Active Manifold Exploration

Ge Huang
Carnegie-Mellon University
gehuang@andrew.cmu.edu

Darby Losey
Carnegie-Mellon University
dlosey@andrew.cmu.edu

Yu Chen
Carnegie-Mellon University
yuc2@andrew.cmu.edu

Abstract

As imaging and electrophysiology experiments grow in number of neurons being monitored, dimensionality reduction has been commensurately growing in use as an analysis technique for denoising and capturing the latent structure of neural data, where each dimension of the original space is defined by the activity of a single neuron. We propose to extend the current factor analysis based techniques through the model space reduction ideas of active learning. Utilizing Hessian local linear embedding in conjunction with this active learning approach, we present a novel method of sampling the underlying neural manifold in order to select a small number of "summary points" that encapsulate the structure of the entire manifold in the original space. This would allow experimenters to pick samples, based on previous simulation or experimental results, that will allow for rapid manifold estimation in the original response space, where samples can be sensory stimuli or motor intentions.

1 Introduction

Neural population dynamics has become a central focus due to the rapid increase in the number of neurons that can be recorded simultaneously in vivo. Following a pattern akin to Moore's law, this number has doubled approximately 7 years (Stevenson & Kording, 2011) and thus has lead to increased interest in how to estimate the low dimensional structure in the high dimensional data space. The state of the art include (Yu *et al.* , 2009) for brain computer interfaces, (Kobak *et al.* , 2016) for supervised dimensionality reduction use-cases, and (Buesing *et al.* , 2012) for temporally smooth spaces. More recently, some statistical neuroscientists are beginning to consider more efficient sampling techniques by leveraging the sampling history or simulations (Lewi *et al.* , 2009; Bak *et al.* , 2016; Cowley *et al.* , 2017a). Additionally, naturally occurring high dimensional data with low dimensional nonlinear structure has also been studied in manifold learning field (Belkin & Niyogi, 2002; Tenenbaum *et al.* , 2000; Roweis & Saul, 2000). Ghashami *et al.* (2016) showed one way of doing manifold estimation kernel principal components analysis (PCA) on streaming data, Schölkopf *et al.* (1998) provided nonlinear PCA for close purpose. Cowley *et al.* (2017b) could handle nonlinear manifold neural activity structure using distance covariance analysis. Belkin & Niyogi (2002) suggested natural clustering could serve as basis for low dimensional recovering, and Voevodski *et al.* (2012) proposed efficient active learning clustering by using a small sample set as landmarks. Additional manifold learning studies can be found in (Bengio *et al.* , 2006; Brand, 2003; Donoho & Grimes, 2003; Hinton & Roweis, 2003; Maaten & Hinton, 2008; Bengio *et al.* , 2013). Nevertheless, such advanced methods have not been widely implemented in neural datasets, due to their non-robustness against neural noise.

31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.

Theoretical progress has been made regarding sample efficient ways to solve related unsupervised learning problems (Settles, 2010, 2012). (Krishnamurthy & Singh, 2013) investigated how to use active sampling to complete a matrix or tensor. Dasarathy *et al.* (2016) showed how to constrain estimates of graphical models with active learning. (Zhang *et al.* , 2011; Li *et al.* , 2014; Zhang *et al.* , 2017) were novel works implementing active learning on exploring nonlinear manifold through geometric methods. Active learning for reduction of expensive experimental steps is growing in computational biology (Naik *et al.* , 2016; Voevodski *et al.* , 2012), but this innovation has yet to make major inroads in neuroscience.

Rapid manifold estimation cuts down redundant electrophysiolocial experimentation. Manifold estimation is an important element in understanding the underlying neural representations (Sadtler *et al.* , 2014) and is thus a request measurement for many experiments. Rapid manifold estimation is especially pertinent for experiments involving human or nonhuman primate subjects, as data collection is especially time sensitive due to factors such as expense and subject motivation. This is especially true in the case of brain-computer interfaces (BCI). BCI involves decoding neural activity, traditionally recorded from motor cortex, to drive an end effector, such as a cursor on a computer screen. As such, BCIs often rely on the subject generating specific neural patterns to represent intentions and therefore BCI activity that samples from within the manifold space is more effective than those that require activity outside the neural manifold (Sadtler *et al.* , 2014). For human patients suffering from severe paralysis, BCI may serve as the only mode of communication. Due to potential instabilities in the neural signal, BCIs often have to be recalibrated (Jarosiewicz *et al.* , 2017). Manifold estimation can be an important component of BCI recalibration, and thus reducing the amount of time necessary for manifold estimation could reduce the burden on the BCI user.

