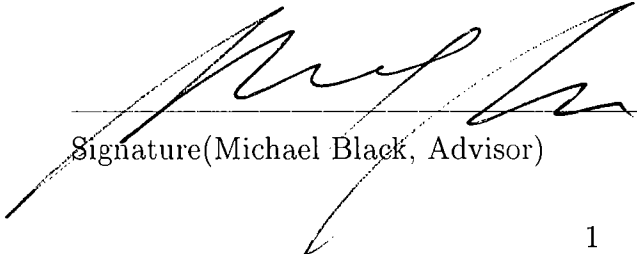


Nonparametric Representation of Neural Activity in Motor Cortex

Yun Gao
Department of Computer Science
Brown University

Submitted in partial fulfillment of the requirements for the Degree of Master of Science in the Department of Computer Science at Brown University.



Signature(Michael Black, Advisor)

April 22, 2002
Date

Acknowledgments

This work was supported by the Keck Foundation and by the National Institutes of Health under grants #R01 NS25074 and #N01-NS-9-2322 and by the National Science Foundation ITR Program award #0113679.

I would like to express sincere gratitude to Professor Michael J. Black, Professor Elie Bienenstock, Professor John P. Donoghue. This thesis would never have been written without their guidance, patience, and support. Also, I would like to thank Mijail Serruya, Matthew Fellows, Liam Paninski and Professor Nicho Hatsopoulos who provided neural data and valuable insights. I am grateful to Professor David Mumford, Professor Stuart Geman and Shy Shoham for many interesting conversations.

Nonparametric Representation of Neural Activity in Motor Cortex

Abstract

It is well accepted that neural activity in motor cortex is correlated to hand motion, previous studies of cosine tuning curve (Georgopoulos *et al.*, 1982) and a modified version (Moran and Schwartz, 1999) are examples at revealing such relationships. Here by analyzing multi-electrode recordings of neural activity and simultaneously recorded hand motion during a continuous tracking task, we introduce a nonparametric representation under a Bayesian framework, and further apply Principle Component Analysis to reduce dimensions and noise. By rigorously comparing all the models with cross-validation, our model are shown superior to previous models at explaining the data. It also reveals the new aspects of MI neural coding: neurons encode speed as well as direction with non-Gaussian pattern of firing; conditional firing rate is better described by regions of smooth activity with relatively sharp discontinuities between them; and only four principle components can count for about 80% of the variance. Finally we explain the superiority of nonparametric models and PCA from Bias/Variance Dilemma (Geman *et al.*, 1992) point of view.

keywords: conditional mean firing rate, regularization, tuning plot, receptive field, Principle Component Analysis, cross-validation, Bias/Variance Dilemma

1 Introduction

The primary motor cortex (MI) plays an important role in movement control, yet its underlying nature of coding is still under debate (Moran and Schwartz, 2000; Georgopoulos and Ashe, 2000; Todorov, 2000; Scott, 2000), of whether MI is coding muscle/joint-based variables, or hand-based variables. Nonetheless it is certain that the neural activity in MI is correlated with hand motion, and describing this relationship is vital for applications such as neuroprosthetics, which is more concerned with higher level end point movement.

Previous studies have proposed several models. the cosine tuning curve (Georgopoulos *et al.*, 1982) and a modified version (Moran and Schwartz. 1999) are successful examples albeit their simple forms. However so far these models are parametric models. with strong assumptions such as unimodality. symmetry. fixed width. infinite smoothness (vs. piecewise smoothness). strict linearity in speed etc. built inside the model. Actually these models are fit to data averaged over thousands of trials. cells and even from different animals. the heterogeneity of individual cells are totally lost during the tremendous averaging. Recent improvement (Georgopoulos *et al.*. 2000) reveals that the profiles of neural activities with respect to direction have more variations and detailed structures than as described by cosine tuning model. yet it is still a parametric model hence the inherited disadvantage of superimposed bias.

Here we explore the use of statistical learning methods for modeling the relationship and specifically nonparametric model-free representations. Our goals are: (i) to investigate the nature of encoding in motor cortex. (ii) to characterize the probabilistic relationship between arm kinematics (hand position or velocity) and activity of a simultaneously recorded neural population.

