Statistical Analysis of the Non-stationarity of Neural Population Codes*

Sung-Phil Kim*, Frank Wood*, Matthew Fellows[†], John P. Donoghue[†], Michael J. Black*

*Department Computer Science,   [†]Department of Neuroscience
Brown University, Providence, RI 02912, USA
Corresponding author: spkim@cs.brown.edu

Abstract—**Neural prosthetic technology has moved from the laboratory to clinical settings with human trials. The motor cortical control of devices in such settings raises important questions about the design of computational interfaces that produce stable and reliable control over a wide range of operating conditions. In particular, non-stationarity of the neural code across different behavioral conditions or attentional states becomes a potential issue. Non-stationarity has been previously observed in animals where the encoding model representing the mathematical relationship between neural population activity and behavioral variables such as hand motion changes over time. If such an encoding model is formed and learned during a particular training period, decoding performance (neural control) with the model may not be consistent during successive periods even when the same task is repeated. It is critical in both laboratory experiments and in clinical settings to be able to evaluate whether the representation of movement encoded by a neural population has changed or not. Such information can be used as a cue to retrain the system or as feedback to an adaptive decoding algorithm. To that end, we develop a statistical methodology to evaluate changes in the neural code over time using a generative probabilistic decoding model. The changes are evaluated by comparing the likelihoods of firing rates given similar distributions of 2D hand kinematics collected while a primate periodically performs a manual cursor control task. A comparison is performed by measuring a distance between probabilistic encoding models trained at different times. The statistical significance of the distance measurements are justified with a systematic statistical hypothesis test. The experimental results demonstrate that the likelihood changes over different periods with the change being greater when more distant periods are compared.**

Index Terms— **Non-stationarity, Neural prostheses, Generative model, Distance measure, Hypothesis test**

I. INTRODUCTION

Neural motor prostheses directly connect the central nervous systems of severely disabled patients to external devices such as computer cursors or robot arms. A great deal of scientific and clinical progress has been made as the result of new microelectrode recoding technology that enables the chronic recording of populations of tens or hundreds of cortical neurons. The current success of motor cortical brain-machine interfaces (BMIs) has also been driven by an understanding of how the brain encodes information about body pose and motion and by new algorithms for transforming neural activity into a reliable control signal. To decode the intended action from the electrical activity recorded from motor cortical neural populations a number of statistical methods have been proposed including multidimensional linear filters [1], [2], Bayesian models using the Kalman filter [3] or the particle filter [4], [5], population vector coding [6], nonlinear neural networks [7], [8], [9], and support vector machines [10].

These methods model the functional mapping between neural activities and kinematic parameters (e.g. 2D hand position and velocity) using stationary statistical assumptions (i.e. model parameters are estimated based on the assumption that functional mapping is time invariant). They may suffer from poor generalization or deteriorating performance after training due to the intrinsic dynamics of the functional mapping. It is well known that neuronal tuning properties can change due to plasticity [11], [12]. There have also been a number of BMI studies reporting the reorganization of functional properties of neural populations [13], or temporal changes of neuronal tuning properties [14]. These non-stationary characteristics of neural populations become critical to practical aspects of neural prostheses, especially for the design and implementation of decoding algorithms. In particular, if the functional mapping between neural activity and kinematic parameters changes over time, the approximation of the mapping by a decoding model learned during a certain time period may not be as accurate at some future time. Therefore, it is important to identify when and how the functional mapping temporally changes such that the decoding models can adapt to it or the model can be retrained. However, the identification of such dynamics of neural populations is still an open problem.

In this paper, we do not address the issue of the identification of neural dynamics, but propose a basic methodology to examine the non-stationary properties of neural populations. This work supports the findings from many other researchers regarding the time-variant functional mapping between neural activity and kinematics. With the proposed methodology, the statistical significance of changes can be determined. This methodology is based on the Bayesian decoding model that has been successfully applied to BMIs. This model-based methodology will help to advance the present decoding models such that they can be applied more readily to non-stationary environments.