We extend the results of (Zhang *et al.* , 2011) to develop a novel unsupervised metric that selects the data points that provide the best structural representation of the neural manifold utilizing Hessian local linear embedding and active learning paradigms. We also present the results on both simulated and real neural data from the visual system of nonhuman primates, collected by the authors, to validate the efficiency and effectiveness of our algorithm. The focus on vision stems from data availability and not necessarily the constraints of our methods.

2 Method Description

2.1 Manifold learning

First we compare recent popular manifold learning algorithms (Wang, 2011), including local linear embedding (LLE) (Roweis & Saul, 2000), isomap (Tenenbaum *et al.* , 2000), Laplacian eigenmaps (LEIGS) (Belkin & Niyogi, 2002), and Hessian local linear embedding (HLLE) (Donoho & Grimes, 2003). As these techniques are unsupervised and don't rely on prior knowledge of the manifold structure, all of them rely on the same strategy for extracting low dimension manifold structure. These algorithms search for the local spatial relationship among neighboring points first, then construct a new global coordinate for each point within the manifold. Stated differently, these algorithms employ an "explore locally then reconstruct globally" approach. Some of the algorithms can both use K nearest neighbor points or data points within a sphere during the explore locally step. Either method provides similar results so we will focus on the K nearest neighbors approach to keep the paper succinct.

LLE uses the linear combination of K nearest neighbors to localize each point by minimizing the reconstruction error over the weights $\sum_i \left| X_i - \sum_{j \in O_K(X_i)} W_{ij} X_j \right|^2$ with constrain $\sum_j W_{ij} = 1$, O_K is the neighbor set. Such method using local points to explore is not sensitive to rotation or linearly stretching of the manifold. The underlying low dimension manifold will keep points tight. Once the weights are obtained, it looks for embedding coordinates which can minimize $\sum_i \left| Y_i - \sum_{j \in O_K(X_i)} W_{ij} Y_j \right|^2$. Since both target equations are quadratic, analytical solutions can be easily obtained.

Isomap uses the shortest path between all pairs of points to "explore locally". The shortest path distance matrix D can be acquired using Floyd-Warshall algorithm Cormen (2009). Let $S = D \circ D$, where $\circ$ denotes the Hadamard product. Then centralize the S by subtracting the row and column

