

Statistical Encoding Model for a Primary Motor Cortical Brain-Machine Interface

Shy Shoham*, *Member, IEEE*, Liam M. Paninski, Matthew R. Fellows, Nicholas G. Hatsopoulos, John P. Donoghue, *Member, IEEE*, and Richard A. Normann, *Member, IEEE*

Abstract—A number of studies of the motor system suggest that the majority of primary motor cortical neurons represent simple movement-related kinematic and dynamic quantities in their time-varying activity patterns. An example of such an encoding relationship is the cosine tuning of firing rate with respect to the direction of hand motion. We present a systematic development of statistical encoding models for movement-related motor neurons using multielectrode array recordings during a two-dimensional (2-D) continuous pursuit-tracking task. Our approach avoids massive averaging of responses by utilizing 2-D normalized occupancy plots, cascaded linear-nonlinear (LN) system models and a method for describing variability in discrete random systems. We found that the expected firing rate of most movement-related motor neurons is related to the kinematic values by a linear transformation, with a significant nonlinear distortion in about 1/3 of the neurons. The measured variability of the neural responses is markedly non-Poisson in many neurons and is well captured by a “normalized-Gaussian” statistical model that is defined and introduced here. The statistical model is seamlessly integrated into a nearly-optimal recursive method for decoding movement from neural responses based on a Sequential Monte Carlo filter.

Index Terms—Discrete distribution, LN model, neural decoding, neuroprosthetics, sequential Monte-Carlo.

I. INTRODUCTION

THE study of neurons in the primary motor cortex (area MI) and their role in the motor control system has strongly relied on two complementary strategies, dictated by experimental constraints as well as diverging views regarding what information is represented by individual neurons. According to the first approach [2], [3] an emphasis is placed on a mechanistic model of the motor control system, in which the contribution of individual neurons is aggregated and transformed to create the motor control signal. Implicit in this approach is the view that

the relation between the activity of individual neurons and the motor output is rather complex, as dictated by the nonlinear, state dependent transformation that leads from motor cortical firing to movement. According to the second approach, the activity of individual neurons can be viewed as representing simple kinematic and dynamic features of the planned or executed motor output. A prime example of this approach is the cosine-tuning curve relating the average firing rate to the direction of movement [1]. The two views are not contradictory [4], [5] and may actually lead to similar models of the motor system. Rather, their main difference is in terms of the emphasis on different functional models “movement = f (activity)” versus “activity = f (movement).” The evolution of models of the second type “activity = f (movement)” in the context of the motor system, is partly a result of experimental constraints that limited the range and nature of movements that could be effectively experimentally controlled or measured. These models, which are closely related to models of sensory and cognitive representation processes, are called encoding models. A complete encoding model of single neurons would capture the way in which the instantaneous firing rate is modulated by the kinematic movement variables, as well as the variability (or noisiness) of this firing [6]–[8]. A more general model will also include interneuronal interactions. Decoding strategies (e.g., the population vector [9] or other methods [10]–[12]) as well as mechanistic models [13] complement encoding models by attempting to explain how the motor system pools together the single-cell cortical activity into a movement representation.

In this paper, we develop an encoding model for individual MI neurons using continuous simultaneous recordings of hand position (in two dimensions) and neural activity during tracking of a randomly moving target. In this experimental scenario, a number of kinematic quantities that may be related to expected firing rate (e.g., position, velocity, and acceleration) are controlled so that they are well sampled within the workspace, approximately stationary over trial time, and are minimally interdependent. However, the time-varying and random, nonrepeated nature of the task makes the encoding relationship challenging to measure, as no averaging across different trials of the same type is possible. In order to characterize the encoding using this dataset, we first introduce nonparametric analyses that allow us to explore the effect of kinematic variables (two at a time) on the expected firing rate. Having demonstrated that the two-dimensional (2-D) projections of the encoding relationship are roughly linear, we use a linear-nonlinear cascade system representation (Wiener cascade) that effectively represents the combined effect of all the kinematic variables. Using the model for expected firing rate we then study the associated neural variability using

Manuscript received June 2, 2004; revised December 5, 2004. The work of S. Shoham was supported in part by a Lewis-Thomas fellowship. The work of J. P. Donoghue was supported by the National Institutes of Health (NIH) under Grant R01NS25074. The work of R. A. Normann was supported in part by the State of Utah Center of Excellence under Contract 95-3365. Asterisk indicates corresponding author.

*S. Shoham is with the Faculty of Biomedical Engineering, the Technion, Israel Institute of Technology, Haifa 32000, Israel (e-mail: sshoham@bm.technion.ac.il).

L. M. Paninski is with the Gatsby Computational Neuroscience Unit, London, WC1N 3AR, U.K.

M. R. Fellows and J. P. Donoghue are with the Neuroscience Department, Brown University, Providence, RI 02912 USA.

N. G. Hatsopoulos is with the Department of Organizational Biology and Anatomy, University of Chicago, Chicago, IL 60637 USA.

R. A. Normann is with the Bioengineering Department, University of Utah, Salt Lake City, UT 84112 USA.

Digital Object Identifier 10.1109/TBME.2005.847542

novel “neural noise” plots, comparing Poisson and an alternative, “normalized-Gaussian” distribution that we develop.

Our study of encoding in the primary motor system is motivated by an attempt to develop effective decoding algorithms for converting spike trains recorded from multiple motor neurons into the motor control signal that they represent—an essential component in an implantable brain-machine interface [14], [15]. Such algorithms may improve the applicability of these devices, possibly overcoming the hardware constraints in the current generation of implantable microelectrode arrays which limit the number of practically available units to several dozens. Optimal decoding approaches, which can be viewed as optimal estimation solutions, ultimately rely on a statistical encoding model [6], like the one developed here. To demonstrate the utility of our encoding model, we implement a recursive Bayesian filter that is optimally tailored to it using the recently developed method of Sequential Monte Carlo filters [16]–[18]. An earlier version of this work has previously appeared in S. Shoham’s Ph.D. dissertation [17].

II. EXPERIMENTAL METHODS

A. Recording Setup

Two macaque monkeys (*M. mulatta*) were operantly conditioned to perform a visually guided manual tracking task (described below). While performing this behavior extracellular signals were recorded with a chronically implanted microelectrode array [19] (Cyberkinetics, Foxborough, MA). The array was implanted in the arm region of the monkeys’ precentral gyrus (primary motor cortex), following a training period lasting several months. The surgical implantation procedures are described elsewhere [20]. The array consists of one hundred 1.0-mm-long silicon electrodes with platinized tips (200–500 k Ω impedances measured with a 1 kHz, 100 nA sine wave) arranged in a rectangular 10 \times 10 grid (0.4-mm spacing). The electrode tips were approximately located in layers III and V. Additional details of the surgical and animal care procedures appear elsewhere [21]. A number of long-term studies have established the chronic recording capability of this electrode-array [22], [23], and an ongoing research effort is underway to establish its efficacy in an implantable brain-machine interface [14], [24]. The procedures were approved by Brown University’s Animals Care and Use Committee.

The neural signals were bandpass filtered (250–7500 Hz, 5th order Butterworth), amplified (5000 \times), digitized (30 kHz sampling), and acquired on a Pentium based PC using a 100-channel data acquisition system [25] (Bionic Technologies LLC., Salt Lake City, UT). Thresholds were manually set, and threshold-crossing events were saved to disk (1.6-ms duration). Multiunit data was sorted off-line using a new automatic and noise-robust spike-sorting algorithm [26].