In this study, we analyze experimental data in which a

primate performs a cursor control task. During a recording session, however, the primate does not always attend to the control task and is occasionally distracted. (See [15] for the analysis of these data showing that the attentional states could be discriminated by classifying the temporal sequence of firing rates.) Therefore, the data can be divided into multiple disjoint attentional periods. We seek to investigate if the functional mapping between neural firing rates and hand kinematics changes between those periods. The functional mapping is specified by a likelihood function that models the probability of neural firing rates conditioned on hand kinematics. The difference between likelihood models trained during different periods is measured by two distance measures: the Bhattacharyya distance and the Kullback-Liebler (KL) divergence. In order to determine whether the measured distance is statistically significant, a hypothesis test is employed with a bootstrap-type random sampling procedure. The overall testing procedures and simulation results are reported in the following sections. Note also that for this comparison of likelihood models to be valid, the statistics of the hand motion in the different periods must be the same; care is taken to ensure this in Section II.B.

While previous studies have addressed changes in the neural code using simple properties such as the "preferred direction" of cells [11], our approach leverages a richer probabilistic model of cortical coding. In particular, the Bayesian encoding model captures the statistical variation in the population activity as it relates to multiple behavioral variables. Since such a model captures the statistical variation in the neural activity it provides a principled foundation on which to build an analysis of non-stationarity.

II. MODELING LIKELIHOOD OF FIRING RATES

A. Data Descriptions

The proposed approach is investigated using neural activity and hand kinematics recorded while monkeys performed a sequential reaching task. We used one dataset from each of 2 monkeys, denoted hereafter by monkey-1 and monkey-2. The task and recording setup used to make the recordings for both data sets were similar to that used in [15]. Briefly, the monkey sat in a primate chair and held a 2-link manipulandum which constrained movements to the horizontal plane. Reach targets and a hand position feedback cursor were presented on a video screen in front of the animal. When a reach target was presented the animal's task was to move the manipulandum so that the feedback cursor moved into the target and remained in the target for 500ms, at which time that target was extinguished and a new reach target was presented in a different location. This was repeated for up to 10 targets per trial. Upon successful completion of a trial the animal received a juice reward. Hand kinematics and neural activity were simultaneously recorded while the animal performed the task. Spikes were manually isolated off-line. Firing rate was estimated by binning spikes into 70 msec time bins,
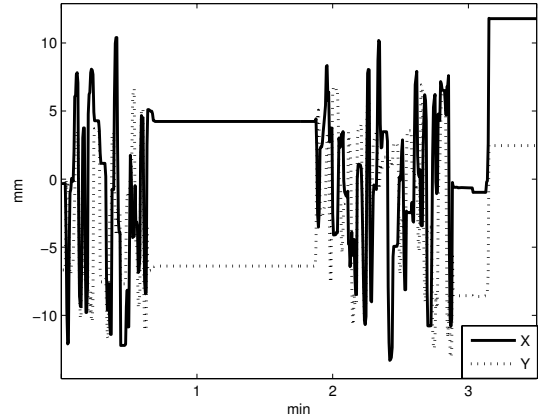

Fig. 1. A sample of the x, y hand trajectory: a monkey periodically attends to a 2D hand motion task. Note several sections in the data where the hand position remains perfectly constant. In these intervals the monkey has removed their hand from the manipulandum and is not performing the task.

and hand position and velocity were computed every 70 msec.

During a typical recording session each monkey spent only a portion of the time actually performing the task. Periodically they were distracted from the task, often letting go of the manipulandum and performing natural arm movements. When the monkeys were not gripping the manipulandum no motion was recorded. These periods of inactivity are easy to identify by inspection as depicted in Fig. 1 where the x-, and y-coordinates of one monkey's hand position trajectory are displayed for a brief segment of the data.

The determination of whether or not a monkey was performing the task was made by detecting significant regions of constant manipulandum position and zero velocity. Since we have no record of what monkeys were doing during such segments, we only use the task-related movement segments for training models of neural coding.

B. Modeling a mapping between firing rates and hand kinematics

Our goal in this study is to investigate temporal changes to the mapping between firing rates and hand kinematics and to develop a principled mathematical model for detecting any such changes. Here we exploit previous work on Bayesian decoding in which a linear Gaussian model was used to relate neural activity to hand kinematics. Such a "likelihood" model can be combined with a prior probability model of hand kinematics to enable recursive Bayesian decoding of motor cortical activity [3]. Here our focus is not on decoding but rather on how the neural encoding may change over time. Consequently we focus on the likelihood term which represents the neural code.

