

Is Neuron Discrimination Preprocessing Necessary for Linear and Nonlinear Brain Machine Interface Models?

Justin C. Sanchez

Department of Pediatrics, Division of Neurology
University of Florida
justin@cnel.ufl.edu

Jose C. Principe

Department of Electrical and Computer Engineering
University of Florida
principe@cnel.ufl.edu

Paul R. Carney

Department of Pediatrics, Division of Neurology
University of Florida
carnepr@peds.ufl.edu

Abstract

Translating brain-machine interface (BMI) research into the clinical setting will require neural decoding algorithms to be implemented in portable, low-power, wireless devices. BMI motor neurophysiological research for estimating hand position from neuronal firing patterns often involves decomposing the cortical activity down to individual neurons using a computationally intensive spike sorting procedure. From a signal processing implementation point of view, we would like to explore what neuron discrimination offers in terms of model decoding performance. In an attempt to reduce computational complexity, we tested the effect of using non-neuron discriminated (NND) data as inputs to common linear and nonlinear BMI models and statistically quantified the difference.

1 Introduction

Early electrophysiological techniques proposed by Adrian for investigating brain function were meticulously focused on the study of single neurons (Adrian 1965). As researchers sought to understand how neuronal activity encodes perception and behavior, the viewpoint from the *single* neuron expanded to encompass the time-dependent communication of ensembles of hundreds of neurons (Nicolelis 1999). With this paradigm shift, came an abundance of techniques to either mechanically manipulate electrodes to isolate single unit activity or sophisticated signal processing techniques to “sort out” the action potentials of single neurons. Again, the idea here in the academic and research settings was to understand brain function from a systems neurophysiologic perspective. In this research environment, it is critical to unequivocally report the rate modulation (peri-stimulus time histograms PSTH, etc.) from single units to develop principles of brain function from functional and environmental changes experienced by the organism. With the development of increased desktop computing power, researchers have employed elegant techniques such as principle component (PCA), wavelet, template, and Bayesian based clustering techniques to spike sort raw microelectrode recordings (Lewicki 1998). The aforementioned spike sorting techniques not only bring with them the ability to separate action potentials from noise but also offer the ability to discriminate the activity of one neuron from another. Recent advances in brain machine interfaces (BMI) have capitalized upon these neurophysiological techniques to demonstrate closed-loop neural control of computer and robotic devices with the single unit activity (Wessberg, Stambaugh et al. 2000; Serruya, Hatsopoulos et al. 2002; Taylor, Tillery et al. 2002). There is no question that these techniques have helped the motor systems community increase their understanding of ensemble neural activity for these tasks.

From a neural engineering perspective, we are beginning to enter another paradigm shift where brain machine interface technology is starting to enter the clinical setting (Donoghue 2004; Leuthardt, Schalk et

al. 2004). As engineers, we are faced with translating the tools and techniques of the research environment into medical devices that are efficient and robust in their tasks. One of the focuses of our BMI research is to develop low-power, wireless digital signal processors (DSP) that are capable of sampling neuronal activity, decoding intent, and transmitting data to an end effector. Moreover, we envision that many of the hardware components that will be utilized in BMI applications will be implemented as implantable devices where there are strict power and bandwidth limitations. With the current or even future (10 years) technology we may not have the computational luxury of spike sorting and discriminating the action potentials single neurons in real-time. The spike sorting approach is not only difficult from the computational side but also from the amount of human involvement where for accurate results spike detection thresholds need to be manually set, principal components need to be clustered, or wave shape templates need to be chosen (Nicolelis, Ghazanfar et al. 1997). It still remains unknown how BMI models will contend with a loss of either of the two experimental advantages that spike sorting offers:

- Action potential wave shapes can be used to filter out signal from noise (i.e. neuronal activity from movement artifact)
- The activity of individual neurons can be labeled and separated out for constructing ensemble time-series

If spike sorting is to be eventually phased out of the BMI experimental paradigm, we must quantify how distinguishing neuronal events (action potentials) from noise and neuronal discrimination influences predictive model performance.