In addition to the neural signals, the x and y position of a two-link manipulandum were digitized by a digitizing tablet (Wacom Technology Corp., Vancouver, WA) at 167 Hz with an accuracy of 0.25 mm (range: 10 \times 10 cm) and saved to disk. These measurements were interpolated to a 1-ms resolution using a smoothing cubic spline (Spline Toolbox, Mathworks, Natick MA). The smoothing spline was also used to analytically calculate the derivatives, in order to avoid large discretization-related

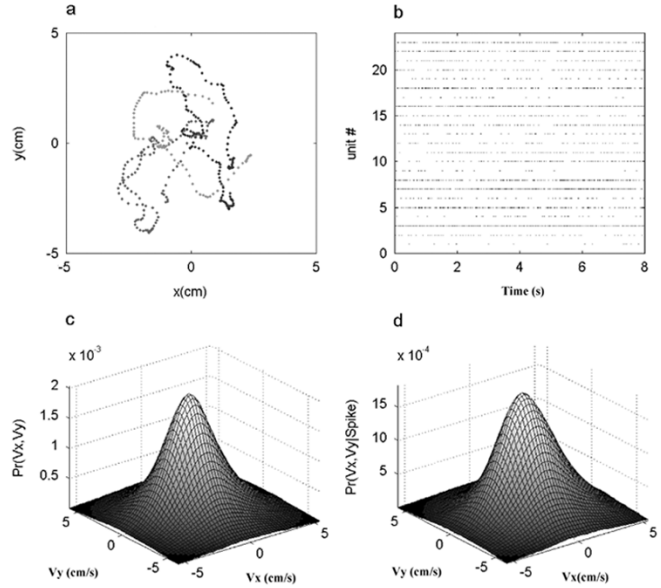


Fig. 1. General data characteristics. Data shown from one session where 23 well-isolated units were recorded during 172 tracks that exceeded 4 s. (a) Movement traces during three successive 8-s-long tracks. Dots represent snapshots once every 50 ms. Tracks are marked by different shades of gray. (b) Raster plot of unit activity during a successful track. (c) Smoothed distribution of movement velocities during the entire recording session. (d) Smoothed velocity histogram conditioned on spikes occurring in unit #7 (with time shift of 125 ms). (a) Hand position traces, (b) spike rasters, (c) velocity distribution, (d) conditional velocity distribution.

artifacts, and the smoothing parameter was chosen manually to produce smooth derivatives with only negligible smoothing of the position time-series.

B. Behavioral Task

The monkeys were operantly conditioned to use a low-friction manipulandum to track a target moving on a computer screen in order to receive juice rewards. The target movement was generated by low-pass filtering a pseudo-random Gaussian-distributed sequence, yielding a time series with a spectral cutoff set at 0.4–0.5 Hz [10-dB drop from maximal power, see also Fig. 11(c) and (d)]. A small black cursor on the screen indicated the manipulandum location, and the monkey’s task was to keep it within the boundaries of the smoothly moving red target circle (1.5-cm radius) for a duration of 8–10 s. Only trials where the monkey acquired the target within a limited time window (1.5 s for one of the monkeys and 4 s for the second) and did not lose it for the entire track duration were considered successful and led to a reward. However, we analyzed data from all trials where the tracking duration exceeded 4 s, including unsuccessful ones.

Additional details of this experiment and additional characteristics of the results are reported in [27]. Due to the experimental design, the correlations between the time series of pairs of different kinematic variables (i.e., hand position in x and y , hand velocities and accelerations) were very small (-0.01 ± 0.06 , mean \pm SD). Fig. 1 illustrates the basic features of the experimental data.

Nomenclature: In the following discussion and derivations we will use \mathbf{x}_t to denote the “state” of the arm at time t . In general, the “state” of the arm can include a large number of

kinematic or dynamic variables, however, as we only measured the manipulandum position (in two dimensions), \mathbf{x}_t will refer to the position and its derivatives (velocity, acceleration...). The number of spikes neuron i fired in the time bin that followed time t is denoted N_t^i . We will denote by λ_t^i the instantaneous firing rate (also known as the stochastic intensity [28]): $\lambda_t^i = P(N_t^i = 1)/\Delta t$ in the limit $\Delta t \rightarrow 0$. In calculating the nonparametric encoding plots we were interested in the conditional probabilities of single spikes, which led us to choose very short time bins $\Delta t = 1$ ms. All the subsequent analyses are based on spike counts, and a 50-ms time bin was chosen to allow for decoding algorithms with sufficient temporal resolution for smooth movement tracking but simple enough to allow for a real-time implementation. A typical average firing rate for the cortical neurons we study is ~ 25 Hz (generally $\lambda < 100$ Hz), and more than 99% of the time bins contained 0–4 spikes. Braces $\langle \rangle$ indicate an average value.

III. EXPERIMENTAL METHODS

In order to understand the general encoding relationship, a model-free method for visualizing how the probability of firing depends on kinematic parameters is required. Previous characterizations of motor encoding (e.g., [1], [29], and [30]) relied heavily on neural responses integrated across extended time-windows or averaged across a large number of similar trials to achieve continuous “tuning curves.” In the context of hippocampal place cells, a simple application of Bayes’ rule was used to describe encoding (or ‘normalized occupancy’) relationships in two dimensions [31]. The basic formula

$$P(N_{t-\tau}^i = 1 | \mathbf{x}_t) = \frac{P(\mathbf{x}_t | N_{t-\tau}^i = 1) P(N_{t-\tau}^i = 1)}{P(\mathbf{x}_t)} = \frac{P(\mathbf{x}_t | \text{spike at } t - \tau)}{P(\mathbf{x}_t)} \cdot \langle \lambda_{t-\tau}^i \rangle \Delta t. \quad (1)$$

can be directly used to explore $P(N_t^i = 1 | \mathbf{x}_t)$, the probability of firing a spike given that the kinematic parameters τ seconds later are \mathbf{x}_t . Given an experimental data trace, placing a point on a discrete grid of the kinematic values \mathbf{x}_t every time the neuron spiked provides an approximation of the density $P(\mathbf{x}_t | N_t^i = 1)$ [Fig. 1(d)]. Placing points on the same grid at every time point (irrespective of whether the cell fired an action potential or not) approximates the density $P(\mathbf{x}_t)$ [Fig. 1(c)]. Dividing these histograms yields an approximate representation of the encoding relationship $P(N_t^i = 1 | \mathbf{x}_t)$.

Plots were created for the position, velocity, and acceleration (two dimensions in each case). In making the plots we tried to focus on regions with a sufficient amount of data, and chose boundaries that included the 1st to 99th percentiles of the state space in each dimension, dividing them into 50 bins. The unnormalized plots were smoothed with a Gaussian kernel ($\sigma = 3$ bins, full width at half maximum = 7.1 bins) and regions where the value of the smoothed unnormalized occupancy histogram [Fig. 1(d)] fell below 1, were rejected. The time delay τ was set at 125 ms. A previous study described the activity of the vast majority of MI neurons as preceding movement by

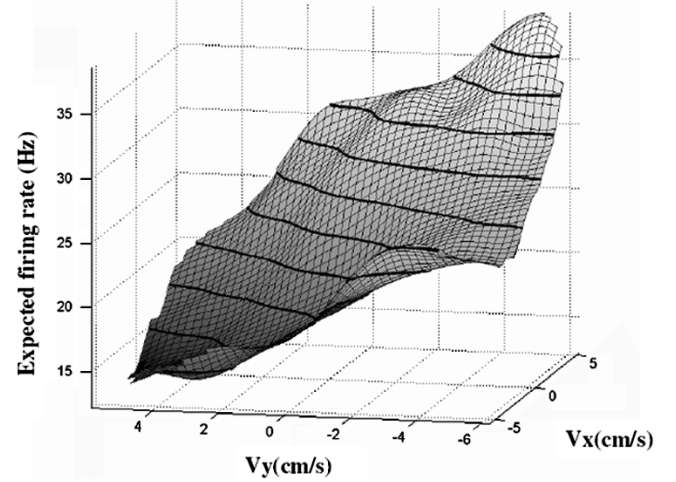


Fig. 2. Velocity-conditional firing rate expressed using nonparametric plots. Contour lines appear at regular intervals. Parallel lines indicate a planar structure.

0–250 ms [29] (roughly uniform distribution), which led to the selection of this number (see also [32] for a similar result). No attempt was made to find individual-neuron delays, however, we verified these results were fairly insensitive to small changes in τ (of order 50 ms).

Nonparametric plots for the encoding of position, velocity, and acceleration in two representative units are shown in Figs. 2 and 3. For the purpose of illustration we selected the two units with the highest signal-to-noise ratio from one recording session. Apart from some high-frequency distortion near the edges, the encoding plots appear as monotonic, distorted planes (note that contour lines for a plane are parallel and equidistant). Comparing the dynamical range on the different plots in Fig. 3 demonstrates that unit 1 (left panels) encodes velocity strongly (firing rate modulated between 10–35 Hz, a range twice as wide as for other variables), whereas unit 2 (right panels) encodes primarily position (firing rate modulated between 7–20 Hz, a range twice as wide as for other variables). However, in neither case is the encoding exclusively for velocity or position. The plots appear planar with respect to the velocity in unit 1 and position in unit 2 (see [27] for a quantitative analysis of the linearity in the nonparametric plots).