Let $X = \{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_N\}$ be a time ordered sequence of hand kinematic parameters such as position, velocity and acceleration and let $Z = \{\vec{z}_1, \vec{z}_2, \ldots, \vec{z}_N\}$ be a corresponding sequence of firing rates where each $\vec{z}_i =$

$[z_i^1, z_i^2, \ldots, z_i^M]^T$ is a vector of the firing rates of M single units or multi-units. For simplicity, we assume $E[\vec{x}] = 0$ and $E[\vec{z}] = 0$ by centering X and Z, respectively.

The likelihood $p(\vec{z}_i|\vec{x}_i)$ of the observed firing rates conditioned on the hand kinematics can be derived from the following generative model of neural activity:

$$\vec{z}_i = f(\vec{x}_i) + \vec{w}_i. \tag{1}$$

where $\vec{w}_i$ corresponds to noise in the observations. If we assume that $f(\cdot)$ is a linear function and $\vec{w}_i$ follows a multivariate Gaussian distribution with zero mean vector and covariance W, (1) can be written as

$$\vec{z}_i = H\vec{x}_i + \vec{w}_i, \tag{2}$$

where H is a linear projection matrix. Note that the firing rates are normally distributed: $\vec{z}_i \sim N(H\vec{x}_i, W)$.

H and W can be estimated from training data using least squares such that [16]

$$\hat{H} = argmin_H \sum_i \left\| \vec{z}_i - H\vec{x}_i \right\|^2 \tag{3}$$

$$\hat{W} = \frac{1}{N-1} \sum_i (\vec{z}_i - \hat{H}\vec{x}_i)(\vec{z}_i - \hat{H}\vec{x}_i)^T \tag{4}$$

Here we focus on a simplified model in which the error covariance matrix $\hat{W}$ is taken to be a diagonal matrix, assuming conditional independence between units. This reduces the chance of overfitting when the training data is limited as it is in our experiments. The ideas developed here however are fully general and apply when a full covariance matrix can be estimated. Note that in the diagonal case the sum of diagonal entries represents the total error variance which can be used to evaluate how well the linear model fits the data.

The linear Gaussian encoding model used here has proven useful for decoding. Combining the likelihood with a linear Gaussian prior over hand kinematics yields a recursive Bayesian decoding method (the Kalman filter) [3]. The approach provides accurate and efficient estimates of hand kinematics from neural firing rates.

III. STATISTICAL EXPERIMENT DESIGN

A. Distance between Likelihoods

In order to investigate the temporal changes of the functional mapping between firing rates and hand kinematics, we compare the likelihood of observations given hand kinematics from different attentional segments. Specifically we train a linear Gaussian model $p_j(\vec{z}|\vec{x})$ using the data in each segment j in which the monkey is attending to the task. A comparison between models can be accomplished by measuring a distance between each $p_j(\vec{z}|\vec{x})$. There are numerous methods to measure the distance among which we opt for using the well-generalized measures such as the Bhattacharyya distance [17] and the Kullback-Leibler (KL) divergence [18].

Consider two distributions $p_1(\vec{y})$ and $p_2(\vec{y})$ with means $\vec{\mu}_1$ and $\vec{\mu}_2$ and covariances Σ_1 and Σ_2, respectively. As we approximate the likelihood based on the linear Gaussian

generative model, the Bhattacharyya distance B_{12} between two Gaussian distributions is given by

$$B_{12} = \frac{1}{8}(\vec{\mu}_1 - \vec{\mu}_2)^T \left(\frac{\Sigma_1 + \Sigma_2}{2}\right)^{-1}(\vec{\mu}_1 - \vec{\mu}_2)$$
$$+ \frac{1}{2}\log\frac{|(\Sigma_1 + \Sigma_2)/2|}{\sqrt{|\Sigma_1|}\sqrt{|\Sigma_2|}}. \tag{5}$$

Since the mean of the Gaussian likelihood is zero, only the second term including covariances need be considered such that

$$B_{12} = \frac{1}{2}\log\frac{|(\Sigma_1 + \Sigma_2)/2|}{\sqrt{|\Sigma_1|}\sqrt{|\Sigma_2|}}. \tag{6}$$