The goal of this study is to first address only the latter issue by testing the effects of neuron labeling on brain machine interface model decoding performance. We would like to understand if neuron labeling provides any performance improvement in BMIs. From an adaptive signal processing modeling point of view, separating individual neurons from a given electrode may not intuitively provide a significant advantage since many of the commonly utilized input-output (I/O) models compute weighted sums of the neuronal data. Spike-sorting preprocessing separates individual neurons per electrode while I/O models re-collapse the data through the inner product calculations. Examples of models that utilize inner product matrix multiplications include the linear finite impulse response (FIR) filters and nonlinear neural networks both which have been extensively studied in the BMI literature (Wessberg, Stambaugh et al. 2000; Sanchez, Kim et al. 2002; Serruya, Hatsopoulos et al. 2002; Kim, Sanchez et al. 2003). In this study, we will evaluate the utility of neuron labeling by training and testing linear (Wiener filter - WF) and nonlinear (recurrent multilayer perceptron – RMLP) models using non-neuron discriminated (NND) and neuron discriminated (ND) neuronal firing patterns and quantify the performance.

2 Data Preparation

Multichannel neuronal firing times were collected synchronously at the University of Florida McKnight Brain Institute using male Sprague-Dawley rats. A total of thirty-two microwire electrodes were chronically implanted bilaterally (+1.0mm anterior, 2.5mm lateral of bregma) in the forelimb region of primary motor cortex (Donoghue and Wise 1982). Electrodes were configured in 2 x 8 arrays using 50 μ m polyimide insulated tungsten wire with 250 μ m separation within the 8 electrode row and 500 μ m spacing between the rows. The firing times of single neurons were recorded using a Tucker-Davis (Gainesville, Florida) Pentusa neural recording system sampling at 24414.1Hz while the rodents performed go no-go lever pressing task in an operant conditioning cage (Med-Associates, St. Albans, VT). The task involved an LED visual stimulus to press a lever with 23.55g per square centimeter of force for a minimum of 0.5s to achieve a water reward. Upon achieving the behavioral criteria (lever choice and press time), a 1kHz tone and a 0.04ml water reward will be delivered to the animal providing feedback that the task has been achieved. The lever position, used as the models' desired signal, was also recorded (with a time shared clock) and digitized with a 384.1Hz sampling rate.

A combination of thresholding and template based spike-sorting techniques yielded neuronal spike firing times from 78 neurons collected from the 32 chronic microwire electrodes. The neuronal firings

were binned (added) in non-overlapping windows of 100 ms, which represents the local firing rate for a neuron. The NND firing patterns were reconstructed by summing the firing rates of the neurons associated with each of the 32 electrodes. For example, if electrode 7 has neurons A, B, C, and D associated with it, the NND sorted firing rate would be the sum of the firing rates for neurons A, B, C, and D at each time bin. This methodology assumes a fixed threshold spike sorting implementation. Data was prepared in this manner to decouple the two advantages of spike sorting. By summing the bin counts we can be confident that noise has been removed and we are only studying the effects of neuron discrimination. We present in Figure 1 an example time-series from electrode #7 for this procedure. For this particular electrode, four action potential waveforms could be discriminated indicating that the tip of the electrode was near three neurons (Neuron 18, 19, 20, 21 in the figure). Summing the firings of each of these neurons will yield Figure 1 subplot 1 which represents the ND data. As we can see, not many of the firings overlap, therefore, the amplitude of the NSS signal does not change significantly. We computed the correlation coefficient between all pairs of neurons ($CC_{18,19} = 0.0544$, $CC_{18,20} = -0.1745$, $CC_{18,21} = 0.3228$, $CC_{19,20} = 0.0043$, $CC_{19,21} = 0.0145$, $CC_{20,21} = -0.0584$) and found that in this example that only two neurons were weakly correlated. The main change in the NND time-series is the degree of sparsity. For the entire ensemble of cells, the number of inputs is collapsed from 78 neurons to 32 electrodes. These spike counts from the ND and NND data were directly used as inputs to the linear and nonlinear models.