These results point to one of the weaknesses of this technique, when the combined contribution to the firing rate of all of the different kinematic variables needs to be evaluated. If the relationships were perfectly linear, we would expect the combined contribution to have this form

$$\lambda_{t-\tau}^i = c_0 + c_1 x_t + c_2 y_t + c_3 \dot{x}_t + c_4 \dot{y}_t + c_5 \ddot{x}_t + \dots \quad (2)$$

However, how to pool together the encoding functions from several nonlinear 2-D plots is not as straightforward. Moreover, relying on a single time delay τ may distort the plots, particularly in the acceleration plots (acceleration being the most rapidly changing signal) [33]. In order to account for nonlinearities in the encoding of multiple parameters, as well as effects occurring at multiple temporal delays, we turn from the fully nonparametric models introduced in this section to a

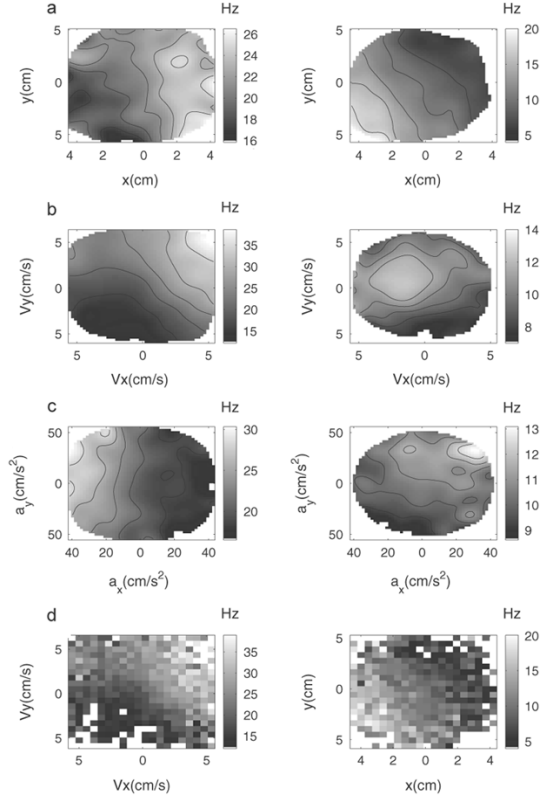


Fig. 3. Nonparametric encoding plots of different kinematic variables. The color scale is in spikes/s units. The left and right panels are for different units [left is the same unit as in Figs. 1(d) and 2]. (a) Position encoding. (b) Velocity encoding. (c) Acceleration encoding. Equidistant contour lines are superposed. (d) nonparametric encoding plots calculated without any smoothing (compare left panel with b-left and right panel with a-right). Lines indicate a planar structure.

nonlinear generalization of the general linear encoding model [which generalizes (2)].

IV. LN ENCODING MODELS

Models comprised of a cascade of a linear filter $\mathbf{h}^i(\tau)$ followed by a general static nonlinearity f , also known as LN- or Wiener-cascade models, have been explored by a number of studies (e.g., [34]–[36]), and form a natural extension to linear system models. These models appear well suited for describing the properties of neurons in the early visual system. In the case of the motor system, since firing tends to precede movement, encoding relationships are usually anti-causal, and can take the form

$$\lambda_t^i = f^i \left(\int_0^T \mathbf{h}^i(\tau) \mathbf{x}_{t+\tau} d\tau \right) \stackrel{\text{Discrete time}}{=} f^i \left(\sum_{k=1}^l \mathbf{h}^i(k) \mathbf{x}_{t+k} \right). \quad (3)$$

In order to obtain $\mathbf{h}^i(k)$ and f^i from our data, we used a simple two-step method similar to the one described in [36]. In the first step standard linear regression of the position (x and y) time series versus the neural response time series $N_{t-\Delta t, t}^i$ is used to determine $\mathbf{h}^i(k)$ (the regression coefficients). Given $\mathbf{h}^i(k)$ it is straightforward to obtain a linear prediction time

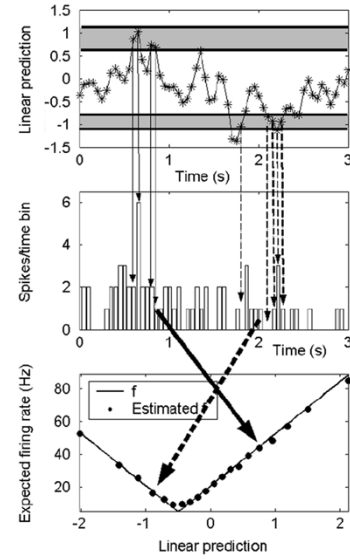


Fig. 4. Illustration of binning method used for estimating LN models from discrete neural data (simulated data). Top: The linear expectation time series \hat{u}_t^i is broken into bins according to the linear prediction value (2 shaded bins illustrated). Center: The spike counts time series is segmented accordingly (solid and dashed arrows point to time points corresponding to the two bins). Bottom: f is evaluated from the relationship between linear prediction and mean firing rate in the corresponding time points. The dashed and solid arrows are pointing at two points that correspond to each of the two bins.

series $u_t^i = \sum_{k=1}^l \mathbf{h}^i(k) \mathbf{x}_{t+k}$ using a convolution operation, and u_t^i can now be used as a basis for evaluating the nonlinearity f^i . This is done using a simple binning approach: the time axis is broken into subsets where u_t^i falls into different bins $u_j - \Delta u_j/2 < \hat{u}_t^i < u_j + \Delta u_j/2$. This nonparametric aggregation of ‘related’ time bins from different parts of the time axis is illustrated in the upper two panels of Fig. 4. The relationship of u_j and the mean number of spikes $\langle N_t^i(j) \rangle$ in the corresponding subset of time bins approximately corresponds to the nonlinearity f^i , as is illustrated in the bottom panel of Fig. 4 ($\langle N_t^i(j) \rangle \cong f^i(u_j)$). To obtain a parametric model of the data we fit the empirical relationship with a polynomial. The fitting procedure outlined will provide an accurate description of the encoding relationship when: 1) \mathbf{x}_t has an elliptically symmetric distribution [36], and 2) the neural response has a one-dimensional dependence on \mathbf{x}_t . Both conditions are satisfied to a good approximation in our data [37]. Determination of an LN system model for unit 1 is illustrated in Fig. 5. The length l of the linear kernel $\mathbf{h}^i(k)$ is not known a priori, nor is the optimal polynomial order. Both parameters affect the model complexity, and an optimal choice has to provide a good fit to the data without being overly complex. Penalizing the data loglikelihood L using Schwarz’s Bayesian Information Criterion (BIC) [38] provides a systematic approach to model selection

$$L_p = \underbrace{\sum_t \sum_{i=1}^n \log(\Pr(N_t^i | \lambda_t^i))}_{L_{\max}} - \underbrace{\frac{\# \text{params}}{2} \cdot \log(\# \text{data points})}_{\text{BIC}} \quad (4)$$

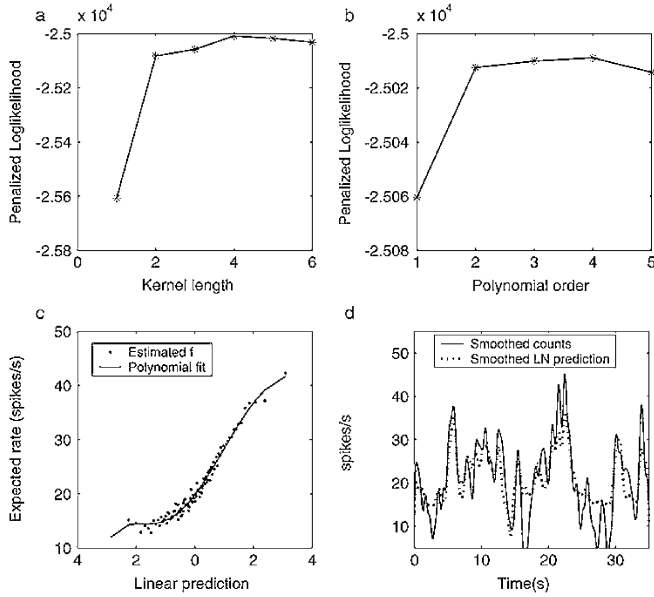


Fig. 5. Results of LN model fitting with neural data. Results are for same unit as in Figs. 1–3. (a) Penalized loglikelihood determination of kernel length. (b) Penalized loglikelihood determination of polynomial order. (c) Static nonlinearity f , and 4th order polynomial fit. (d) Performance of LN model in predicting the firing rate (concatenation of several trials). The spike counts and prediction were smoothed with a hamming window of width 1 s.

where we calculate $\Pr(N_t^i | \lambda_t^i)$ using the statistical models introduced later. An optimal choice is one where the penalized likelihood L_p is maximal: $l = 4$ in Fig. 5(a) (for each of the x and y dimensions). This implies that the encoding relationship for this unit is predominantly a function of position, velocity, acceleration, and jerk. The inferred nonlinearity for this unit [Fig. 5(b)] is a polynomial of order 4 (note that the penalized loglikelihood is nearly flat in the range 2–4). The model is able to describe expected rate modulations in the range 12–45 Hz [Fig. 5(c)]. The collection of units in this study had modulations in the range 0–60 Hz (equivalent to a maximal expectation of 3 spikes/bin). Directly scrutinizing the nonlinearity reveals that it can be mostly described by a piecewise-linear combination of two parts. The power of the inferred LN model in predicting the time-varying firing rate of one such cell can be qualitatively demonstrated by comparing low-pass filtered versions of the prediction to a smoothed version of the binned spike counts [Fig. 5(d)]. The prediction is clearly able to track some of the significant excursions of the firing rate, but generally does not span its full dynamical range (since not all the factors that modulate firing are captured by our measurements and model). The correlation coefficient between the two time series was 0.64; 40% of the units recorded in this study had a correlation coefficient larger than 0.5 (maximal cc: 0.78, mean: 0.4).