Next, the KL divergence between $p_1(\vec{y})$ and $p_2(\vec{y})$ is defined by

$$D(p_1\|p_2) = \int_{\vec{y}} p_1(\vec{y})\log\frac{p_1(\vec{y})}{p_2(\vec{y})}. \tag{7}$$

For two M-dimensional Gaussian distributions, the KL divergence can be expressed as

$$D(p_1\|p_2) = tr(\Sigma_2^{-1}\Sigma_1) + \log|\Sigma_2| - \log|\Sigma_1| - M. \tag{8}$$

where $tr(\cdot)$ represents the trace of a matrix. Since the original KL divergence is asymmetric, a symmetrized version of the distance may be required which can be defined by

$$D_s(p_1\|p_2) = \frac{1}{2}\big(D(p_1\|p_2) + D(p_2\|p_1)\big)$$
$$= \frac{1}{2}\big(tr(\Sigma_2^{-1}\Sigma_1) + tr(\Sigma_1^{-1}\Sigma_2)\big) - M \tag{9}$$

B. Statistical Methodology

While the above measures give a distance between two likelihoods and hence two encoding models they do not directly tell us whether the neural coding is the same or not. For this, we evaluate whether the difference between models is statistically significant by developing a hypothesis test. This test can be done in such a way that the temporal variation of the likelihood is examined; given multiple disjoint attentional segments, the earliest one is set as a basis to which the rest of the segments are compared using the statistical test. This reveals how much the likelihood changes over time.

One must be careful however that the statistical difference between likelihoods is not the result of a difference in the hand kinematics between the two segments. To ensure that the statistics of the hand motions are the same in each segment we impose conditions on the training data such that the distributions of the hand kinematics in different segments are approximately identical. This is obviously a very strong constraint due to the fact that the hand trajectory of a monkey may be a nonstationary time series. Consequently we randomly sample hand motions from each segment such that the distributions of motions from each segment are the same. The training data then empirically satisfies the condition that the hand motions must be drawn from identical distributions. Then, the statistical test of the likelihood difference is performed only on sample sets that have similar hand kinematic

distributions. With this preprocessing, the temporal changes of the likelihood will specifically indicate the change of the functional properties of neural populations.

An overall procedure of the statistical test is as follows; from two given segments, two disjoint sample sets are randomly selected. A statistical test for identical hand kinematic distributions is performed between the two selected sets. If the test reveals that hand kinematic distributions are not statistically different, the likelihood for each set is modeled using equations (3) and (4) and the distances between the likelihood models are measured; otherwise, new sets are randomly selected again. This process is repeated by running Monte Carlo simulations to generate the empirical distribution of the distance measurements, similar to the bootstrap approach. The generation of distribution is performed 1) *within* the basis segment, and 2) *between* one of the rest of the segments and the basis, respectively. For this pair of empirical distributions, a statistical test is performed with the null hypothesis that two distributions are not statistically different. If this null hypothesis is rejected at a given significance level, we can state that the likelihood approximated from a given segment is statistically different from the one in the basis segment. This test is performed for each of the rest of the segments.

Now, let us describe our statistical methodology in detail. Let $\{\vec{z}_{1i}, \vec{x}_{1i}\}_{i=1}^{N_1}$ and $\{\vec{z}_{2i}, \vec{x}_{2i}\}_{i=1}^{N_2}$ be disjoint randomly selected sample sets both from the basis segment, or one from the basis segment and the other from a different segment. Before proceeding to model the likelihood, the similarity of $p(\vec{x}_1)$ and $p(\vec{x}_2)$ must be examined. Here, we employ another statistical hypothesis test to see if the two distributions are equal to each other. Suppose the hand kinematics follow a multivariate Gaussian distribution, which has been widely assumed in generative decoding models, such that $p(\vec{x}_1) \sim N(\vec{\mu}_1, \Sigma_1)$ and $p(\vec{x}_2) \sim N(\vec{\mu}_2, \Sigma_2)$. Then, a comparison of mean and covariance is sufficient to check if $p(\vec{x}_1)$ and $p(\vec{x}_2)$ are equally distributed. Hence, we first employ the likelihood ratio test for the multivariate Gaussian distributions to test equal covariance matrices [19]. The null hypothesis for the test is given by