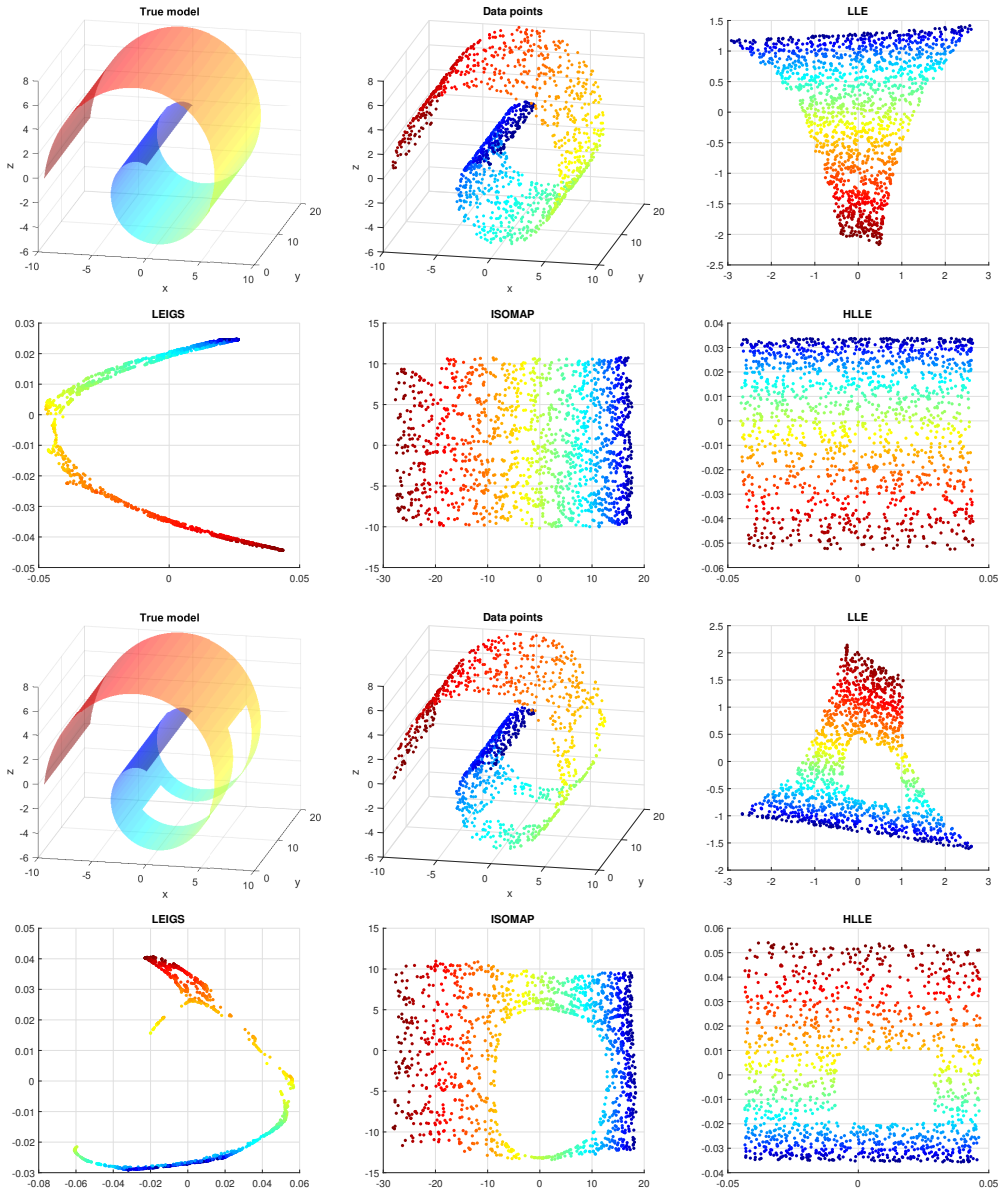

Figure 1: A comparison of manifold learning algorithms on simulated data. The first row is the results of swiss-roll dataset, the second row is about swiss-roll with a hole. LLE can reconstruct both cases, but it loss the geodesic distance for a folded shape. The distance between points on the inner end is different from that on the outer end. LEIGS cannot handle too folded shapes. Isomaps using shortest path to "explore locally", so it maintains the geodesic property very well, but it is sensitive to the topology of the structure, the hole will stretch the shape around that. So far the HLLE performances the best on both folded shapes.

mean. S reflects the geodesic distance between points within the manifold. The principle covariance analysis (PCA) is used to find the new coordinate axis, along which the data points are most widely spanned.

Laplacian eigenmaps also starts with pairwise shortest paths like Isomap. The goal is finding an axis where the projections of data points stay as close as possible. In another word, we aim to minimize $\sum_{i,j}(y_i - y_j)^2 W_{i,j}$, where $W_{i,j} = e^{-\frac{||x_i - x_j||^2}{\sigma}}$, σ is a kernel parameter. The further points get more

weight in order to push them together. After a few simple steps of calculation, the target is equivalent to searching for eigenvectors with minimal eigenvalues of $Ly = \lambda Dy$, where D is a diagonal matrix, with $D_{ii} = \sum_j W_{i,j}$. $L = D - W$. This method has a geometric meaning. Let $f : \mathcal{M} \to R$ be a function mapping manifold points to a real number. If we add a small disturb δ at point x, we see the consequence disturb of the function is $|f(x+\delta) - f(x)| \approx |\langle f(x), \delta \rangle| \leq ||\nabla f|| \cdot ||\delta||$. First approximation comes from local linear approximation, the second inequality comes from the definition of operator norm. If we want to find a low dimension manifold, we want project the points to stay as closer, so we want to minimize, $\int_{\mathcal{M}} ||\nabla f(x)||^2$ where f has unit L^2 norm. It has been shown that $\int_{\mathcal{M}} ||\nabla f(x)||^2 = \int_{\mathcal{M}} \mathcal{L}(f) \cdot f$. The Laplacian-Beltrima operator is $\mathcal{L} = \mathrm{div}\nabla(f)$.