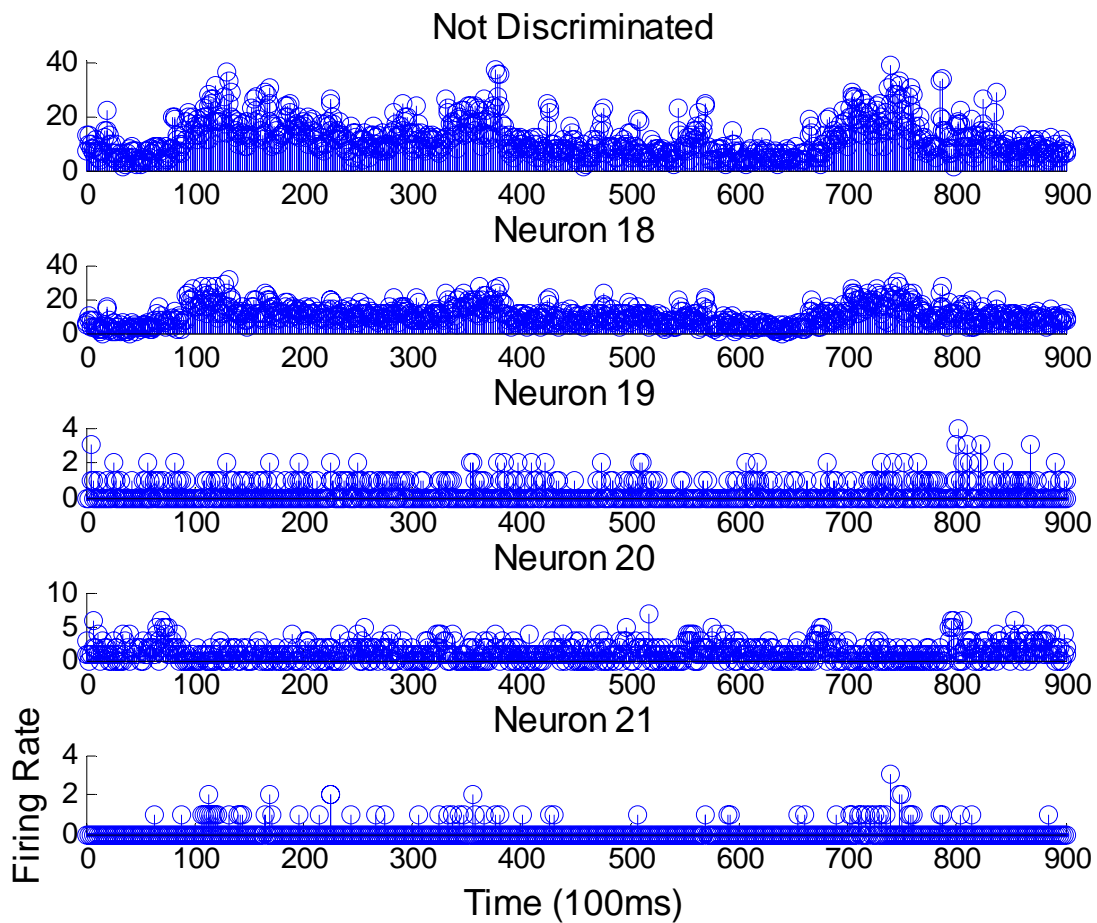


Figure 1. Example NND and ND time-series for electrode #7.

3 Model Topology

Figures 2 and 3 show the topology of WF and RMLP that are used in our studies. Both topologies were trained with 10 minutes (6000 samples) of neuronal and behavioral recordings. For the WF, the output is a weighted linear combination of neuronal inputs \mathbf{x} . The WF utilizes a 10-tap delay line (most recent 1 second of neural activity) to compute each output. The vector form of this operation is given by (1) (Sanchez, Carmena et al. 2003). The optimal MSE solution is given by (2), where \mathbf{d} is the hand trajectory.

$$\mathbf{y}(t) = \mathbf{W}\mathbf{x}(t) \quad (1)$$

$$\mathbf{W} = \mathbf{R}^{-1}\mathbf{P} = E(\mathbf{x}^T \mathbf{x})^{-1} E(\mathbf{x}^T \mathbf{d}) \quad (2)$$

The RMLP uses only the instantaneous neural activity to compute each output. The hidden layer consists of 5 *tanh* processing elements (PEs) that are fully connected to each other with a feedback matrix. The number of hidden PEs used in each topology was optimized to produce the best testing performance with the smallest number of degrees of freedom (Sanchez, Erdogmus et al. 2002). The state vector (7 states) of the hidden layer in (3) is a nonlinear function of the linear combination of input and previous state. The feedback of the state creates memory and allows representations on multiple timescales. The output layer has 1 sigmoidal PE and produces the lever position output as in (4). Each hidden PE is a nonlinear adaptive basis for the output that projects the high dimensional neuronal data. These projections are then linearly combined to form the outputs (position predictions) of the RMLP. Optimal weights are determined by minimizing MSE using BPTT (Haykin 1994). Training was stopped using the method of cross-validation (batch size of 1000 pts.) to maximize the generalization of the network (Vapnik 1999). Even though the RMLP is a parsimonious model, there still are a significant number of weights in the input layer. To further improve the generalization of this model we implement a regularization technique (weight decay) in the input layer by subtracting a small constant (0.00001) from each weight at each update (Príncipe, Euliano et al. 2000). This procedure ensures that weights that are not being significantly updated will be reduced to zero. Each RMLP for NND and ND data was trained through 50 Monte Carlo simulations and the network that produced the smallest training mean square error (MSE) was selected throughout this study.