$$H_0 : \Sigma_1 = \Sigma_2 = \Sigma. \tag{10}$$

Let Q_1 and Q_2 be defined by

$$Q_k = \sum_i^{N_k} (\vec{x}_{ki} - \bar{x}_k)(\vec{x}_{ki} - \bar{x}_k)^T, \quad k = 1, 2 \tag{11}$$

where $\bar{x}_k$ is the maximum likelihood estimate of mean, i.e. $\bar{x}_k = \frac{1}{N_k} \sum_i \vec{x}_{ki}$ for $k = 1, 2$. Then the maximum likelihood estimate of Σ_1 and Σ_2 are determined as

$$\hat{\Sigma}_k = \frac{1}{N_k - 1} Q_k, \quad k = 1, 2. \tag{12}$$

The joint likelihood function for two distributions is defined as the product of two individual likelihood functions, where each likelihood function is represented as a Gaussian density function. If we substitute the maximum likelihood estimates of mean and covariance obtained above into each Gaussian density function, we obtain the maximum likelihood joint likelihood function represented as

$$L_{12}(\hat{\vec{\mu}}_1, \hat{\vec{\mu}}_2, \hat{\Sigma}_1, \hat{\Sigma}_2) = \frac{\exp(-\frac{NM}{2})}{(2\pi)^{\frac{NM}{2}} |\hat{\Sigma}_1|^{\frac{N_1}{2}} |\hat{\Sigma}_2|^{\frac{N_2}{2}}} \tag{13}$$

where L_{12} denotes the joint likelihood function of $\vec{x}_1$ and $\vec{x}_2$, M is the dimension of $\vec{x}$ and $N = N_1 + N_2$. The alternative maximum likelihood function with the null hypothesis, i.e. $\Sigma_1 = \Sigma_2 = \Sigma$, can also be similarly represented as

$$L_{12}(\hat{\vec{\mu}}_1, \hat{\vec{\mu}}_2, \hat{\Sigma}, \hat{\Sigma}) = \frac{\exp(-\frac{NM}{2})}{(2\pi)^{\frac{NM}{2}} |\hat{\Sigma}|^{\frac{N}{2}}}, \tag{14}$$

where $\hat{\Sigma}$ is the maximum likelihood estimate of Σ, determined as $\hat{\Sigma} = \frac{1}{N-1}(Q_1 + Q_2)$. The likelihood ratio statistic is then given by

$$\begin{aligned}
\ell &= \frac{L_{12}(\hat{\mu}_1, \hat{\mu}_2, \hat{\Sigma}, \hat{\Sigma})}{L_{12}(\hat{\mu}_1, \hat{\mu}_2, \hat{\Sigma}_1, \hat{\Sigma}_2)} \\
&= \frac{|\hat{\Sigma}|^{-\frac{N}{2}}}{|\hat{\Sigma}_1|^{-\frac{N_1}{2}} |\hat{\Sigma}_2|^{-\frac{N_2}{2}}} \\
&= c_{12} \frac{|Q_1|^{-\frac{N_1}{2}} |Q_2|^{-\frac{N_2}{2}}}{|Q_1 + Q_2|^{-\frac{N_1+N_2}{2}}}, \tag{15}
\end{aligned}$$

where

$$c_{12} = \frac{N^{\frac{NM}{2}}}{N_1^{\frac{N_1 M}{2}} N_2^{\frac{N_2 M}{2}}}. \tag{16}$$

Using the large sample theory, it is shown that $-2 \log \ell$ approximately follows $\chi^2_{M(M+1)/2}$. Then, we can apply the χ^2 test for the null hypothesis.