A multi-electrode array (Figure 1) is used to simultaneously record the activity of 25 neurons in the arm area of primary motor cortex (MI) in awake, behaving, macaque monkeys. This activity is recorded while the monkeys manually track a smoothly and

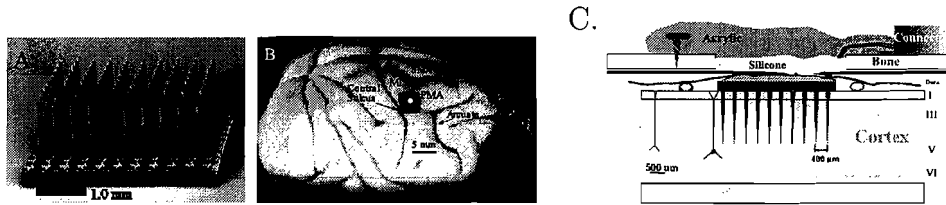


Figure 1: Multi-electrode array. *A.* 10X10 matrix of electrodes. Separation $400\mu\text{m}$ (size 4×4 mm). *B.* Location of array in the MI arm area. *C.* Illustration of implanted array.

randomly moving visual target on a computer monitor (Paninski *et al.*, 2001). Statistical learning methods are used to derive Bayesian estimates of the conditional probability of firing for each cell given the kinematic variables (we consider only hand velocity here). Specifically, we use non-parametric models of the conditional firing, learned using regularization (smoothing) techniques with cross-validation, and further deduct noise by Principle Component Analysis. Our results suggest that the cells encode information about the position and velocity of the hand in space and they have much richer structures than previously shown.

An early version of this work appeared in (Gao *et al.*, 2002)

2 Methods

2.1 Behavioral Task and Recordings

The design of the experiment and data collection is described in detail in (Paninski *et al.*, 2001). Summarizing, a ten-by-ten microelectrode array (Bionic Technologies Inc., Salt Lake City, UT) with 1.5-mm-long platinized tip silicon probes (impedances between 200 and 500 $\text{k}\Omega$; 1 nA, 1 kHz sine wave) arranged in a square grid ($400\mu\text{m}$ separation) is chronically implanted in the arm area of primary motor cortex (MI) of a Macaque monkey (Figure 1) (Donoghue *et al.*, 1998; Hatsopoulos *et al.*, 1998; Maynard *et al.*, 1999; Paninski *et al.*, 2001).

Previous behavioral tasks have involved reaching in one of a fixed number of direc-

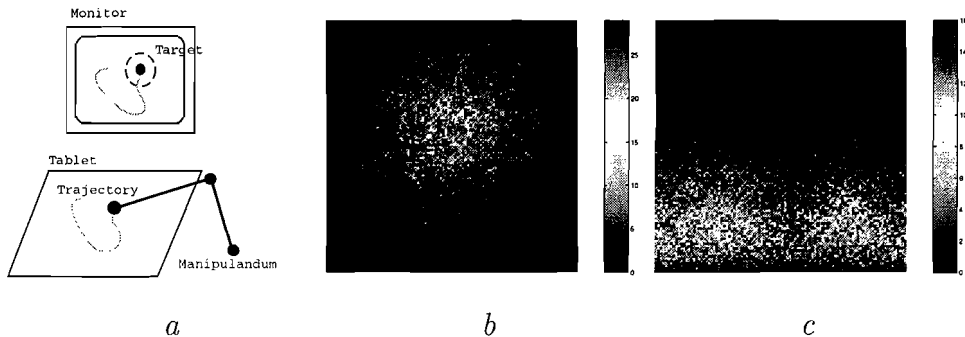


Figure 2: Smooth tracking task. (a) The target moves with a smooth random walk. Distribution of the position (b) and velocity (c) of the hand. Color coding indicates the frequency with which different parts of the space are visited. (b) Position: horizontal and vertical axes represent x and y position of the hand. (c) Velocity: the horizontal axis represents direction, $-\pi \leq \theta < \pi$, and the vertical axis represents speed, r .

tions (Fu *et al.*, 1995; Georgopoulos *et al.*, 1982; Wessberg *et al.*, 2000). To model the relationship between continuous, smooth, hand motion and neural activity, we use a more complex scenario where the monkey performs a continuous tracking task in which the hand is moved on a 2D tablet while holding a two-link, low-friction manipulandum (Wacom Technology Corp., Vancouver, WA) that controls the motion of a feedback dot on a computer monitor (Figure 2a) (Paninski *et al.*, 2001).