Determination of the optimal polynomial orders allows us to break the neural units into three categories: nonlinear (polynomial order > 1), linear (order = 1) and unresponsive (order = 0). Representative units in these categories are presented in Fig. 6. The unresponsive units generally had low average firing rates (0–15 Hz). In the two monkeys from which we recorded (one dataset analyzed from each monkey), unresponsive units accounted for 17% and 7%, respectively

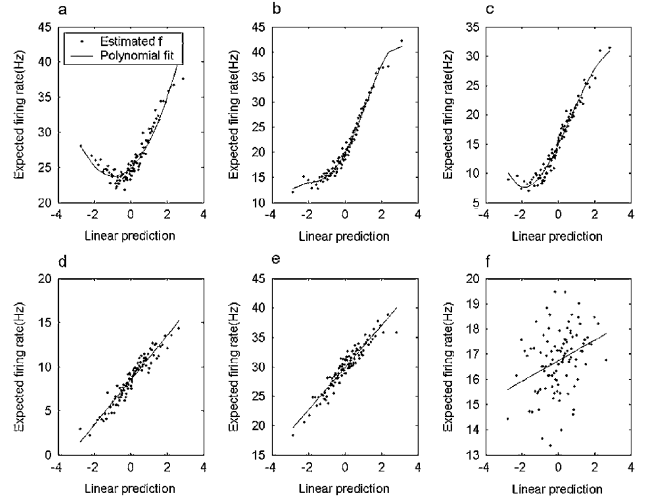


Fig. 6. Calculated static nonlinearity in six representative units. The Expected rate has units of spikes/s. A 3rd order polynomial fit line is also shown for each unit. (a)–(c) Nonlinear functions. (d), (e) Linear encoding functions. (f) Unresponsive unit—no significant encoding of kinematic variables.

(4/23 and 2/27), linear units accounted for 52% and 59%, respectively (12/23 and 16/27), with nonlinear units accounting for the rest—31% and 34%, respectively.

The optimal linear kernel length for the different units in our data fell in the range 1–6, implying that the encoding we observe is typically a mixture of various kinematic variables. Therefore, the encoding relationship must be viewed as more general than the cosine tuning curve [1], or the model proposed more recently by Moran and Schwartz [29] in which the firing rate is seen to be a function exclusively of the direction of motion and its speed. Instead, it is more consistent with the complex dependence on multiple movement parameters revealed by multiple regression analysis [32], [39]. To further illustrate this point, we note that a subset of the cells we analyzed actually coded for position more strongly than for velocity (e.g., unit 2 in Fig. 3) and that for such units the predictive power of our new model was much stronger than that of the velocity tuned models (Fig. 7). Correctly accounting for these units is of particular importance when attempting to decode arm position.

V. NEURAL NOISE PLOTS

Modeling the “expected firing rate” $\lambda_t^i = f^i(\mathbf{x}_t)$ does not provide a complete description of the neural firing process and describing the variability of the neuron firing is also necessary in our framework for likelihood calculations. A favored approach to this problem is to assume an inhomogeneous Poisson model with rate λ_t^i [6], [31]. This provides a relatively simple statistical model, however, in some neural systems the Poisson model is clearly inappropriate (e.g., [40]), and a previous study has indicated a significantly smaller-than-Poisson variability in MI unit activity [12]. While some of the units we recorded here had Poisson firing statistics, a significant proportion did not. The variance versus mean behavior of two representative units is illustrated in Fig. 8(a), and it can be seen that while unit 1 is reasonably well described as Poisson, unit 2 is clearly not.

One principled way of characterizing non-Poisson statistics in spike trains is the method of time rescaling [7], [8], where

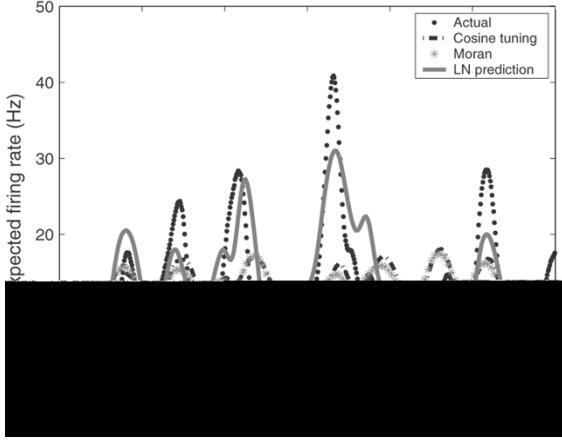


Fig. 7. Failure of velocity based models to predict firing rate in a strongly position encoding unit. The plot compares the low-pass filtered firing of a neural unit with the filtered predictions of three models (several trials concatenated). The cosine tuning model predicts the firing rate solely as a function of the direction of motion, while the Moran model uses velocity and speed as input. Parameters for both models were obtained using a maximum-likelihood procedure. The LN model kernel had a length of 2 (in both x and y) and, therefore, represents a function of position and velocity. A hamming window of 1.5 s was used to filter the time traces. The correlation coefficients when comparing the real and predicted traces are 0.32, 0.28, and 0.79, respectively.

the expected firing rate λ_t^i is used to rescale the time axis, and subsequently can lead to a non-Poisson, history dependent statistical model of the firing. However, as this work is focused on using spike counts in intermediate-size bins (50 ms), an alternative method was developed for describing the statistics of the discrete random spike-counts as a function of the expected firing rate. The method [see Fig. 8(b)] is conceptually similar to the adaptive binning method used above to calculate the non-linearity in the LN cascade model. Here, the expected firing rate λ_t^i time-series (3) is calculated, and used to segment the set of time bins into subsets where λ_t^i falls into different bins $\lambda_j - \Delta\lambda_j/2 < \lambda_t^i < \lambda_j + \Delta\lambda_j/2$. For each subset we find the probabilities for the different values of the number of spikes N_t^i and plot $\Pr(0)$, $\Pr(1)$, $\Pr(2)$, etc. versus the expected value λ_j . The neural noise plot in Fig. 8(c) illustrates that the Poisson model adequately describes the empirical count distributions for unit 1

$$\Pr(N | \lambda) = \frac{e^{-\lambda} \lambda^N}{N!}. \quad (5)$$

In contrast, unit 2 does not fit the Poisson model [Fig. 8(d)]. The discrepancy primarily manifests as “clustering”: higher than expected probability for 1 and 2 spikes, while the probability of 0 and 3 spikes is lower than expected. This clustering leads to the smaller-than-Poisson conditional variance in Fig. 8(a).