Once the null hypothesis of equal covariance is accepted at a significance level α, the next test for equal means given equal covariance can be done with a relatively simple method, called a Hotelling T^2 test [19]. The null hypothesis for the mean is set to

$$H_0 : \vec{\theta} = \vec{\mu}_1 - \vec{\mu}_2 = \vec{0}. \tag{17}$$

Let $S_p \equiv (Q_1 + Q_2)/(N_1 + N_2 - 2)$. The the null T^2 statistic is calculated as

$$T_0^2 = \frac{N_1 N_2}{N_1 + N_2} (\bar{x}_1 - \bar{x}_2)^T S_p^{-1} (\bar{x}_1 - \bar{x}_2). \tag{18}$$

H_0 is rejected at the α significance level if

$$T_0^2 \geq \frac{M(N_1 + N_2 - 2)}{N_1 + N_2 - M - 1} F^\alpha_{M, N_1+N_2-M-1}. \tag{19}$$

If the above hypotheses of equal mean and covariance are accepted for $p(\vec{x}_1)$ and $p(\vec{x}_2)$, then we can move on to approximate the likelihood from each sample set. If at least one of them is rejected, new random sets are sampled and tested again.

After approximating likelihoods, the Bhattacharyya distance and the symmetrized KL divergence between two likelihoods are measured. This procedure is repeated through Monte Carlo simulations, generating a set of

distance measurements. After simulations, we obtain the set of distances within the basis segment if two data sets are sampled from the same basis segment. The same procedure can be performed for two data sets one of which is sampled from the basis segment and the other is sampled from a different segment. Then, we can test if there is a significant difference between two distance measurement sets. Since the distribution of distance is unknown (and possibly non-normal), the nonparametric one-sided Kolmogorov-Smirnov (KS) test is used to test the null hypothesis given by

$$H_0 : F_1(d) \le F_2(d) \tag{20}$$

where d is a random variable representing the distance measure (Bhattacharyya or symmetrized KL divergence) and $F(d)$ is the cumulative density function (cdf) of d. Here we set $F_1(d)$ as the cdf of the distances within the basis segment and $F_2(d)$ as the cdf of the distances between the basis segment and another segment. This hypothesis is tested by the empirical cdf $\hat{F}_1(d)$ and $\hat{F}_2(d)$ created from given samples, with a test statistic of the maximum difference between $\hat{F}_1(d)$ and $\hat{F}_2(d)$. The rejection of H_0 means that $F_1(d) > F_2(d)$; this indicates that the differences between distributions are much smaller for samples drawn from the basis. Consequently, the distance between different segments is statistically larger than the distance within the basis segment. With this test, we can justify the change of the likelihood of firing rates given hand kinematics over subsequent attentional segments.

IV. EXPERIMENTAL RESULTS

A. Experimental Setup

The spike counts of 44 single/multi units recorded from monkey-1 and 53 single/multi units from monkey-2 are used. All these units are recorded from the $M1$ cortical area of each monkey. The spike counts and the hand kinematic data are centered to have zero mean. Four subsequent attentional segments are extracted from each monkey's data, respectively. The intervals from the basis segment to the remaining segments are as follows; [210, 568, 1,715] (seconds) for monkey-1, and [235, 588, 786] for monkey-2. From each segment, $N = 500$ samples including bin count and hand kinematic parameters (the x-, and y-coordinates of position and velocity, respectively) are randomly selected. The significance levels for the likelihood ratio test, the T^2 test and the KS test are all set to 0.05. 10,000 Monte Carlo runs are performed for each comparison.

B. Results

1) The KS Test: Table I displays the mean and standard deviation of the Bhattacharyya distance and the symmetrized KL divergence measured for each segment. Note that the statistics in the first row are obtained from two sample sets within the basis segment, and the other statistics in the remaining rows are between the corresponding segment and the basis. The null hypothesis of the KS test described in (20) is always rejected in every case; no matter
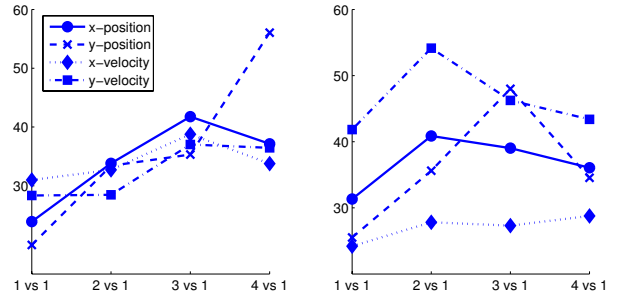

Fig. 2. The average angles (in degrees) between the linear projection vectors for position and velocity are evaluated over multiple attentional segments from two recording sets: monkey-1 (left) and monkey-2 (right).

what distance we measure between the basis segment and any other segment.