The manipulandum constrains hand movement to the plane parallel to the floor. Its x and y positions are digitized at 167 Hz with accuracy of 0.25 mm and stored. The monkey receives a reward upon completion of a successful trial in which the manipulandum is moved to keep the feedback dot within a pre-specified distance of the target for a certain amount of duration. The path of the target is chosen to be a Gaussian process that effectively samples the space of hand positions and velocities: measured hand positions and velocities have a roughly Gaussian distribution (Figure 2b and c) (Paninski *et al.*, 2001). No sample of path is repeated, each stimulus appears only once. Neural signals recorded simultaneously with behavioral task are amplified and sampled at 30 kHz per channel (Bionic), waveforms crossed an experimenter-set thresh-

old are stored (from 0.4 ms before to 1.5 ms after the threshold crossing), and spike sorting is performed off-line to isolate the activity of individual cells (Maynard *et al.*, 1999). Recordings from 25 motor cortical cells are measured simultaneously with hand kinematics.

2.2 Data Preprocessing

In the following analysis, only trials last more than 4 sec long are used to minimize data contamination. The raw recordings of hand positions are smoothed using quintic natural B-spline (Woltring, 1986), and subsampled evenly at every 50 ms; tangential velocities are computed as derivatives at these grid points, then converted to polar coordinates, $\mathbf{v} = [r, \theta]^T$, here r is speed, and moving direction θ is the angle with horizontal axis, with $-\pi \leq \theta \leq \pi$.

Spike trains are binned using 50 ms windows, and in the following analysis we use spike counts 100 ms lag prior to corresponding kinematical samples. this particular choice of lag is made according to study in Moran and Schwartz (1999) and have been tested on the data set.

2.3 Nonparametric Model of the Activity

Here we only consider the neural activity conditioned on velocity. There are more detailed investigations of all possible end point kinematics variables and their interactions elsewhere (Paninski *et al.*, 2001). The velocity space $[0, \text{maximal speed}] \times [-\pi, \pi]$ is discretized on a 100×100 grid. Let $N_{\mathbf{v}}$ be the number of visits to state \mathbf{v} . $f_{\mathbf{v}}(k, c)$ be the spike counts within 50 ms window and 100 ms lag of c -th cell during the i -th visits of \mathbf{v} . Empirical mean firing rate of cell c is then:

$$\hat{f}_{\mathbf{v}}(c) = \sum_k f_{\mathbf{v}}(k, c) / N_{\mathbf{v}}$$

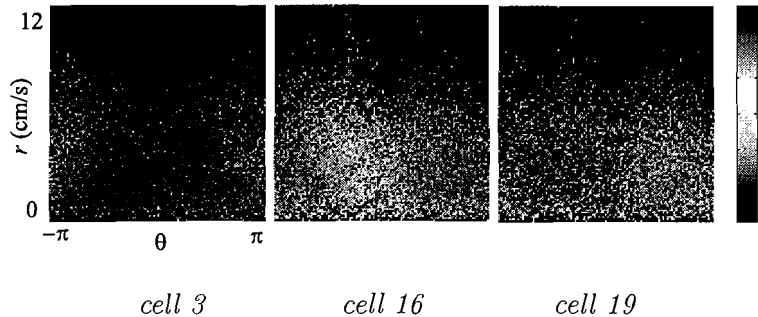


Figure 3: Observed mean conditional firing rate for three cells given hand velocity. The horizontal axis represents the direction of movement, θ , in radians (“wrapping” around from $-\pi$ to π). The vertical axis represents speed, r , and ranges from 0 cm/s to 12 cm/s.

Shown in Figure 3 are examples of three cells. In the following index c is omitted without causing confusion.

