22 Dynamic Analyses of Neural Representations Using the State-space Modeling Paradigm

Emery N. Brown and Riccardo Barbieri

Neuroscience Statistics Research Laboratory, Department of Anesthesia and Critical Care, Massachusetts General Hospital, Division of Health Sciences and Technology, Harvard Medical School/Massachusetts Institute of Technology, Boston, Massachusetts 02114-2696

ABSTRACT

Understanding how neurons represent biological signals in their ensemble spiking activity is a fundamental challenge in neuroscience. A key property of the neural representations is that they are plastic. That is, with repeated exposure to the same stimulus, a neuron changes its firing properties. Understanding these fundamental properties of neurons is essential for understanding how the nervous system responds to any stimulus including, light, sound, temperature, contact, a cognitive task, or pharmacological substances. This suggests that methods used to analyze neurophysiological data should be designed specifically to characterize ensemble spiking activity and neural plasticity. In this chapter, we describe how a broad class of neural spike train decoding algorithms for studying ensemble representations and adaptive filter algorithms for analyzing neural plasticity can be derived in a unified way using the state-space modeling paradigm. We show the decoding algorithm in the analysis of position decoding from the ensemble spiking activity of rat hippocampal neurons. In addition, we illustrate the adaptive filter algorithm by studying the temporal dynamics of the spatial receptive field of a hippocampal neuron. In the final section, we discuss directions for future research.

INTRODUCTION

Understanding how neurons represent a biological signal in their joint spiking activity is a fundamental challenge in neuroscience (Georgopoulos et al. 1986; Bialek et al. 1991; Wilson and McNaughton 1993; Stanley et al. 1999; Brown et al. 2004; Wu et al. 2004). During the past 12 years, the development of multiunit electrode arrays that allow the simultaneous recording of a large number of neurons is an important experimental advance that now makes it possible to study these ensemble representations (Wilson and McNaughton 1993). In addition, neural representations are plastic. That is, with experience, neurons in many brain regions change their spiking responses to relevant stimuli or biological signals (Kaas et al. 1983; Merzenich et al. 1984; Jog et al. 1999; Gandolfo et al. 2000; Mehta et al. 2000). Deciphering how groups of neurons represent information in their ensemble spiking activity and the dynamics of neural plasticity are key elements to understanding how the nervous system responds to any stimulus—whether it be light,

sound, temperature, contact, a cognitive task, or a pharmacological substance. These characterizations are crucial for defining both the normal physiological states of neural systems as well as pathological conditions that occur in diseases such as drug addiction. These observations suggest that methods used to analyze neurophysiological data should be designed specifically to characterize ensemble spiking activity and the plasticity of the individual neurons (Brown et al. 2004).

Signal Processing Algorithms for Neural Spike Train Data Analysis

Neural spike train decoding algorithms are commonly used with signal processing methods for analyzing how neural systems represent biological signals (Brown et al. 1998; Barbieri et al. 2004; Brockwell et al. 2004; Wu et al. 2004). These algorithms use a wide range of methods to estimate from neural spiking activity the most likely value of a biological signal. More recently, these algorithms have been one of several strategies used to design controls for neural prosthetic devices and other brain-machine interfaces (Chapin et al. 1999; Wessberg et al. 2000; Donoghue 2002; Serruya et al. 2002; Taylor et al. 2002; Mussallam et al. 2004). Developing optimal strategies for constructing and testing decoding algorithms is therefore an important challenge in computational neuroscience. Similarly, adaptive filtering signal processing algorithms offer an approach to analyzing the dynamics of neural receptive fields (Brown et al. 2001; Frank et al. 2002, 2004; Wirth et al. 2003; Eden et al. 2004). Both the neural spike train decoding algorithms and the adaptive filtering algorithms can be constructed using the state-space modeling paradigm. State-space modeling is an established framework used to study dynamical processes in engineering, computer science, and statistics (Kitagawa and Gersh 1996; Smith and Brown 2003).

GOALS OF THE CHAPTER

In this chapter, we derive a class of broadly applicable neural spike train decoding algorithms and adaptive filter algorithms using the state-space modeling paradigm. This approach provides a unified framework for analyzing both neural ensemble representations of biological signals and neural plasticity. First, we review the relevant neurophysiology of the rat hippocampus. Second, we review the state-space paradigm and derive the neural spike train decoding algorithms. Third, we illustrate the decoding algorithm in the analysis of position decoding from ensemble neural spiking activity of simultaneously recorded rat hippocampal neurons. Fourth, we show how the adaptive filter algorithm can be derived as a special case of the decoding algorithm, and we illustrate its use with an analysis of the temporal dynamics of hippocampal place receptive fields. In the final section, we present a summary and discuss gaps and opportunities.