TABLE I
DISTANCE MEASUREMENT MEAN (STANDARD DEVIATION)

Compared segments	Monkey-1		Monkey-2	
	Bhatt.	KL Div.	Bhatt.	KL Div.
1 vs 1	0.17(0.09)	1.64(1.55)	0.27(0.10)	2.45(2.58)
2 vs 1	0.76(0.19)	9.85(6.50)	0.78(0.22)	8.73(7.90)
3 vs 1	0.62(0.16)	6.58(3.73)	0.53(0.14)	5.03(2.45)
4 vs 1	2.05(0.22)	51.07(45.75)	0.42(0.17)	4.98(13.13)

2) Comparison of Linear Projections: A difference between encoding models is viewed in another way by evaluating the angle between the linear matrices relating hand kinematics to firing rates (Equation (1)). When we estimate the linear matrices H_1 and H_2 from (2) for two given sample sets, the angles between two corresponding column vectors of H_1 and H_2 (i.e. linear mapping vectors between the firing rates and individual hand kinematics) are evaluated. The average of the angles over Monte Carlo runs for each segment are illustrated in Fig. 2. These results show that the linear projection vectors between firing rates and each hand motion parameter including position and velocity tend to change over attentional segments.

For reference, the mean squared errors of the likelihood approximation in (3) evaluated from the basis segment to the last segment are as follows; 0.08±0.01, 0.08±0.01, 0.07±0.01, 0.07±0.01 for monkey-1, and 0.08±0.01, 0.08±0.01, 0.08±0.01, 0.08±0.01 for monkey-2. This demonstrates that the linear Gaussian fitting for every segment is fairly consistent.

V. CONCLUSIONS, DISCUSSION AND FUTURE WORK

We have proposed a statistical approach for detecting the changes in the neural code relating population firing rates and hand kinematics. We have observed such changes occur over time when monkeys sporadically perform a cursor control task. Here we leverage a learned linear Gaussian model of motor cortical coding and the fact that the learned models in different segments of the data are statistically different. This observed change in neural coding of hand

kinematics over different attentional periods is consistent with previous observations of the temporal change of neuronal tuning properties. Our approach however provides a formal statistical method for analyzing such changes.

The result may have practical implications for neural prosthetic systems. Such systems will have a calibration or training phase and our statistical tests could be used to detect changes in the neural population code that might suggest the need to retrain the system. It is worth noting however that our analysis is performed *off-line*. In an on-line control task the neural population may change in different ways that continue to allow, or even improve, neural control of devices. While our statistical approach could be used to detect such changes, their analysis remains future work.

In our analysis, spiking activity was determined off-line by a human using standard spike sorting software. It remains an open question whether spike detection was consistent across segments of the data. A change in spike detection rates could result in a change in our linear model and this would be detected as a change by our method. The impact of spike sorting error on the non-stationarity of the data deserves further study.

We would like to remark that the statistical results demonstrated here are based on the linear Gaussian assumption. While this has been shown to be useful in practice [3], it is only an approximation of the neural code. Hence, it would be interesting to see if similar results are produced with different modeling schemes. One might employ nonparametric approaches to approximate the conditional likelihood of firing rates given hand kinematics. Then, an information theoretic measure such as the symmetrized KL divergence can be easily estimated for such nonparametric models. Among many possibilities, one can estimate the probabilistic models using Parzen windows from samples and adopt Renyi's entropy to define the KL divergence [20]. This approach could provide more reasonable distance measures without requiring the linear Gaussian assumption.

However, building our statistical approach upon a decoding modeling framework (e.g. the generative modeling approach in this paper) is advantageous since the direct link between our statistical analysis and decoding might help in neural prosthetic applications. For instance, understanding how the likelihood changes over time could be used to develop adaptive decoding models that cope with non-stationarity. This insight is a path towards future applied research.

REFERENCES

[1] M. D. Serruya, N. G. Hatsopoulos, L. Paninski, M. R. Fellows, and J. P. Donoghue, "Brain-machine interface: Instant neural control of a movement signal," *Nature*, vol. 416, pp. 141–142, 2002.