We view the measurements like Figure 3 as a stochastic and sparse realization of some underlying model that relates the neural firing to hand motion. Each model can be thought of as a type of “tuning function” or “receptive field” (Paninski *et al.*, 2001) that characterizes the response of the cell conditioned on hand velocity. Previous studies have suggested several parametric models including cosine tuning function (Georgopoulos *et al.*, 1982) and a modified cosine function (Moran and Schwartz, 1999). The disadvantages of these parametric models is the strong assumption superimposed about the shape of the tuning curves. Here we explore a totally model-free non-parametric representation of the underlying activity and adopting a Bayesian formulation. seek a maximum *a posterior* (MAP) estimate of a cell’s conditional firing.

We adopt a weak smoothness assumption, namely, Markov Random Field (MRF) (Geman and Geman, 1984) as prior distribution, that is, the neighboring states with velocities close to each other should have close conditional mean firing rate. Let \mathbf{g} be the array of true underlying conditional neural firing and \mathbf{f} be the corresponding observed

mean firing. We seek the posterior probability

$$p(\mathbf{g} \mid \mathbf{f}) = \Pi_{\mathbf{v}} (\kappa \Pi_k p(f_{\mathbf{v}}(k) \mid g_{\mathbf{v}}) \Pi_{i=1}^{\eta} p(g_{\mathbf{v}} \mid g_{\mathbf{v}_i})) \quad (1)$$

where κ is a normalizing constant independent of \mathbf{g} , $g_{\mathbf{v}}$ is the true mean firing at velocity \mathbf{v} , $g_{\mathbf{v}_i}$ represents the firing rate for the i^{th} neighboring velocity of \mathbf{v} , and the neighbors are taken to be the four nearest velocities ($\eta = 4$).

The first term on the right hand side represents the *likelihood* of observing a particular firing rate $f_{\mathbf{v}}(i)$ given that the true rate is $g_{\mathbf{v}}$. The second term is a spatial *prior* probability that encodes our expectations about Δg , the variation of neural activity in velocity space. In the special case of Gaussian likelihood and Gaussian prior, finding MAP of Equ 1 is equivalent to minimizing energy

$$E(\mathbf{g}) = \sum_{\mathbf{v}} \left(\sum_k (f_{\mathbf{v}}(k) - g_{\mathbf{v}})^2 + \lambda \sum_{i=1}^{\eta} (g_{\mathbf{v}} - g_{\mathbf{v}_i})^2 \right)$$

which is the standard form of regularization (Szeliski, 1989). here λ is a parameter adjusting the amount of smoothing. Under two extremes: when λ is zero, the optimal \mathbf{g} is the empirical conditional mean: when λ is infinity, the optimal \mathbf{g} is totally flat.

One assumption here is that the cells are conditionally independent, independence is hard to test in general, here we only test a weak form, i.e. independence of pairs of cells, by constructing χ^2 test's two way table at each state. Due to very sparse frequencies of visit, most of the states (99.5%) do not have enough measurements to meet the requirement of the test, for those do, at less than 10% of these states the cell pairs reject the null (p-value < 0.05). Overall this assumption is fine under current setting, although incorporating dependence among cells has been done elsewhere (Hatsopoulos *et al.*, 1998).

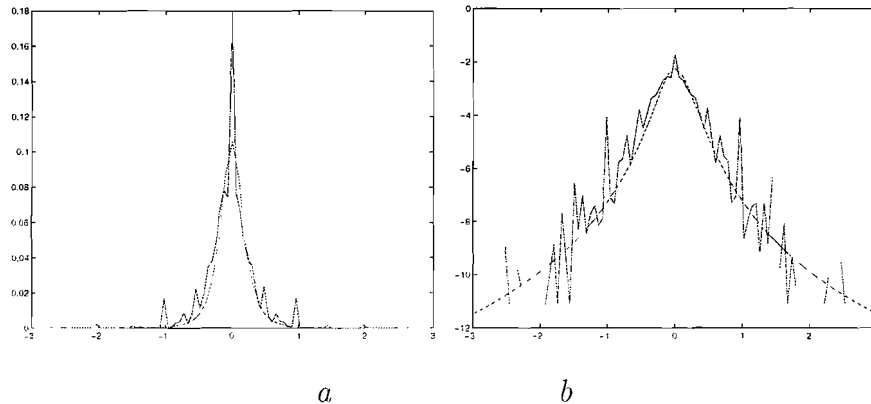


Figure 4: Prior probability of firing variation (Δg). (a) Log probability of firing variation computed from training data (blue). Log of proposed prior model (red) plotted for $\sigma = 0.28$. (b) Logarithm of the distributions shown to provide detail.