$$\mathbf{y}_1(t) = f(\mathbf{W}_1\mathbf{x}(t) + \mathbf{W}_f\mathbf{y}_1(t-1) + \mathbf{b}_1) \quad (3)$$

$$\mathbf{y}_2(t) = \mathbf{W}_2\mathbf{y}_1(t) + \mathbf{b}_2 \quad (4)$$

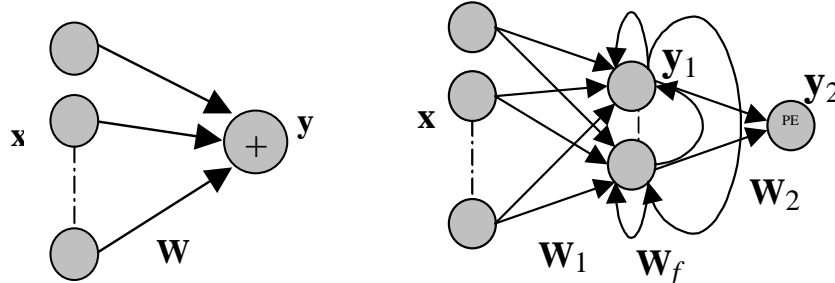


Figure 2. Wiener filter topology. Figure 3. Fully connected, state RMLP.

4 Performance Results

Once each of the linear and nonlinear topologies was trained with NND and ND data, we fixed the model parameters and presented 10 minutes of novel neural and behavioral data. During this test segment each model topology predicted the lever position trajectory that included eleven consecutive presses. Model performance was quantified by computing both the probability of predicting a lever press when the animal was actually pressing and the probability of predicting a release when the animal was not pressing the lever. The probabilities were computed after applying a hard threshold to the output time series. For each model, the threshold level was set by iteratively searching for the value that simultaneously

maximized the probabilities of pressing and releasing prediction. Our results shown in Table 1 indicate that for the linear model the testing CC values decreased for NND input data with pressing taking the larger decrease of 7%. In terms of the nonlinear model, the probability of correctly predicting pressing or releasing remained unchanged for the NND data. We statistically verified the difference in model performance between NND and ND data using a two-sample T-test that tested the hypothesis that the two independent pressing or releasing test vectors come from distributions with equal means (“null hypothesis”). A T-test value of 0 indicates that the means are equal with 95% confidence. The value 1 indicates that the null hypothesis can be rejected. In this study with the models and behaviors chosen, choosing to not discriminate neurons in the data did produce a significant decrease in BMI pressing performance for the linear model during pressing. For the RMLP nonlinear model, choosing to not discriminate neurons did not statistically affect BMI performance in pressing or releasing.

We qualitatively compare the numerical results with the testing trajectory traces for two representative reaching movements in Figure 4. In general, the ND nonlinear RMLP produces the best trajectory in terms of false movements and timing of capturing the edges of the pressing pulses. For the first press the ND RMLP tended to undershoot the movement while on the second press the NND RMLP extended the press. The NND linear trajectory contained many false presses combined with anticipatory, undershooting, and overshooting the trajectory compared with the ND linear model output.

Table 1. Model Testing Performance.

	Linear - WF		Nonlinear - RMLP	
	ND	NND	ND	NND
$P_p(O_{press} \lambda_{press})$	0.93	0.86	0.93	0.93
	ttest = 1		ttest = 0	
$P_r(O_{release} \lambda_{release})$	0.93	0.90	0.93	0.93
	ttest = 0		ttest = 0	