VI. THE NORMALIZED-GAUSSIAN DISCRETE DISTRIBUTION

The Poisson model fails to describe the empirical spiking statistics in a large proportion of the units we recorded. Several studies (e.g., [7]) replaced the Poisson model using nonexponential models for the inter-spike interval such as the gamma distribution. However, processes with such ISI distributions lead to nonanalytical forms for the distributions of discrete spike

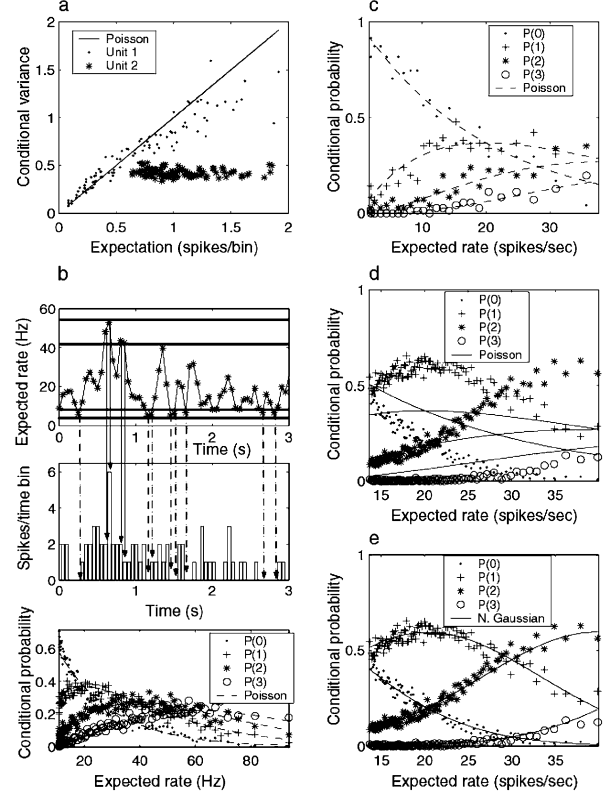


Fig. 8. Statistical properties and noise analysis of motor unit data. (a) Spike count variance versus mean in two representative units. Each point corresponds to one percentile of the time bins, broken according to the prediction of the LN model. (b) Constructing a neural noise plot (simulation data). (top panel) The LN-expected firing rate time series is binned—two bins are illustrated. (center panel) The spike counts time series is segmented accordingly. The proportion of bins having specific number of spikes (0, 1, 2, ...) is evaluated. (bottom panel) The neural noise plot illustrates the conditional probabilities of the different discrete outcomes. Each of the bins contributed one point to each of the discrete probability curves. The lines correspond to a Poisson distribution. (c) Neural noise distribution for Unit 1, compared with a Poisson model. (d) Comparison with Poisson model for Unit 2. (e) Comparison with the Normalized Gaussian model for Unit 2.

counts, and we were unable to find a major discrete distribution [41] that was able to capture the statistics of our empirical distributions. Motivated by the observed decoupling of the variance and the mean [e.g., Fig. 8(a)] as well as the shape of the empirical distributions $\Pr(N | \lambda)$ we observed in the neural noise plots [Fig. 8(c)], we set to adapt the normal distribution

$$\Pr(N | \lambda) = c \exp\left(-\frac{(N - \lambda)^2}{2\sigma^2}\right). \quad (6)$$

to the case where N takes only discrete positive values. One way to do this is to treat the parameter c in (6) as a free parameter, chosen to normalize the probabilities. However, the resulting distribution will no longer have a mean λ , which will limit its usefulness. Instead, we look for a form that satisfies simultaneously the following two constraints:

$$\text{Total Probability: } \sum_{N=0,1,2,\dots} \Pr(N | \lambda) = 1 \quad (7)$$

$$\text{Expectation: } \sum_{N=0,1,2,\dots} N \cdot \Pr(N | \lambda) = \lambda. \quad (8)$$

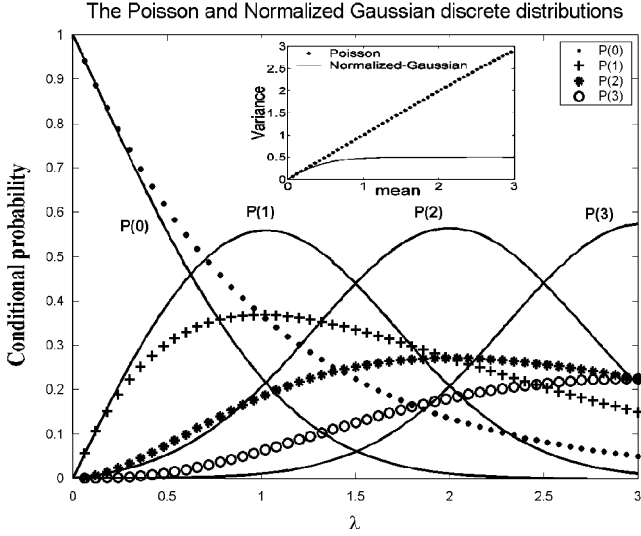


Fig. 9. The Poisson and Normalized Gaussian discrete distributions. The Normalized Gaussian distribution shown here (solid lines) has a parameter $\sigma = 0.71$. The Normalized Gaussian distribution behaves like a Gaussian with standard deviation σ at large values of λ . For small values of σ it approaches the Poisson distribution (symbols) near $\lambda = 0$ (binomial range). The inset compares the behavior of the variance as a function of the mean for the two distributions.

Both constraints cannot in general be satisfied using a single free parameter. A different normalization that is able to simultaneously satisfy these constraints (using two auxiliary variables: $a(\lambda)$ and $b(\lambda)$) has the following form:

$$\begin{cases} P(0|\lambda) = a(\lambda) \exp(-\lambda^2/2\sigma^2) \\ P(1|\lambda) = b(\lambda) \exp(-(\lambda-1)^2/2\sigma^2) \\ P(2|\lambda) = b(\lambda) \exp(-(\lambda-2)^2/2\sigma^2) \\ \vdots \end{cases} \quad (9)$$

The normalizing auxiliary variables: $a(\lambda)$ (zero probability) and $b(\lambda)$ (the rest), can be numerically evaluated from the two constraint [(7) and (8)] for a given value of N and the parameters λ and σ . The simplest way to evaluate these variables is first to evaluate $b(\lambda)$ from (8), and then substitute it in (7) to obtain $a(\lambda)$. We call the resulting distribution the Normalized-Gaussian discrete distribution. Fig. 9 illustrates the Normalized-Gaussian probability of different values of N over a range of different λ 's, and compares it with the Poisson distribution in the same range. We note that the Normalized-Gaussian distribution behaves like a Gaussian distribution at large values of λ and (for small values of σ) approaches a Poisson distribution near $\lambda = 0$. The plot insert demonstrates that while in a Poisson-distributed random variable the variance is equal to the mean, a Normalized-Gaussian variable has the property of saturating variance, and for large means the variance is essentially independent of the mean. Adjusting the value of the dispersion parameter σ results in a different width of the individual bell curves (not shown). In our application σ is fit numerically by minimizing the sum-squared distortion between the observed and calculated distribution in the neural noise plots.

The satisfactory fit that the Normalized-Gaussian distribution provides to our data is illustrated in Fig. 8(e), which shows neural noise plots for a unit for which the Poisson model fails. In fact, we found that the new noise model works well for units with all three types of behaviors we noted in the previous section

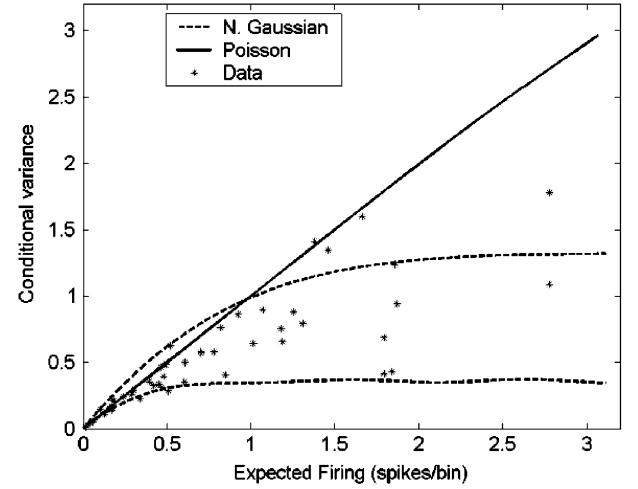


Fig. 10. Variance versus expected firing for all recorded units. Each point represents the variance versus the highest expected firing rate (in spikes/bin) for each unit. The N. Gaussian lines (dashed) are for $\sigma = 0.6$ (lower line) and $\sigma = 1.15$. The solid line is for the Poisson distribution (variance = mean).

(and Fig. 6). Fig. 10 presents a composite view of the variance versus mean behavior of all the 50 units we studied here. About one half of the units lie significantly below the identity line, and were well fit by N. Gaussian models with $0.6 \leq \sigma \leq 1.15$ (the N. Gaussian lines in Fig. 10 show the bounds of this parameter range). The N. Gaussian distribution converges at low σ and low firing rate to the binomial distribution and, therefore, also captures the behavior of the large proportion of units that had low firing rates. In only 8 of the 44 responsive units (all in monkey 2) the Poisson model had a higher penalized likelihood (4) relative to the N. Gaussian model.