2.3.1 Likelihood

Here we consider two *generative models* of the neural spiking process within 50 ms: a Poisson model, p_P , and a Gaussian model, p_G :

$$p_P(f | g) = \frac{1}{f!} g^f e^{-g}, \quad p_G(f | g) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(f - g)^2}{2\sigma^2}\right).$$

2.3.2 Spatial Prior

The MRF prior states that the firing, $g_{\mathbf{v}}$, at velocity \mathbf{v} depends only on the firing at neighboring velocities. A simple choice of prior model for the distribution of Δg is Gaussian distribution, namely,

$$p_G(\Delta g) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(\Delta g)^2}{2\sigma^2}\right)$$

which corresponds to an assumption that the firing rate varies smoothly.

A closer look at the data leads to “robust” prior: a histogram of the differences between adjacent velocities (we have done this separately for r and θ but the differences exhibit the same same spatial statistics and hence the two are combined here) is

constructed and then normalized (see Figure 4). The logarithm of the distribution is also shown to give more detail about its shape. In particular, the heavy tailed distribution exhibits the same sort of spatial variation observed in images of natural scenes (Huang and Mumford, 1999) and implies piecewise smooth data; that is, when there are distinct regions of smooth activity with sharp discontinuities between them. This is an advantage of nonparametric models over parametric models, in both cosine and related models (Georgopoulos *et al.*, 1982; Moran and Schwartz, 1999; Georgopoulos *et al.*, 2000) they all have infinite smooth derivatives, none of them can model this piecewise smoothness. The data is well fit by Student's t-distribution of degree 3:

$$p_R(\Delta g) = \frac{2\sigma^3}{\pi(\sigma^2 + \Delta g^2)^2}$$

whose negative logarithm belongs to a family of robust statistical error functions (Black and Anandan, 1996) which have been used extensively in machine vision and image processing for smoothing data with discontinuities (Black and Rangarajan, 1996; Geman and McClure, 1987).

Here σ is the smoothing parameter. The optimal value of σ is computed for each cell using cross validation. More detailed description is in 2.5.

2.3.3 MAP estimate

Missing data is a computational problem here because of the sparse sampling of the state space. we impute the average conditional mean rate as the initial guess before computing the MAP estimate.

The solution to the Gaussian+Gaussian (i.e. Gaussian likelihood and Gaussian prior) model can be computed in closed form but for the Poisson+Robust model no closed form solution for \mathbf{g} exists and an optimal Bayesian estimate could be achieved with simulated annealing (Geman and Geman, 1984). Instead, we derive an approximate solution for \mathbf{g} in (1) by minimizing the negative logarithm of the distribution using standard regularization techniques (Black and Rangarajan, 1996; Terzopoulos,

1986; Blake and Zisserman, 1987) with missing data, the learned prior model, and a Poisson likelihood term (Nowak and Kolaczyk, 2000). Simple gradient descent (Black and Rangarajan, 1996; Blake and Zisserman, 1987) with deterministic annealing provides a reasonable solution.

2.4 Principle Component Analysis

Generally speaking nonparametric models may fit the noise in training data, here we further apply Principle Component Analysis (PCA) to the MAP estimate of above nonparametric method for dimension reduction, which essentially reduce the variance. Each tuning plot of 25 cells is columnized into a 10000×1 vector: T_i from the training data. S_i from the testing data. a set of principle components or bases $\{b_j\}$ is made up by the eigenvectors corresponding to the largest a few eigenvalues of the covariance matrix of the T_i 's, the number nb chosen by 80% variance criteria together with the descending rate of the eigenvalues. Then each S_i is projected on the subspace of the bases. the coefficients $\{a_j\}$ are chosen such that:

$$\min \sum_{v=1}^{10000} N(v) (S_i(v) - \sum_{j=1}^{nb} a_j b_j(v))^2$$

where v is the state index. $N(v)$ is the frequency of visiting state v . Note the weighted sum of squared error is used here to put more confidence on the states with more measures. The projection \hat{S}_i of S_i is then

$$\hat{S}_i(v) = \sum_{j=1}^{nb} a_j b_j(v)$$

the noise reduced tuning plot we are seeking after reshaped into the usual square matrix.