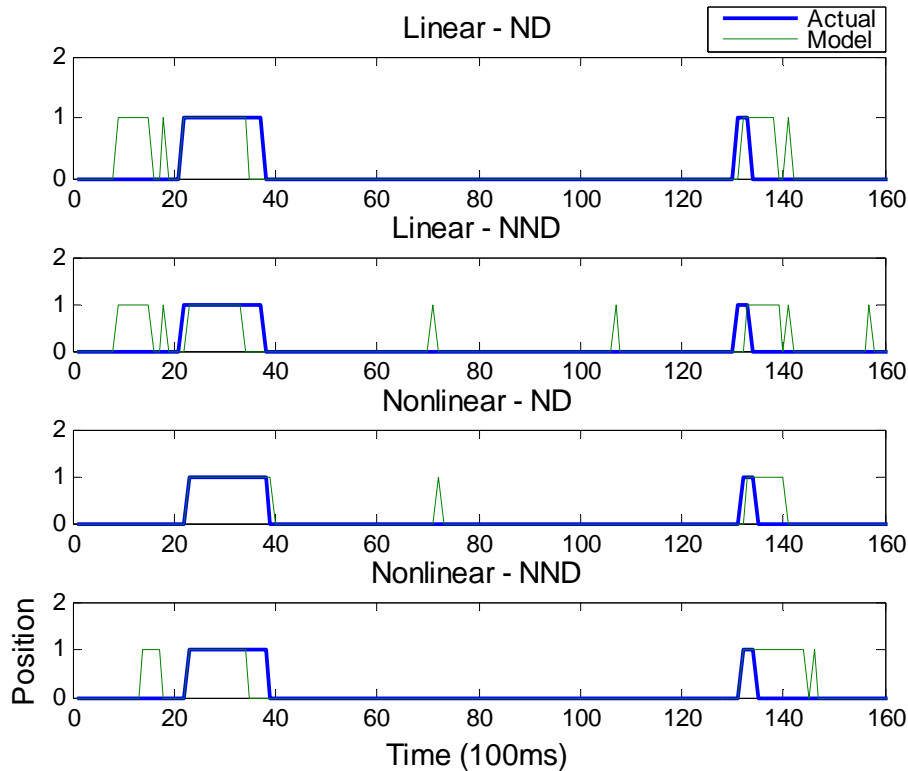


Figure 4. Testing trajectories for two lever presses.

Since we are interested in how the aggregate of NND neurons affects model performance we analyzed the magnitude of the weights in each of the models. Interestingly, both linear and nonlinear models trained with NND place large weight values on electrode #7. Moreover, neurons 18, 20, and 21 were assigned large values during training in both the linear and nonlinear models. For this particular electrode, we infer that from the model's perspective it is irrelevant if neurons are labelled in the time series as long as the temporal modulation of the activity is preserved in the NND data.

5 Discussion

The vast majority of BMI papers for estimating hand position from neuronal firing patterns includes decomposing the cortical activity to individual neurons. Neuronal spike-sorting is a computationally intensive procedure that often involves extensive human discretion and even error. If BMIs are to be feasibly implemented in clinically robust portable devices, solving the spike sorting issues is going to be a daunting task. For this reason, we explored the effect of neuron discrimination on common linear and nonlinear BMI model performance. Overall we found that neuron discrimination does not have a profound effect on model decoding performance. For nonlinear models the performance remains unchanged while for linear models the performance dropped at most 7%. Of course, this drop in performance must be considered against the fact that the linear and nonlinear approaches have a different number of model parameters (Linear – ND 780, NND 320, Nonlinear – ND 602, NND 280) which can affect generalization. Experimentally we would like to state up front that our data preparation in this study assumes that useful neuronal events (action potentials) related to behaviour have already been selected from the noise in the signal. The next step of our studies is to perform the modelling analysis with only a single threshold applied to the raw data to select events which will be used as inputs. Preliminary studies have shown that lever position prediction can still be achieved with high probability without any spike sorting whatsoever. However, we have encountered a trade-off between the threshold level, number of false presses, and number of missed presses especially when the SNR of the raw neuronal recording is low. Therefore, we are researching BMI architectures that can first effectively extract relevant features from noise and then perform the prediction.

From a signal processing point of view, we need to ask ourselves what neuron discrimination *provides* for model performance. At the most fundamental level, spike sorting isolates unit activity into a single model input that is assigned a weight. If during the optimization procedure the neuron's activity does not correlate with behavior, its weight can be kept at a small value so that changes in the input do not affect the output. To put it simply, with ND data the model has the ability to "shut off" noisy neurons. Interestingly, in the results presented here, the ability to "shut off" noisy neurons did not provide significant benefit to the performance in the RMLP and only marginal benefit to the Wiener filter. So, either there were not many noisy neurons in the ensemble or only a few of the neurons were relevant for behavior and were assigned large weights. We have performed extensive analysis on the importance of subsets of neurons for the models presented here and have found that only about 1/4 of ensemble of cells is critically important in constructing the mapping (Sanchez, Carmena et al. 2003). Therefore, when we collapse the spike sorted data back down to the electrode we are only creating a fewer number of noisy channels what will not have their weights updated anyway.