Hessian LLE considers local Hessian instead of local Laplace-Beltrsmi operator, i.e. to minimize the Hessian functional $\mathcal{H}(f) = \int_{\mathcal{M}} ||H_x^{\tan}(f)||_F^2$, $H_x^{\tan}$ denotes the local tangent Hessian evaluated at x with respect to the tangent hyperplane through x. As has been proved, the Hessian functional $\mathcal{H}$ has $d+1$ dimension null space. For discrete expression on dataset, we can only use the K nearest neighbors to construct local Hessian, so for each point x_i only consider the function as a vector of $f_i = [f(x_{i_1}), f(x_{i_2}), ..., f(x_{i_k})]^T$. $\mathcal{H}(f)|_{W_i} = f_i^T W_i f_i$. W_i is a $K \times K$ matrix whose null space contains coordinate functions of the tangent space. Thus basis for the Hessian functional can be constructed using singular vectors and their 2-way interaction (Hadamard product between vectors) of $M^i = [(x_{i_1} - \bar{x}), (x_{i_k} - \bar{x})]$. Now the problem is reduced to find the eigenvectors with smallest eigenvectors of the sum of W_i for all points.

2.2 Active Manifold learning

Inspired by (Zhang *et al.* , 2011), we develop an algorithm combining the manifold learning and active learning to select the most informative points representing the the manifold. Zhang *et al.* (2011)'s algorithm achieved this goal through minimizing two target equations by incorporating LLE,

$$\begin{aligned} \epsilon(q_1, q_2, ..., q_m) &= \sum_{i=1}^{k} ||q_{s_i} - x_{s_i}||^2 + \mu \sum_{i=1}^{m} ||q_i - \sum_{j=1}^{m} W_{ij} q_j||^2 \\ &= \sum_{i=1}^{k} ||q_{s_i} - x_{s_i}||^2 + \mu \mathrm{tr}(Q^T M Q) \end{aligned} \tag{1}$$

where $M = (I - W)^T (I - W)$. Instead of seeking low dimension representation, (Zhang *et al.* , 2011)'s method is looking for reconstructed coordinates Q still in high dimension first. The selected points are not too far away from corresponding original points (first term), while maintaining the manifold structure (second term) using local nearest neighbors' weights W.

$$e(x_{s_1}, x_{s_2}, ..., x_{s_k}) = ||X - Q||^2 \tag{2}$$

The first equation is a quadratic function of Q, so the exact solution can be achieved. The optimal solution is a function of a subset of data points $\{x_{s_1}, x_{s_2}, ..., x_{s_k}\}$. The second equation chooses that subset to minimize the difference between original data points and reconstructed data points using sequential method, which just selects the points x_{s_i} one by one.

One advantage of this method is that it merges active learning and manifold learning by simply combining the target equations. Our method adopts this framework by replacing the manifold learning optimization component. Since LLE does not perform well on non-convex structure, and could lose the intrinsic geodesic distance in low dimension space as shown in figure 1, we choose to use HLLE. As introduced in last section, HLLE explores the local structure via Hessian operator, and the optimization is also quadratic.

$$\begin{aligned} \epsilon(q_1, q_2, ..., q_m) &= \sum_{i=1}^{k} ||q_{s_i} - x_{s_i}||^2 + \mu \int_{\mathcal{M}} ||\mathbf{H}_x^{\tan}(f)||_F^2 \mathrm{d}x \\ &= \sum_{i=1}^{k} ||q_{s_i} - x_{s_i}||^2 + \mu \mathrm{tr}(Q^T K Q) \end{aligned} \tag{3}$$

$$e(x_{s_1}, x_{s_2}, ..., x_{s_k}) = ||X - Q||^2 \tag{4}$$

Akin to active HLLE, the framework can also be extended to LEIGS.