2.5 Evaluation

To have a quantitative comparison of various models, we use the log likelihood criteria with cross-validation. More specifically, during cross-validation, each time 10 trials

out of 180 are left out for testing and models are fit on the remaining training data, yielding estimated conditional mean rate $\lambda(\mathbf{v})$, then on test data set we compute the log likelihood of the observed spike counts given the model. This provides a measure of how well the model captures the statistical variation in the training set. The whole procedure is repeated 18 times for different test data sets.

Additionally, the sparse sampling of each trial determines the heterogeneity of the log likelihood results, hence only comparison on the same trial is valid, or equivalently taking log likelihood ratio (LLR) of two methods on each trial, that is:

$$LLR(\text{model 1. model 2}) = \log \left(\frac{\Pr(\text{observed firing} \mid \text{model 1})}{\Pr(\text{observed firing} \mid \text{model 2})} \right)$$

Then to test whether one model is significantly better than the other. we apply a nonparametric test --- Wilcoxon signed rank test (Splus, MathSoft Inc.. WA) since the histogram of these 18 numbers usually are not fit by any parametric probability density functions.

The likelihood used here is by assuming independent Poisson distribution of the bin counts. each bin with the predicted mean obtained by $\lambda(\mathbf{v})$ plug in observed velocity. The validness of this assumption is tested by goodness of fit of the residues to Gaussian distribution after normalization. namely:

$$r_i = \frac{3}{2} \frac{c_i^{\frac{2}{3}} - \lambda(\mathbf{v}_i)^{\frac{2}{3}}}{\lambda(\mathbf{v}_i)^{\frac{1}{6}}}$$

where c_i is observed count of a bin. here Anscombe residual is used instead of Pearson residual ($r_i = \frac{c_i - \lambda(\mathbf{v}_i)}{\sqrt{\lambda(\mathbf{v}_i)}}$) because of the high skewness of Poisson with small mean (McCullagh and Nelder, 1989) . 22 out of 25 cells have p-value 1, the other three have p-value $< 10^{-4}$. Overall this is a fine assumption.

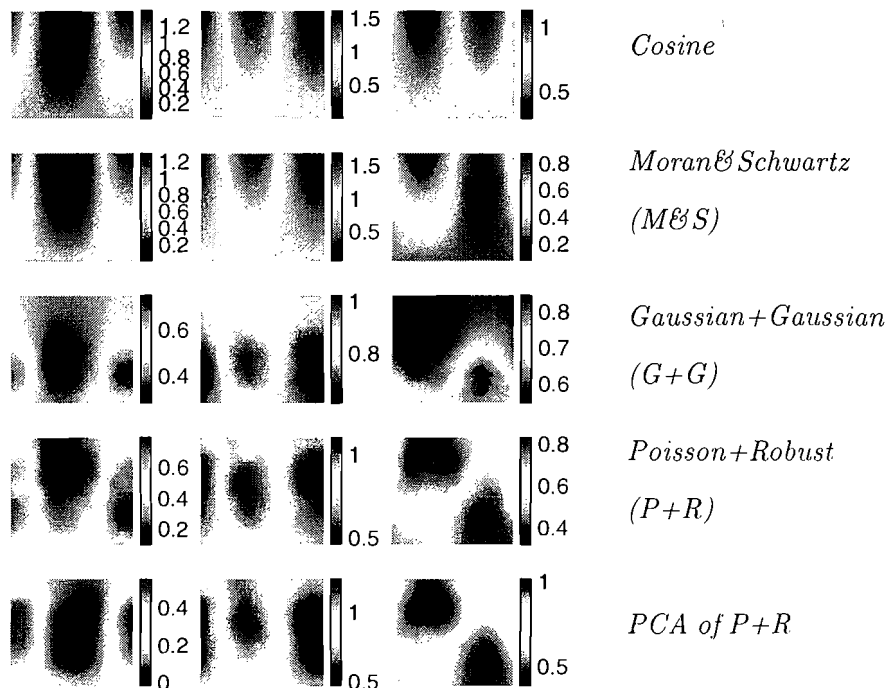


Figure 5: Estimated firing rate for cells in Figure 3 using different models.