For implementation in low-power hardware the use of NND provides an advantage in the amount of processing that needs to be performed locally at the recording site. At that location, the only preprocessing is a spike detection routine that has the capability of picking out spikes from noise and movement artifacts. Locally we can also implement a binning procedure and transmit wirelessly only the firing rates for each electrode. Such an implementation would require maximally 8 bits of resolution per electrode.

In summary, by sampling the firing information from a neighborhood of neurons we are still able to capture an equivalent amount of information while gaining advantages. The reduction of the dimensionality drastically decreases network training time, data requirements, and preprocessing

computational complexity. The results presented in this analysis indicate that for the task of predicting hand/arm pressing trajectories from neuronal firing patterns neuron discrimination is not a requirement for maintaining high performance. Removing the discrimination procedure simplifies the BMI architecture and increases the feasibility of implementation in low-power wireless hardware. This simplification however comes at the cost of additional responsibility for accurate selection of neuronal events from noise.

Acknowledgement

This work was supported by the McKnight Brain Institute: Brain Spinal Cord Injury Trust Fund.

References

- Adrian, E. (1965). Nobel Lectures, Physiology or Medicine 1922-1944. Amsterdam, Elsevier.
- Donoghue, J. P. (2004). From Mind to Movement: Developing Neurotechnologies to Restore Lost Function. Neural Information Processing Systems, Whistler, B. C.
- Donoghue, J. P. and S. P. Wise (1982). "The motor cortex of the rat: cytoarchitecture and microstimulation mapping." J. Comp. Neurol. **212**(12): 76-88.
- Haykin, S. (1994). Neural networks: a comprehensive foundation. New York: Toronto, Macmillan; Maxwell Macmillan Canada.
- Kim, S. P., J. C. Sanchez, et al. (2003). "Divide-and-conquer approach for brain machine interfaces: nonlinear mixture of competitive linear models." Neural Networks **16**(5-6): 865-871.
- Leuthardt, E. C., G. Schalk, et al. (2004). "A brain-computer interface using electrocorticographic signals in humans." Journal of Neural Engineering **1**: 63-71.
- Lewicki, M. S. (1998). "A review of methods for spike sorting: the detection and classification of neural action potentials." Network: Computation in Neural Systems **9**(4).
- Nicolelis, M. A., A. A. Ghazanfar, et al. (1997). "Reconstructing the engram: simultaneous, multisite, many single neuron recordings." Neuron **18**(4): 529-537.
- Nicolelis, M. A. L. (1999). Methods for Neural Ensemble Recordings. Boca Raton, CRC Press.
- Principe, J. C., N. R. Euliano, et al. (2000). Neural and adaptive systems: fundamentals through simulations. New York, Wiley.
- Sanchez, J. C., J. M. Carmena, et al. (2003). "Ascertaining the importance of neurons to develop better brain machine interfaces." IEEE Transactions on Biomedical Engineering **61**(6): 943-953.
- Sanchez, J. C., D. Erdogmus, et al. (2002). A comparison between nonlinear mappings and linear state estimation to model the relation from motor cortical neuronal firing to hand movements. SAB Workshop on Motor Control in Humans and Robots: on the Interplay of Real Brains and Artificial Devices, University of Edinburgh, Scotland.
- Sanchez, J. C., S. P. Kim, et al. (2002). Input-output mapping performance of linear and nonlinear models for estimating hand trajectories from cortical neuronal firing patterns. International Work on Neural Networks for Signal Processing, Martigny, Switzerland, IEEE.
- Serruya, M. D., N. G. Hatsopoulos, et al. (2002). "Brain-machine interface: Instant neural control of a movement signal." Nature **416**: 141-142.
- Taylor, D. M., S. I. H. Tillery, et al. (2002). "Direct cortical control of 3D neuroprosthetic devices." Science **296**(5574): 1829-1832.
- Vapnik, V. (1999). The Nature of Statistical Learning Theory. New York, Springer-Verlag.
- Wessberg, J., C. R. Stambaugh, et al. (2000). "Real-time prediction of hand trajectory by ensembles of cortical neurons in primates." Nature **408**(6810): 361-365.