Note: our goal in using the statistical models above is to be able to calculate the likelihood of particular experimental outcomes. A few difficulties arise infrequently in this application: the inherent model error in the statistical model may lead to over-weighing outcomes with a very low likelihood, the expected firing rate (calculated from the LN model) may become negative and the constraints in (7) and (8) lead to small negative values of $a(\lambda)$ for certain values of λ and σ . To address all of these issues we truncate the likelihood from below at a small value ϵ ($\epsilon = 0.02$ was chosen here).

VII. MODEL-BASED DECODING

Taken together, (3) and (9) define the likelihood of obtaining a certain number of spikes per bin in neuron i . To simplify further calculations, we assume conditional independence between the simultaneous output of n neurons (see, e.g., [12] and [42]–[44] for different perspectives on the issue of dependencies in motor cortical coding). The independence assumption leads to the following likelihood of the combined firing

$$\begin{aligned} \Pr(\{N_{t-\Delta t, t}^i\} | \{\mathbf{x}_t, \mathbf{x}_{t+1} \dots \mathbf{x}_{t+\tau}\}) \\ = \prod_{i=1}^n NG \left(N_{t-\Delta t, t}^i | f^i \left(\sum_{k=1}^l \mathbf{h}^i(k) \mathbf{x}_{t+k} \right), \sigma^i \right). \end{aligned} \quad (10)$$

A full statistical characterization of our experimental system also requires a model for the dynamics of the arm. As in similar applications in the applied estimation literature [45], [46]

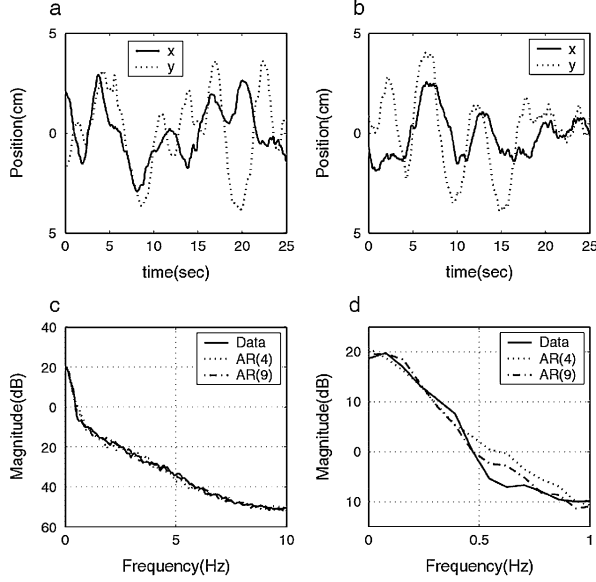


Fig. 11. Performance of auto regressive models of a monkey's hand movement. (a) Experimentally obtained traces of the monkey's hand movement (x and y coordinates). (b) Time series generated by 4th order autoregressive process fit to data from (a). (c) power spectra for data in (a) and (b), as well as best AR model (9th order). (d) Blow up of (c) for low frequencies. (a) Hand position trace, (b) simulated AR(4) trace, (c) power spectra, (d) power spectra (zoom).

we choose the multivariate autoregressive (AR) model, a simple model which describes a large range of realistic motion traces and leads to tractable estimation procedures

$$\mathbf{x}_t = \sum_{k=1}^P \mathbf{A}_k \mathbf{x}_{t-k} + \mathbf{G} \mathbf{w}_t \quad (11)$$

where \mathbf{w} is a vector of independent, unit variance random disturbances (Gaussian distributed), $\{\mathbf{A}_{1..P}\}$ and \mathbf{G} are matrices (estimated using the ARFIT package for fitting multivariate AR models [47]), and \mathbf{x}_t contains the x and y position at time t . The full time trace was obtained by pasting detrended movement traces (to minimize discontinuities). The results of fitting an auto regressive model to one of our data traces are illustrated in Fig. 11 (Note the good approximation provided by an order 4 model ($P = 4$)).

With the encoding and movement models characterized it is now possible to define an algorithm for decoding the arm's movement from the neural responses in a nearly optimal fashion which is tailored to the encoding statistics. Recently, algorithms for recursive Bayesian estimation based on sequential Monte Carlo methods [16] have gained popularity for applications involving nonlinear non-Gaussian observation models (i.e., encoding models), as is our case. The implementation proceeds by recursive application of two computational procedures at every time step, leading from a conditional probability distribution at time $t - \Delta t$, through an intermediate distribution, to a new conditional probability distribution at time t

$$\begin{aligned} P(\mathbf{x}_{t-\Delta t} | \mathbf{N}_{0,t-\Delta t}) &\xrightarrow{\text{PREDICTION}} P(\mathbf{x}_t | \mathbf{N}_{0,t-\Delta t}) \\ &\xrightarrow{\text{UPDATE}} P(\mathbf{x}_t | \mathbf{N}_{0,t}). \end{aligned} \quad (12)$$

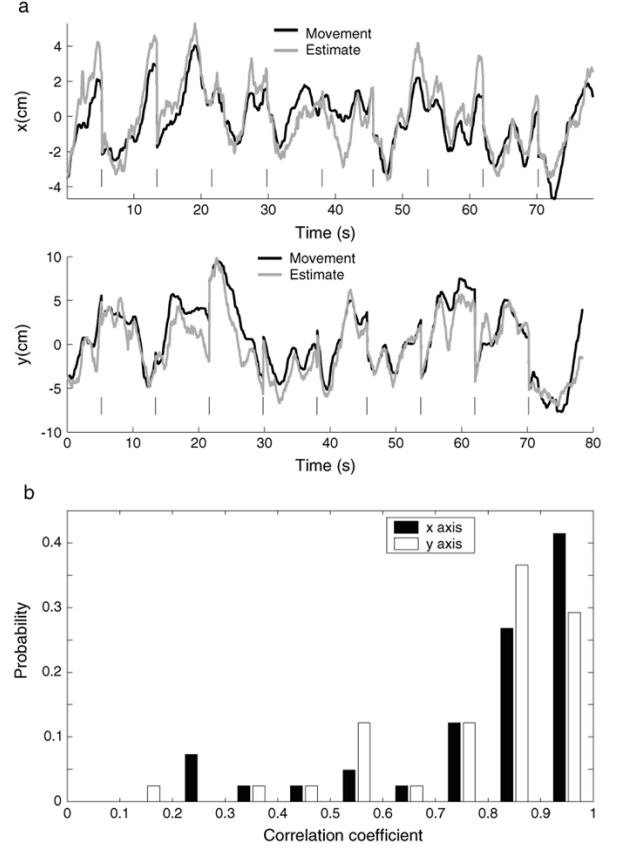


Fig. 12. Performance of a Sequential Monte Carlo filter using 17 simultaneously recorded motor cortical units. The units were selected from the 27 isolated in monkey 2. (a) Estimation performance during 10 successful trials in the x (upper panel) and y (lower panel) dimensions. Trials were stitched together at the locations indicated by the vertical lines. (b) Histogram of correlation coefficients for all 41 trials recorded.

As in all Monte Carlo methods, the conditional probability distributions are approximately represented by a set of representative random samples (or particles), and the computations involve simple operations such as particle movement [using the movement model (11)] and resampling [using the likelihood model (10)]. A detailed account of particle filter implementations is available elsewhere [16]–[18]. Each sample in our implementation is a vector with eight elements (eight-dimensional state space), corresponding to the positions x and y in 4 different time delays. Thus, the filter essentially uses the firing up to time t to predict the hand position at time $t + 3\Delta t$. We implemented the filter using 3000 particles, and our results for position decoding using the 17 best units (selected from the 27 recorded in monkey 2) are illustrated in Fig. 12. In over 2/3 of the trials the model-based trajectory estimate had a correlation coefficient larger than 0.8 when compared to the actual position trajectory. The correlation coefficients mean value was ~ 0.8 . In contrast, trajectories computed using a vector-based optimal linear estimator (OLE) [10] (with a single time bin set at zero delay, as in [18]) had an average correlation coefficient of 0.25 with the actual trajectory (not shown).