3 Results

Various models (cosine, a modified cosine (Moran and Schwartz, 1999), Gaussian+Gaussian, Poisson+Robust and PCA of Poisson+Robust) are fit to the training data.

Observe that the pattern of firing is not Gaussian. Moreover, some cells appear to be tuned to motion direction, θ , and not to speed, r , resulting in vertically elongated patterns of firing. Other cells (e.g. cell 19) appear to be tuned to particular directions and speeds: this type of activity is not well fit by the parametric models.

The PCA yields surprisingly small amount of principle components, the first 4 eigenvalues count for about 80% of the variance, with the steepest descending rate between the first and the second ones.

The bases confirmed that neurons are coding both direction as well as speed. The first principle component is low speed tuning with no distinction on directions, the rest three are direction tuning with different phases.

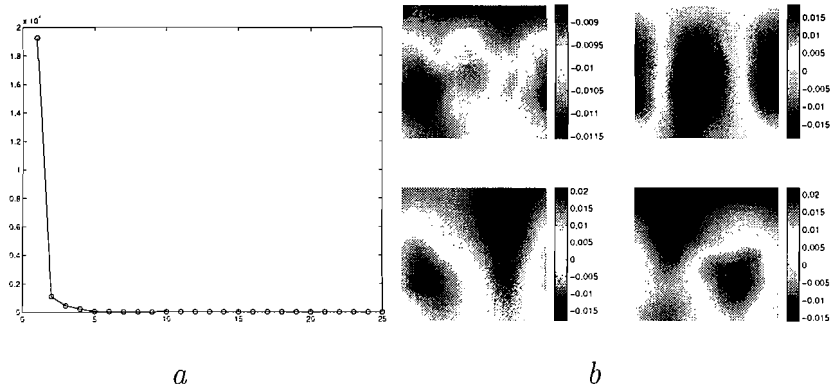


Figure 6: PCA of Poisson+Robust plot. (a) the eigenvalues; (b) the first four principle components.

<i>Method:</i>	Log Likelihood Ratio	p-value
G+G over Cosine	24.9181	7.6294e-06
G+G over M/S	15.8333	0.0047
P+R over Cosine	50.0685	7.6294e-06
P+R over M/S	32.2218	7.6294e-06
PCA of P+R over P+R	233.6086	7.6294e-06

Table 1: Log likelihood ratio of pairs of models and the significance level by Wilcoxon signed rank test

Table 1 shows a quantitative comparison using cross-validation. The positive values in Table 1 indicate that the non-parametric models do a better job of explaining new data than the parametric models with the Poisson+Robust fit providing the best description of the data, and PCA post-processing has a significant improvement. This P+R model implies that the conditional firing rate is well described by regions of smooth activity with relatively sharp discontinuities between them.

4 Discussion

The better fit of nonparametric models and the great improvement of PCA are not surprising. It can be well explained from the Bias/Variance (Geman *et al.*, 1992) point of view, which says the total error given a data set can be decomposed as bias and variance; bias is systematic error of the estimate, variance describes the dependence of the estimate on a particular realization of the data set. Here parametric models have strong bias of how the firing is coding movements and very little variance of the parameter fitting, hence they lose all the variabilities of different cells and different representations. By increasing the number parameters, such as Georgopoulos *et al.* (2000). it allows more variability, hence it will definitely do better. Nonparametric models are model-free. so no bias at all. but they have high variance. i.e. the model can fit the noise of training data. so different realizations can lead to very different estimate. The reason why nonparametric models are doing better than parametric models here is because they reduce the strong bias of the parametric models with a slight increase in variance and hence achieve a lower total error. PCA reduces the variance of nonparametric models without increasing much of the bias. so it increases the log likelihood significantly.

Besides data fitting. nonparametric models are saying the cells have much more variabilities than as described by the cosine model. Cells can be direction as well as speed tuned. or mixture of both. or even more variables which modeling is currently inhibited by the limited amount of data. The simple cosine tuning model is a good starting point but definitely not the end.

References

Black, M. and Anandan, P. (1996). The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields. *Computer Vision and Image Understanding*

63(1), 75–104.