3 Results

3.1 Manifold Learning

Our model not only performed well on the more regular shapes as shown in Figure 1, it also successfully unfolded the more noisy and complicated neural data. This data includes the response of 23 neurons to 180 images from 4 classes, with each class containing 45 images. As shown by Figure 2, PCA failed to learn the intrinsic structure of the neural space spanned by those 180 images. With PCA, the four classes overlap. In contrast, our model was able to separate them,as shown by the four clearly grouped clusters. The red boxes denote the 20 "summary points" returned by our model.

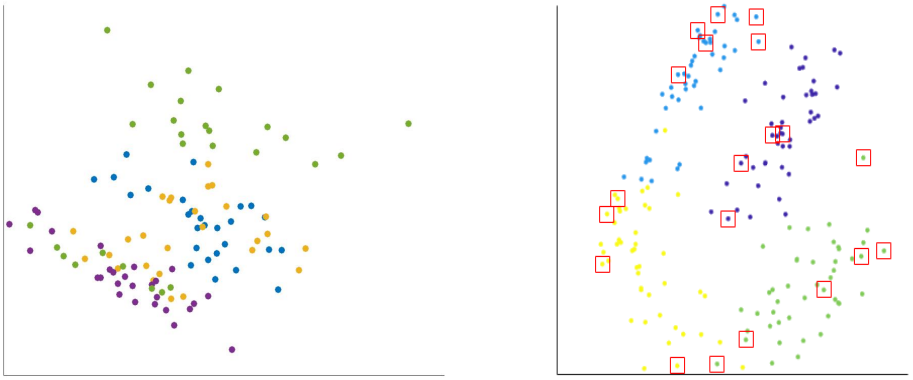

Figure 2: Left: PCA; Right: manifold learned by our model. Different colors represent images from different classes.

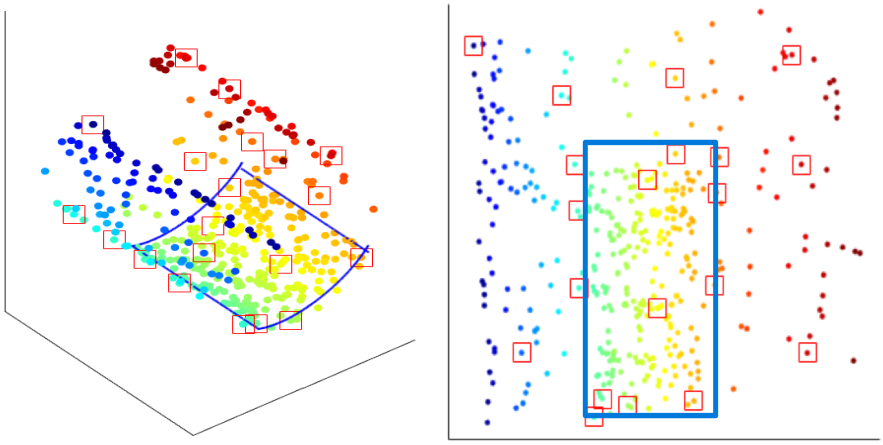

Figure 3: Left: sampling from swiss roll; Right: unfolding in 2d.

3.2 Sampling within the manifold

Besides unfolding the swiss roll, our model was able to sample most representative points from the data as shown in Figure 3. This figure shows 400 points, with half of them contained with a small subsection, denoted by the blue box. Our model avoids the influence of point density while sampling,

instead sampling uniformly along the manifold space. Thus the samples selected from within the blue box are roughly proportional to the area of the blue box in comparison to the area of the entire manifold.

We also tested it on simulated V1 neural data. Gabor is the most commonly used model for V1. We picked 24 Gabor neurons, Figure 4, including 8 orientations and 3 frequencies (0.55, 1.1 and 2.2 cycles/degree) which are the typical used to model V1 neurons. 800 randomly sampled images from ImageNet (Deng *et al.* (2009)) were then presented to the 24 Gabor neurons with effectively size of one degree. The modulus of the complex response is then taken as the Gabor response.

