl monkey's cortex while the animal was performing a food reaching task. Multiple micro-wire arrays record from 104 neural cells in multiple cortical areas: posterior parietal cortex, left and right primary motor cortex, and dorsal premotor cortex. Synchronous recordings provide the reaching hand's position in three dimensions. The spikes are binned at 100 ms, and then filtered with a 3-tap moving average filter, the hand position is also at 10 Hz. In this dataset, there is over 38 min of data during which the animal would reach and eat food sitting on two different tray, and in-between reaches the animal would return its hand to a resting location.

Applying *t-SNE* with a perplexity of 15 to 100 s of filtered data results in the embedding shown in Fig. 3; the original hand trajectories are colored by the instantaneous location in the latent neural space. The clear distinction in coloring

for the different segments indicate that across multiple trials the reach, return, and resting segments correspond to disparate locations in latent space. This implies that the neural responses for each movement have distinct firing patterns from each other. The responses for multiple reaches and returns are relatively more similar to each other than the portions that correspond to resting. Overall, the visualization results revealed strong association with movement, which is expected on data that was used for real-time movement reconstruction [1].

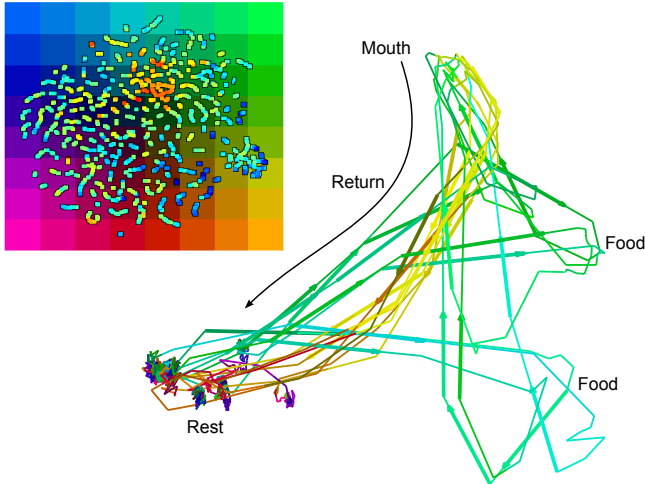Hand trajectories colored by location in latent neural space

Fig. 3. The movement trajectories colored by the corresponding neural state's location in the latent space. Similar colors for similar movements indicate a useful latent space embedding. (Inset) Position in latent space defines the color of the corresponding movement. Each point is also colored coded by the average firing rate across all units (red being the highest and dark blue the lowest ensemble firing). In the latent space plot, the cluster of points on the lower right are represented by yellow and correspond to the return movement, points on the upper center and right are represented by light green or cyan and correspond to the reach movements, the points corresponding to the rest state spread across the remainder of the latent space.

To quantify the performance across more trials, we use 7 non-overlapping 300 s segments, since this is 3 times as many points we scale the perplexity 3-fold to 45, and measure the movement predication (5) using the original count data, the filtered count data, the latent space produced by *t-SNE*, and the first two components from principle component analysis (PCA). From Table I, the mean value of γ indicates that the nearest-neighbor in the original space is a poor predictor, filtering the count data significantly increases performance, and *t-SNE* maintains most of the structure in just two-dimensions, whereas the PCA projection is not useful for prediction.

IV. CONCLUSION

The results obtained using this method can provide an answer to the question of whether similar movement trajectories have similar neural representation. It does this by producing a visualization that maintains the neighborhoods

TABLE I
γ PREDICTION PERFORMANCE (5) FOR BOTH DATASETS

	III-A		III-B	
	Mean	St. Dev.	Mean	St. Dev.
Original	0.408	0.016	0.427	0.017
Filtered	0.293	0.018	0.078	0.016
t-SNE w/ Filtered	0.294	0.018	0.087	0.018
PCA w/ Filtered	0.340	0.016	0.444	0.016

of the original data, which is useful for prediction. The method requires no explicit model or trial-based averaging. In addition, it can be used with minimal preprocessing, i.e. spike binning and low-pass filtering of the rates, and *t-SNE* only requires a single user-defined parameter, the perplexity of the original space. For the movement task, the visualization neatly segmented the different portions of movement. Similar attempts for visualization were made using PCA and ensemble averaging, but these methods were not successful in capturing any distinction between the segments of the movements. Overall, this technique is useful for exploratory analysis of neural recordings without assumptions or model formulation. However, since the method is stochastic the user is left with no explicit model for the embedding; thus, further modeling that captures the structure seen in the visualization is still required.

V. ACKNOWLEDGMENTS

REFERENCES

[1] J. Wessberg, C. Stambaugh, J. Kralik, P. Beck, M. Laubach, J. Chapin, J. Kim, S. Biggs, M. Srinivasan, and M. Nicolelis, "Real-time prediction of hand trajectory by ensembles of cortical neurons in primates," *Nature*, vol. 408, no. 6810, pp. 361–365, 2000.

[2] M. A. Lebedev and M. A. Nicolelis, "Brain-machine interfaces: past, present and future," *Trends in Neurosciences*, vol. 29, no. 9, pp. 536 – 546, 2006.

[3] M. Churchland, B. Yu, M. Sahani, and K. Shenoy, "Techniques for extracting single-trial activity patterns from large-scale neural recordings," *Current Opinion in Neurobiology*, vol. 17, no. 5, pp. 609–618, 2007.

[4] B. Yu, A. Afshar, G. Santhanam, S. I. Ryu, K. Shenoy, and M. Sahani, "Extracting dynamical structure embedded in neural activity," in *Advances in Neural Information Processing Systems 18*, Y. Weiss, B. Schölkopf, and J. Platt, Eds. Cambridge, MA: MIT Press, 2006, pp. 1545–1552.

[5] B. M. Yu, J. P. Cunningham, G. Santhanam, S. I. Ryu, K. V. Shenoy, and M. Sahani, "Gaussian-Process Factor Analysis for Low-Dimensional Single-Trial Analysis of Neural Population Activity," *Journal of Neurophysiology*, vol. 102, no. 1, pp. 614–635, July 2009.

[6] S. Roweis and L. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, p. 2323, 2000.

[7] M. Stopfer, V. Jayaraman, and G. Laurent, "Intensity versus identity coding in an olfactory system," *Neuron*, vol. 39, no. 6, pp. 991 – 1004, 2003.

[8] B. M. Broome, V. Jayaraman, and G. Laurent, "Encoding and decoding of overlapping odor sequences," *Neuron*, vol. 51, no. 4, pp. 467 – 482, 2006.

[9] L. van der Maaten and G. Hinton, "Visualizing Data using t-SNE," *Journal of Machine Learning Research*, vol. 9, pp. 2579–2605, 2008.

[10] J. Friedman and L. Rafsky, "Graph-theoretic measures of multivariate association and prediction," *The Annals of Statistics*, vol. 11, no. 2, pp. 377–391, 1983.