Black, M. and Rangarajan, A. (1996). On the unification of line processes, outlier rejection, and robust statistics with applications in early vision. *International Journal of Computer Vision* **19(1)**, 57–92.

Blake, A. and Zisserman, A. (1987). *Visual Reconstruction*. The MIT Press, Cambridge, MA.

Donoghue, J., Sanes, J., Hatsopoulos, N., and Gaal, G. (1998). Neural discharge and local field potential oscillations in primate motor cortex during voluntary movements. *J. Neurophysiology* **79**, 159–173.

Fu, Q.-G., Flament, D., Coltz, J., and Ebner, T. (1995). Temporal encoding of movement kinematics in the discharge of primate primary motor and premotor neurons. *J. Neurophysiology* **73(2)**, 836–854.

Gao, Y., Black, M., Bienenstock, E., Shoham, S., and Donoghue, J. (2002). Probabilistic inference of arm motion from neural activity in motor cortex. *Advances in Neural Information Processing Systems 14*.

Geman, S., Bienenstock, E., and Doursat, R. (1992). Neural networks and the bias/variance dilemma. *Neural Computation* **4**, 1–58.

Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distributions and Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6(6)**, 721–741.

Geman, S. and McClure, D. (1987). Statistical methods for tomographic image reconstruction. *Bulletin of the Int. Stat. Inst.* **LII-4**, 5–21.

Georgopoulos, A. and Ashe, J. (2000). One motor cortex, two different views. *Nature Neuroscience* **3**, 963–963.

- Georgopoulos, A., Kalaska, J., Caminiti, R., and Massey, J. (1982). On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex. *J. of Neuroscience* **2**, 1527–1537.
- Georgopoulos, A., Schwartz, A., and Kettner, R. (2000). Directional tuning profiles of motor cortical cells. *Science* **36**, 73–79.
- Hatsopoulos, N., Ojakangas, C., Paninski, L., and Donoghue, J. (1998). Information about movement direction obtained from synchronous activity of motor cortical neurons. *Proc. Nat. Academy of Sciences* **95**, 15706–15711.
- Huang, J. and Mumford, D. (1999). The statistics of natural images and models. *IEEE Conference on Computer Vision and Pattern Recognition* 541–547.
- Maynard, E., Hatsopoulos, N., Ojakangas, C., Acuna, B., Sanes, J., Normann, R., and Donoghue, J. (1999). Neuronal interaction improve cortical population coding of movement direction. *J. of Neuroscience* **19(18)**, 8083–8093.
- McCullagh, P. and Nelder, J. (1989). *Generalized Linear Models*. Chapman & Hall, London, UK.
- Moran, D. and Schwartz, A. (1999). Motor cortical representation of speed and direction during reaching. *J. Neurophysiology* **82**, 2676–2692.
- Moran, D. and Schwartz, A. (2000). One motor cortex, two different views. *Nature Neuroscience* **3**, 963–963.
- Nowak, R. and Kolaczyk, E. (2000). A statistical multiscale framework for Poisson inverse problems. *IEEE Inf. Theory* **46(5)**, 1811–1825.
- Paninski, L., Fellows, M., Hatsopoulos, N., and Donoghue, J. (2001). Temporal tuning properties for hand position and velocity in motor cortical neurons. *submitted, J. Neurophysiology*.

- Scott, S. (2000). Reply to "one motor cortex, two different views". *Nature Neuroscience* **3**, 964–965.
- Szeliski, R. (1989). *Bayesian Modeling of Uncertainty in Low-level Vision*. Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Terzopoulos, D. (1986). Regularization of inverse visual problems involving discontinuities. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **8(4)**, 413–424.
- Todorov, E. (2000). Reply to "one motor cortex, two different views". *Nature Neuroscience* **3**. 963–964.
- Wessberg, J., Stambaugh, C., Kralik, J., Beck, P., Laubach, M., Chapin, J., Kim, J., Biggs, S., Srinivasan, M., and Nicolelis, M. (2000). Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. *Nature* **408**, 361–365.
- Woltring, H. (1986). A FORTRAN package for generalized, cross-validatorspline smoothing and differentiation. advances in engineering software. *Advances in Engineering Software* **8(2)**. 104–113.