[2] J. Wessberg *et al.*, "Real-time prediction of hand trajectory by ensembles of cortical neurons in primates," *Nature*, vol. 408, no. 6810, pp. 361–365, 2000.

[3] W. Wu, Y. Gao, E. Bienenstock, J. P. Donoghue, and M. J. Black, "Bayesian population coding of motor cortical activity using a Kalman filter," *Neural Computation*, vol. 18, pp. 80–1118, 2005.

[4] A. E. Brockwell, A. L. Rojas, and R. E. Kass, "Recursive Bayesian decoding of motor cortical signals by particle filtering." *Journal of Neurophysiology*, vol. 91, no. 4, pp. 1899–1907, 2004.

[5] Y. Gao, M. Black, E. Bienenstock, S. Shoham, and J. Donoghue, "Probabilistic inference of arm motion from neural activity in motor cortex," *Advances in Neural Information Processing Systems*, vol. 14, pp. 221–228, 2002.

[6] D. Taylor, S. Tillery, and A. Schwartz, "Direct cortical control of 3D neuroprosthetic devices," *Science*, vol. 296, no. 5574, pp. 1829–1832, 2002.

[7] J. K. Chapin and M. A. Nicolelis, "Principal component analysis of neuronal ensemble activity reveals multidimensional somatosensory representations," *Journal of Neuroscience Methods*, vol. 94, pp. 121–140, 1999.

[8] Y. Gao, M. J. Black, E. Bienenstock, W. Wu, and J. P. Donoghue, "A quantitative comparison of linear and non-linear models of motor cortical activity for the encoding and decoding of arm motions," in *1st International IEEE/EMBS Conference on Neural Engineering*, Capri, Italy, March 2003, pp. 189–192.

[9] Y. N. Rao *et al.*, "Learning mappings in brain-machine interfaces with echo state networks," in *IEEE Int. Conf. on Acou., Speech, and Sig. Proc.*, March 2005.

[10] L. Shpigelman, Y. Singer, R. Paz, and E. Vaadia, "Spikernels: Predicting arm movements by embedding population spike rate patterns in inner-product spaces," *Neural Computation*, vol. 17(3), pp. 671–690, 2004.

[11] C. Padoa-Schioppa, C. R. Li, and E. Bizzi, "Neuronal activity in the supplementary motor area of monkeys adapting to a new dynamic environment," *J Neurophysiol.*, vol. 91, no. 1, pp. 449–473, 2004.

[12] J. N. Sanes and J. P. Donoghue, "Plasticity and primary motor cortex," *Annual Review of Neuroscience*, vol. 23, pp. 393–415, 2000.

[13] J. M. Carmena, M. A. Lebedev, R. E. Crist, J. E. O'Doherty, D. M. Santucci, D. F. Dimitrov, P. G. Patil, C. S. Henriquez, and M. A. L. Nicolelis, "Learning to control a brain-machine interface for reaching and grasping by primates," *PLoS Biology*, vol. 1, no. 2, pp. 1–16, 2003.

[14] S. I. H. Tillery, D. M. Taylor, and A. B. Schwartz, "Training in cortical control of neuroprosthetic devices improves signal extraction from small neuronal ensembles," *Reviews in the Neurosciences*, vol. 14, pp. 107–119, 2003.

[15] F. Wood, Prabhat, J. P. Donoghue, and M. J. Black, "Inferring attentional state and kinematics from motor cortical firing rates," in *27th Int. Conf. of the IEEE Engr. in Med. and Bio. Soc.*, September 2005, pp. 1544–1547.

[16] W. Wu and et al, "Neural decoding of cursor motion using a Kalman filter," in *Advances in Neural Information Processing Systems 15*. MIT Press, 2003, pp. 133–140.

[17] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, 2nd ed. New York: Academic, 1990.

[18] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, ser. Wiley Series in Telecommunications, D. L. Schilling, Ed. New York: John Wiley and Sons, 1991.

[19] G. A. F. Seber, *Multivariate Observations*, ser. Wiley Series in Probability and Mathematical Statistics. New york: John Wiley and Sons, 1984.

[20] D. Erdogmus, "Information theoretic learning: Renyi's entropy and its applications to adaptive system training," Ph.D. dissertation, University of Florida, Gainesville, FL, May 2002.