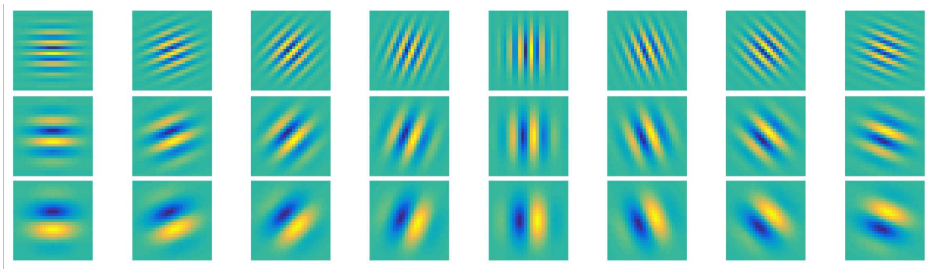

Figure 4: 24 Gabor filters in 8 orientations and 3 frequencies.

Sparsity is an important character of V1 neurons. The sparseness we used here is from Vinje (2000) and defined as $S = [1 - (\sum_i r_i/n)^2 / \sum_i (r_i^2/n)]/(1 - 1/n)$ where r_i is a neuron's response to the i_th image. A neuron will not respond to most image stimulus or responds to many images in the weak way. There are only 5% of images in the dataset that can evoke relatively higher responses. This causes a lot of redundancy in experiments. As experimental time for animals is restricted, it requires an efficient way to choose images in order to avoid those that evoke either no response or a response similar to all other images. Gabor filters have been commonly used in modeling the function of V1 neurons. In our simulation, we use the output of Gabor filters over images to represent the neural response. Our algorithm selects images corresponding to data points evenly distributed in neural response space, so it can efficiently cover all different neural response strengths within 20 stimulus (5). Under random selection, about 125 images are needed and may still miss many areas of neural response space (Woloszyn & Sheinberg, 2012).

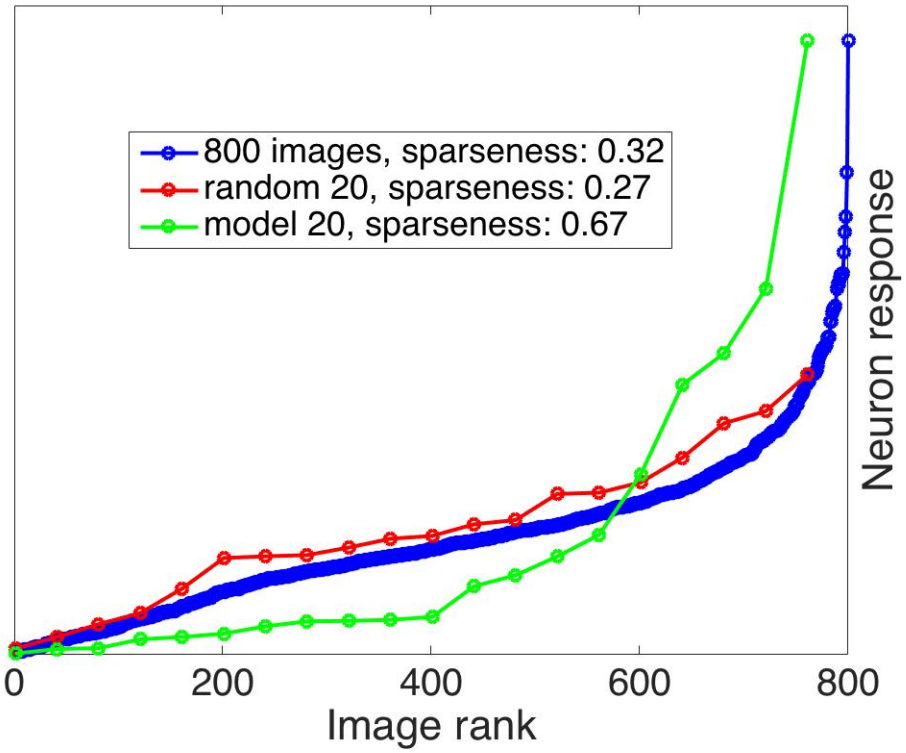

Figure 5: Tuning curve from the original 800 images (blue), 20 images randomly sampled from the images space (red), and 20 images sampled by our model from the neural space (green).