VIII. DISCUSSION

Our development of statistical encoding models for the activity of MI neurons has introduced a few useful new tools and pro-

vided some insight into properties of motor encoding. We began by characterizing the encoding of kinematics by MI neuron firing rates. The majority of the neurons we have studied here had firing rates that were linearly related to the kinematic variables. This linearity is not surprising, as it is consistent with the classical cosine shaped tuning curve (with respect to the direction of movement), and was thus anticipated by theoretical studies [5], [48]. However, a possibility that cannot be discounted based on our data is that the linearity results from an underlying nonlinear relationship that is artificially linearized by being experimentally sampled in a narrow range of positions, velocities, and accelerations. Since we are probably underestimating the full-range nonlinearity, it is interesting to note that over one third of the responsive units we have studied exhibited significant nonlinearity, and that most of the highly informative neurons were nonlinear. Nonlinear motor encoding may reflect pervasive nonlinearities throughout the motor control system (which models typically linearize [3]), and a specific multiplicative form of motor encoding nonlinearity has recently been suggested based on psychophysical data [49]. Experimental methods that enable tracking the arm position over a wide range of motions are now available (e.g., the Shape Tape, Measurand Inc., Fredrickton, NB, Canada), and this study illustrates the importance of performing such experiments in order to obtain a more complete picture of motor encoding processes.

This analysis also shows that motor cortical units encode complex functions of the arm position and its derivatives, which does not lend support to the view that motor cortex neurons encode only simplified global features of movement (like the direction of movement). To emphasize this point, we have provided a direct illustration of the extremely weak predictive power of velocity-based encoding models in a representative unit, and contrasted it with the more complete characterization provided using the LN cascade models (Fig. 7). We note that the cascade approach can also be extended to the study of neuronal interactions [44].

We found that the variability around the expected firing rate is markedly sub-Poisson in many of the movement-related units we recorded, and those units may exhibit a nearly constant variance across a wide dynamic range. In contrast with other studies that have looked at variability of motor responses (e.g., [12]), we have attempted to use our encoding model to account for some of the variability that is a result of the movement-related modulations in firing rate. The remaining variability, which we capture using the Normalized-Gaussian noise model, is due in part to using an imperfect model, in part to the imperfect observation of motion (only movement in 2-D was measured in our experiments), and in part to the inherent statistics of neuronal firing. Our use of a predictive model in conjunction with the neural noise plot introduced here allowed us to faithfully represent signal-dependent distributions that are necessary for model-based decoding; however, we are unable to distinguish between the different sources of variance with the existing dataset. This difficulty notwithstanding, it is extremely likely that after ‘explaining’ additional variance the intrinsic neural statistical properties will have very little variance—markedly sub-Poisson. The nonpoisson nature of the variability has a number of implications. Square-root transforming of binned counts (e.g., [29] and [50]) is a variance stabilizing transformation suitable for Poisson processes where the noise variance

depends on the signal, and is ill-suited for the analysis of data with the properties we have described. Interestingly, the signal-dependent noise observed in the peripheral motor system, which may explain a number of phenomenological scaling laws observed in motor control [51] does not depend on signal dependent noise in MI, but rather on the organization and properties of the peripheral motor-unit pool [52].

In order to address the observed statistical properties of neural firing we developed the Normalized Gaussian distribution, a new discrete probability distribution. We anticipate that the new distribution will find many additional applications in studies of neural systems, particularly cortical systems. While we have used it in conjunction with a predictive model, in other types of experiments the peri-stimulus time histogram (PSTH) of responses to highly reproducible experimental conditions could provide an approximate firing rate model. This probability distribution is easy to compute, provides a signal independent variance at high spike counts and smoothly transitions to a binomial-like behavior at low spike counts. It is also flexible enough to fit different data with a range of non-Poisson variability characteristics. The statistics of responses in extended time-windows in the visual cortex [53] are well fit by a truncated Gaussian, similar to the statistics observed in the primary motor cortex [12], strengthening our belief in the general applicability of the Normalized-Gaussian distribution.

Finally, we have demonstrated that the encoding model presented can be integrated with recursive likelihood-based estimation procedures to yield practical filters for a brain-machine interface. The filter we implemented yielded a very large performance gain versus the optimal linear vector-based method (which out-performs the population vector [10]). This result was partially due to the recursive filter’s ability to use the temporal smoothness of the motion trajectory while the OLE is an instantaneous estimator. Other filtering methods available for this application include finite-impulse response linear filters [27], [54]–[56], recursive Kalman filtering [see [57]] and neural network based filters [56] (note that [27] and [57] tested their results using the same experimental data that we used here). Based on fundamental results in estimation theory, the optimal filter for a linear encoding model, with an additive Gaussian noise process is a linear one. This appears not very far from the situation we have described here for motor cortical units. Nevertheless, we expect that the model-based approach used here will offer future advantages in this brain-machine interface application. First, the models used are compact and the small number of parameters makes for good generalization. Second, they provide clear measures for which units are “good encoders,” providing for improved generalization performance. Third, our filter is recursive and so can deal well with short decoding epochs. Fourth, it can easily deal with recording nonstationarity (unit adaptation and loss). And finally, it is quite likely that most neurons used in actual brain-machine interfaces will exhibit rich nonlinearities in their activity patterns, which will also be well-captured by our approach.

ACKNOWLEDGMENT

The authors would like to thank Y. Gao, D. J. Warren, Prof. S. Geman, Prof. E. Bienenstock, Prof. M. Black, Prof. E. Brown,

and Prof. S. Nagarajan as well as two anonymous referees for valuable input during the preparation of this manuscript.