4 Discussion

A more attractive method for neuroscientists, but also a more challenging task, is to adaptively select the stimulus in order to uncover the unknown portions of the neural response manifold in an online setting. However, this requires further study into the relationship between stimulus and neural response and it is not sufficient to simply understand the neural response manifold. More specifically, greater research needs to be done to figure out how to design stimulus avoiding evoking the existing response.

For the visual system recordings and simulations, we visualized the low dimension neural response manifold to see if the selected dataset was evenly distributed in order to verify our results. Since there is no specific training tasks for the neural response dataset, like there were for Zhang *et al.* (2011), it is hard to quantify the validation of results. The tuning curve method can provided a partial verification, but more sophisticated methods for verification need to be developed. For motor system data, we could potentially perform learning task or using decoder through brain-computer interface. The decoding accuracy over the selected data set versus the whole dataset can be used to evaluate the algorithm. We leave these ideas as an avenue for future research.

Acknowledgments

We would like to thank Professor Nina Balcan, Arun Sai Saggala, Faisal Baqai and Ben Cowley for useful discussion.

References

Bak, Ji Hyun, Choi, Jung, Witten, Ilana, Akrami, Athena, & Pillow, Jonathan W. 2016. Adaptive optimal training of animal behavior. *Pages 1939–1947 of: Neural Information Processing Systems*.

Belkin, Mikhail, & Niyogi, Partha. 2002. Laplacian eigenmaps and spectral techniques for embedding and clustering. *Pages 585–591 of: Advances in neural information processing systems*.

Bengio, Yoshua, Larochelle, Hugo, & Vincent, Pascal. 2006. Non-local manifold parzen windows. *Pages 115–122 of: Advances in neural information processing systems*.

Bengio, Yoshua, Courville, Aaron, & Vincent, Pascal. 2013. Representation learning: A review and new perspectives. *IEEE transactions on pattern analysis and machine intelligence*, **35**(8), 1798–1828.

Brand, Matthew. 2003. Charting a manifold. *Pages 985–992 of: Advances in neural information processing systems*.

Buesing, Lars, Macke, Jakob H, & Sahani, Maneesh. 2012. Spectral learning of linear dynamics from generalised-linear observations with application to neural population data. *Advances in Neural Information Processing Systems (NIPS)*, 1–9.

Cormen, Thomas H. 2009. *Introduction to algorithms*. MIT press.

Cowley, Ben, Williamson, Ryan, Clemens, Katerina, Smith, Matthew, & Yu, Byron. 2017a. Adaptive sampling for a population of neurons. *In: Advances in Neural Information Processing Systems*.

Cowley, Benjamin, Semedo, Joao, Zandvakili, Amin, Smith, Matthew, Kohn, Adam, & Yu, Byron. 2017b. Distance Covariance Analysis. *Pages 242–251 of: Artificial Intelligence and Statistics*.

Dasarathy, Gautamd, Singh, Aarti, Balcan, Maria-Florina, & Park, Jong H. 2016. Active Learning Algorithms for Graphical Model Selection. *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics*, 1356–1364.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. 2009. ImageNet: A Large-Scale Hierarchical Image Database. *In: CVPR09*.

Donoho, David L, & Grimes, Carrie. 2003. Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data. *Proceedings of the National Academy of Sciences*, **100**(10), 5591–5596.

Ghashami, Mina, Perry, Daniel, & Phillips, Jeff M. 2016. Streaming Kernel Principal Component Analysis. *International Conference on Artificial Intelligence and Statistics*, **41**, 1–16.

Hinton, Geoffrey E, & Roweis, Sam T. 2003. Stochastic neighbor embedding. *Pages 857–864 of: Advances in neural information processing systems*.

Jarosiewicz, Beata, Sarma, Anish A, Saab, Jad, Franco, Brian, Cash, Sydney S, Eskandar, Emad N, & Hochberg, Leigh R. 2017. Retrospectively supervised click decoder calibration for self-calibrating point-and-click brain–computer interfaces. *Journal of Physiology-Paris*.

Kobak, Dmitry, Brendel, Wieland, Constantinidis, Christos, Feierstein, Claudia E., Kepecs, Adam, Mainen, Zachary F., Qi, Xue-Lian, Romo, Ranulfo, Uchida, Naoshige, & Machens, Christian K. 2016. Demixed principal component analysis of neural population data. *eLife*, **5**(APRIL2016), 1–36.

Krishnamurthy, Akshay, & Singh, Aarti. 2013. Low-Rank Matrix and Tensor Completion via Adaptive Sampling. *Pages 1–20 of: Advances in Neural Information Processing Systems*.

Lewi, J., Butera, R., Schneider, D.M., Woolley, S.M.N., & Paninski, L. 2009. Designing neurophysiology experiments to optimally constrain receptive field models along parametric submanifolds. *Advances in Neural Information Processing Systems 21 - Proceedings of the 2008 Conference*, 945–952.

Li, Cheng, Liu, Haifeng, & Cai, Deng. 2014. Active learning on manifolds. *Neurocomputing*, **123**, 398–405.

Maaten, Laurens van der, & Hinton, Geoffrey. 2008. Visualizing data using t-SNE. *Journal of Machine Learning Research*, **9**(Nov), 2579–2605.

Naik, Armaghan W., Kangas, Joshua D., Sullivan, Devin P., & Murphy, Robert F. 2016. Active machine learning-driven experimentation to determine compound effects on protein patterns. *eLife*, **5**(FEBRUARY2016), 1–21.

Roweis, Sam T, & Saul, Lawrence K. 2000. Nonlinear dimensionality reduction by locally linear embedding. *science*, **290**(5500), 2323–2326.

Sadtler, Patrick T, Quick, Kristin M, Golub, Matthew D, Chase, Steven M, Ryu, Stephen I, Tyler-Kabara, Elizabeth C, Byron, M Yu, & Batista, Aaron P. 2014. Neural constraints on learning. *Nature*, **512**(7515), 423–426.

Schölkopf, Bernhard, Smola, Alexander, & Müller, Klaus-Robert. 1998. Nonlinear component analysis as a kernel eigenvalue problem. *Neural computation*, **10**(5), 1299–1319.

Settles, Burr. 2010. Active learning literature survey. *University of Wisconsin, Madison*, **52**(55-66), 11.

Settles, Burr. 2012. Active learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, **6**(1), 1–114.

Stevenson, Ian H, & Kording, Konrad P. 2011. How advances in neural recording affect data analysis. *Nature neuroscience*, **14**(2), 139–142.

Tenenbaum, Joshua B, De Silva, Vin, & Langford, John C. 2000. A global geometric framework for nonlinear dimensionality reduction. *science*, **290**(5500), 2319–2323.

Vinje, W. E. 2000. Sparse Coding and Decorrelation in Primary Visual Cortex During Natural Vision. *Science*, **287**(5456), 1273–1276.

Voevodski, Konstantin, Balcan, Maria-florina, Roglin, Heiko, Teng, Shang-Hua, & Xia, Yu. 2012. Active clustering of biological sequences. *The Journal of Machine ...*, **13**, 203–225.

Wang, Jianzhong. 2011. *Geometric structure of high-dimensional data and dimensionality reduction*. Springer.

Woloszyn, Luke, & Sheinberg, David L. 2012. Effects of Long-Term Visual Experience on Responses of Distinct Classes of Single Units in Inferior Temporal Cortex. *Neuron*, **74**(1), 193–205.

Yu, Byron M., Cunningham, John P., Santhanam, G, Ryu, Si, Shenoy, Kv, & Sahani, M. 2009. Gaussian-Process Factor Analysis for Low-Dimensional Single-Trial Analysis of Neural Population Activity (vol 102, pg 614, 2009). *Journal of Neurophysiology*, **102**(April 2009), 614–635.

Zhang, Lijun, Chen, Chun, Bu, Jiajun, Cai, Deng, He, Xiaofei, & Huang, Thomas S. 2011. Active learning based on locally linear reconstruction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **33**(10), 2026–2038.

Zhang, Lining, Shum, Hubert PH, & Shao, Ling. 2017. Manifold Regularized Experimental Design for Active Learning. *IEEE Transactions on Image Processing*, **26**(2), 969–981.