REFERENCES

- [1] A. P. Georgopoulos, J. F. Kalaska, R. Caminiti, and J. T. Massey, "On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex," *J. Neurosci.*, vol. 2, pp. 1527–1537, 1982.
- [2] D. R. Humphrey, "Relating motor cortex spike trains to measures of motor performance," *Brain Res.*, vol. 40, pp. 7–18, 1972.
- [3] E. Todorov, "Direct cortical control of muscle activation in voluntary arm movements: A model," *Nat. Neurosci.*, vol. 3, pp. 391–398, 2000.
- [4] F. A. Mussa-Ivaldi, "Do neurons in the motor cortex encode movement direction? An alternative hypothesis," *Neurosci. Lett.*, vol. 91, pp. 106–111, 1988.
- [5] T. D. Sanger, "Theoretical considerations for the analysis of population coding in motor cortex," *Neural Comput.*, vol. 6, pp. 29–37, 1994.
- [6] E. N. Brown, L. M. Frank, D. Tang, M. C. Quirk, and M. A. Wilson, "A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells," *J. Neurosci.*, vol. 18, pp. 7411–7425, 1998.
- [7] R. Barbieri, M. C. Quirk, L. M. Frank, M. A. Wilson, and E. N. Brown, "Construction and analysis of nonpoisson stimulus-response models of neural spiking activity," *J. Neurosci. Methods*, vol. 105, pp. 25–37, 2001.
- [8] E. N. Brown, R. Barbieri, V. Ventura, R. Kass, and L. M. Frank, "The time-rescaling theorem and its application to neural spike train data analysis," *Neural Comput.*, vol. 14, pp. 325–346, 2001.
- [9] A. P. Georgopoulos, A. B. Schwartz, and R. E. Kettner, "Neuronal population coding of movement direction," *Science*, vol. 233, pp. 1416–1419, 1986.
- [10] E. Salinas and L. F. Abbott, "Vector reconstruction from firing rates," *J. Comput. Neurosci.*, vol. 1, pp. 89–107, 1994.
- [11] T. D. Sanger, "Probability density estimation for the interpretation of neural population codes," *J. Neurophysiol.*, vol. 76, pp. 2790–2793, 1996.
- [12] E. M. Maynard, N. G. Hatsopoulos, C. L. Ojakangas, B. D. Acuna, J. N. Sanes, R. A. Normann, and J. P. Donoghue, "Neuronal interactions improve cortical population coding of movement direction," *J. Neurosci.*, vol. 19, pp. 8083–8093, 1999.
- [13] A. V. Lukashin, B. R. Amirikian, and A. P. Georgopoulos, "A simulated actuator driven by motor cortical signals," *Neuroreport*, vol. 7, pp. 2597–2601, 1996.
- [14] J. P. Donoghue, "Connecting cortex to machines: Recent advances in brain interfaces," *Nat. Neurosci.*, vol. 5, pp. 1085–1088, 2002. Suppl.
- [15] M. A. L. Nicolelis, "Brain-machine interfaces to restore motor function and probe neural circuits," *Nat. Rev. Neurosci.*, vol. 4, pp. 417–422, 2003.
- [16] A. Doucet, N. D. Freitas, and N. Gordon, *Sequential Monte Carlo Methods in Practice*. Berlin, Germany: Springer-Verlag, 2001.
- [17] S. Shoham, "Advances toward an implantable motor cortical interface," Ph.D. Dissertation, Dept. Bioeng, Univ. Utah, Salt Lake City, 2001.
- [18] A. E. Brockwell, A. L. Rojas, and R. E. Kass, "Recursive Bayesian decoding of motor cortical signals by particle filtering," *J. Neurophysiol.*, vol. 91, pp. 1899–1907, 2004.
- [19] K. E. Jones, P. K. Campbell, and R. A. Normann, "A glass/silicon composite intracortical electrode array," *Ann. Biomed. Eng.*, vol. 20, pp. 423–437, 1992.
- [20] E. M. Maynard, E. Fernandez, and R. A. Normann, "A technique to prevent dural adhesions to chronically implanted microelectrode arrays," *J. Neurosci. Methods*, vol. 97, pp. 93–101, 2000.
- [21] J. P. Donoghue, J. N. Sanes, N. G. Hatsopoulos, and G. Gaal, "Neural discharge and local field potential oscillations in primate motor cortex during voluntary movements," *J. Neurophysiol.*, vol. 79, pp. 159–173, 1998.
- [22] E. M. Maynard, C. T. Nordhausen, and R. A. Normann, "The Utah intracortical electrode array: A recording structure for potential brain-computer interfaces," *Electroencephalogr. Clin. Neurophysiol.*, vol. 102, pp. 228–239, 1997.
- [23] P. J. Rousche and R. A. Normann, "Chronic recording capability of the Utah intracortical electrode array in cat sensory cortex," *J. Neurosci. Methods*, vol. 82, pp. 1–15, 1998.
- [24] 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.
- [25] K. S. Guillory and R. A. Normann, "A 100-channel system for real time detection and storage of extracellular spike waveforms," *J. Neurosci. Methods*, vol. 91, pp. 21–29, 1999.
- [26] S. Shoham, M. R. Fellows, and R. A. Normann, "Robust, automatic spike sorting using mixtures of multivariate t-distributions," *J. Neurosci. Methods*, vol. 127, pp. 111–122, 2003.
- [27] L. Paninski, M. R. Fellows, N. G. Hatsopoulos, and J. P. Donoghue, "Spatiotemporal tuning of motor cortical neurons for hand position and velocity," *J. Neurophysiol.*, vol. 91, pp. 515–532, 2004.
- [28] D. L. Snyder and M. I. Miller, *Random Point Processes in Time and Space*, 2nd ed. New York: Springer, 1991.
- [29] D. W. Moran and A. B. Schwartz, "Motor cortical representation of speed and direction during reaching," *J. Neurophysiol.*, vol. 82, pp. 2676–2692, 1999.
- [30] B. Amirikian and A. P. Georgopoulos, "Directional tuning profiles of motor cortical cells," *Neurosci. Res.*, vol. 36, pp. 73–79, 2000.
- [31] K. Zhang, I. Ginzburg, B. L. McNaughton, and T. J. Sejnowski, "Interpreting neuronal population activity by reconstruction: Unified framework with application to hippocampal place cells," *J. Neurophysiol.*, vol. 79, pp. 1017–1044, 1998.
- [32] J. Ashe and A. P. Georgopoulos, "Movement parameters and neural activity in motor cortex and area 5," *Cereb. Cortex*, vol. 4, pp. 590–600, 1994.
- [33] D. W. Moran and A. B. Schwartz, "Motor cortical activity during drawing movements: Population representation during spiral tracing," *J. Neurophysiol.*, vol. 82, pp. 2693–2704, 1999.
- [34] I. W. Hunter and M. J. Korenberg, "The identification of nonlinear biological systems: Wiener and Hammerstein cascade models," *Biol. Cybern.*, vol. 55, pp. 135–144, 1986.
- [35] M. J. Korenberg and I. W. Hunter, "Two methods for identifying Wiener cascades having noninvertible static nonlinearities," *Ann. Biomed. Eng.*, vol. 27, pp. 793–804, 1999.
- [36] E. J. Chichilnisky, "A simple white noise analysis of neuronal light responses," *Network*, vol. 12, pp. 199–213, 2001.
- [37] L. Paninski, "Convergence properties of three spike-triggered analysis techniques," *Network*, vol. 14, pp. 437–464, 2003.
- [38] G. Schwarz, "Estimating the dimensions of a model," *Ann. Statist.*, vol. 6, pp. 461–464, 1978.
- [39] Q. G. Fu, D. Flament, J. D. Coltz, and T. J. Ebner, "Temporal encoding of movement kinematics in the discharge of primate primary motor and premotor neurons," *J. Neurophysiol.*, vol. 73, pp. 836–854, 1995.
- [40] M. J. Berry, D. K. Warland, and M. Meister, "The structure and precision of retinal spike trains," *Proc. Nat. Acad. Sci.*, vol. 94, pp. 5411–5416, 1997.
- [41] N. L. Johnson and S. Kotz, *Discrete Distributions*. Boston, MA: Houghton Mifflin, 1969.
- [42] A. Riehle, S. Grun, M. Diesmann, and A. Aertsen, "Spike synchronization and rate modulation differentially involved in motor cortical function," *Science*, vol. 278, pp. 1950–1953, 1997.
- [43] N. G. Hatsopoulos, L. Paninski, and J. P. Donoghue, "Sequential movement representations based on correlated neuronal activity," *Exp. Brain Res.*, vol. 149, pp. 478–486, 2003.
- [44] L. Paninski, S. Shoham, M. R. Fellows, N. G. Hatsopoulos, and J. P. Donoghue, "Superlinear population encoding of dynamic hand trajectory in primary motor cortex," *J. Neurosci.*, vol. 24, pp. 8551–8561, 2004.
- [45] A. Gelb, *Applied Optimal Estimation*. Cambridge, MA: MIT Press, 1974.
- [46] P. S. Maybeck, *Stochastic Models, Estimation, and Control*. New York: Academic, 1979, vol. 1–3.
- [47] T. Schneider and A. Neumaier, "Algorithm 808: ARfit—A Matlab package for the estimation of parameters and eigenmodes of multivariate autoregressive models," in *ACM Trans. Math. Softw.*, vol. 27, 2001, pp. 58–65.
- [48] K. Zhang and T. J. Sejnowski, "A theory of geometric constraints on neural activity for natural three-dimensional movement," *J. Neurosci.*, vol. 19, pp. 3122–3145, 1999.
- [49] E. J. Hwang, O. Donchin, M. A. Smith, and R. Shadmehr, "A gain-field encoding of limb position and velocity in the internal model of arm dynamics," *PLoS Biol.*, vol. 1, p. E25, 2003.
- [50] A. B. Schwartz, "Motor cortical activity during drawing movements: Population representation during sinusoid tracing," *J. Neurophysiol.*, vol. 70, pp. 28–36, 1993.
- [51] C. M. Harris and D. M. Wolpert, "Signal-dependent noise determines motor planning," *Nature*, vol. 394, pp. 780–784, 1998.

- [52] K. E. Jones, A. F. De, C. Hamilton, and D. M. Wolpert, "Sources of signal-dependent noise during isometric force production," *J. Neurophysiol.*, vol. 88, pp. 1533–1544, 2002.
 - [53] E. D. Gershon, M. C. Wiener, P. E. Latham, and B. J. Richmond, "Coding strategies in monkey V1 and inferior temporal cortices," *J. Neurophysiol.*, vol. 79, pp. 1135–1144, 1998.
 - [54] D. R. Humphrey, E. M. Schmidt, and W. D. Thompson, "Predicting measures of motor performance from multiple cortical spike trains," *Science*, vol. 170, pp. 758–762, 1970.
 - [55] A. B. Schwartz, "Direct cortical representation of drawing," *Science*, vol. 265, pp. 540–542, 1994.
 - [56] J. Wessberg, C. R. Stambaugh, J. D. Krallk, P. D. Beck, M. Laubach, J. K. Chapin, J. Kim, S. J. Biggs, M. A. Srinivasan, and M. A. L. Nicolelis, "Real-time prediction of hand trajectory by ensembles of cortical neurons in primates," *Nature*, vol. 408, pp. 361–365, 2000.
 - [57] W. Wu, M. J. Black, D. Mumford, Y. Gao, E. Bienenstock, and J. P. Donoghue, "Modeling and decoding motor cortical activity using a switching Kalman filter," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 933–942, Jun. 2004.
- Shy Shoham** (S'97–M'01), photograph and biography not available at the time of publication.
- Liam M. Paninski**, photograph and biography not available at the time of publication.
- Matthew R. Fellows**, photograph and biography not available at the time of publication.
- Nicholas G. Hatsopoulos**, photograph and biography not available at the time of publication.
- John P. Donoghue**, photograph and biography not available at the time of publication.
- Richard A. Normann**, photograph and biography not available at the time of publication.