DYNAMIC PROPERTIES OF RAT HIPPOCAMPAL NEURONS

The Hippocampus: A Model for Neural Information Processing

The rat hippocampus is an experimental system widely used to study how neural systems represent information. Located in the temporal lobes, the hippocampus is a bilateral

region critical for the formation and storage of both short- and long-term memories (Scoville and Milner 1957; Squire 1982; Cohen and Eichenbaum 1993; Rudy and Sutherland 1995). It is now appreciated that, within the rat hippocampus, specialized pyramidal neurons known as place cells tune their receptive field properties to the animal's current environment and thereby allow the animal to develop a spatial representation (O'Keefe and Dostrovsky 1971). That is, as an animal moves through its environment, each one of a subset of the hippocampal place cells demarcates, within a few minutes, its own region in the environment by firing spikes only when the animal is within that region. The region of the environment in which the cell fires is termed its place field. Large numbers of hippocampal place cells tile each environment with overlapping place fields. The recorded shapes of these fields depend on the structure of the environment and the task the animal is required to perform. When the animal forages randomly in an open environment, these fields resemble approximately two-dimensional Gaussian surfaces (Fig. 1), whereas when the animal runs on a linear track, the fields resemble a one-dimensional Gaussian curve. The hippocampal theta rhythm with its frequency range of 6–14 Hz, the animal's running velocity, and its direction of motion are additional factors known to modulate place cell spiking activity (O'Keefe and Reece 1993; Wilson and McNaughton 1993). The spatial information carried by hippocampus place cells is believed to be an important component of the rat's spatial navigation system (McNaughton et al. 1996).

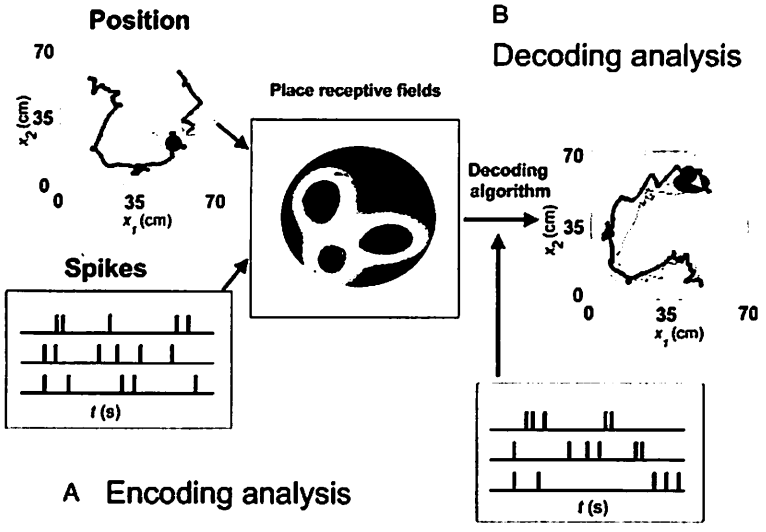

FIGURE 1. Decoding of position from rat ensemble neural spiking activity. (*A*) Encoding analysis in which the relation between the biological signal (trajectory of the rat in the environment, *solid black line in the position panel*) and spiking activity (*Spikes panel*) is estimated as place receptive fields for three neurons. (*B*) Decoding analysis in which the estimated place receptive fields are used in the point process filter decoding algorithm to compute the predicted position (*thin black line*) of the rat in the environment from new spiking activity of the neural ensemble recorded during the decoding stage. The predicted position is compared with the observed position (*thick black line*) during the decoding stage. The blue oval defines a 95% confidence region centered at that location. (Reprinted, with permission, from Brown et al. 2004.)

Multiunit Electrode Arrays for Recording Ensemble Neural Activity

During the past 12 years, the development of multiunit electrode arrays that allow the simultaneous recording of a large number of individual neurons represents an important experimental advance in neuroscience (Wilson and McNaughton 1993; Buzsaki 2004). Once implanted, a multiunit electrode array can be used in some cases to study population neural activity for up to 1 year. The technical capability to simultaneously record the spiking activity of large numbers of hippocampal place cells—between 20 and 100—along with the animal's position in its environment has made feasible formal quantitative study of how rats represent spatial information in short-term memory. We show how the state-space paradigm may be used to analyze the representation of the animal's position in its environment maintained by the ensemble spiking activity of the hippocampal place cell neurons recorded from multiunit electrode arrays.

Dynamics of Hippocampal Place Receptive Fields

Once formed, the place receptive fields of hippocampal neurons are not static. That is, rat hippocampal neurons have been shown to change their spatial receptive fields as the animal executes either a novel or familiar behavioral task (Mehta et al. 2000; Frank et al. 2002, 2004). For example, when the experimental environment is a linear track, these spatial receptive fields have been shown to migrate in the direction opposite to the cell's preferred direction of firing relative to the animal's movement, and increase in scale and maximum firing rate (Mehta et al. 2000). This behavior can be seen in the spiking activity of the place cell neuron recorded on the U-shaped track shown in Figure 2. To display all of the experimental data on a single graph in space and time, we show a linear representation of the U-shaped track. At the outset of the experiment, the spiking activity is sparse, centered near 90 cm, and ranges between 75 and 100 cm, whereas by the end of the experiment, it is appreciably more intense, centered near 70 cm, and ranges between 50 and 90 cm. Because receptive field plasticity is a characteristic of many neural systems, analysis of these dynamics from experimental measurements is crucial for understanding how different brain regions learn and adapt their representations of relevant biological information. Therefore, we show below how the state-space paradigm may be used to track the dynamics of hippocampal place receptive field properties on a millisecond timescale.

THE STATE-SPACE MODELING PARADIGM

The State Equation

We assume that neural representations can be studied with the state-space framework (Kitagawa and Gersh 1996; Smith and Brown 2003). The state-space model consists of two equations: a state equation and an observation equation. The state equation defines an unobservable process whose evolution is tracked across time. Such state models with unobservable processes are often referred to as hidden Markov or latent process models

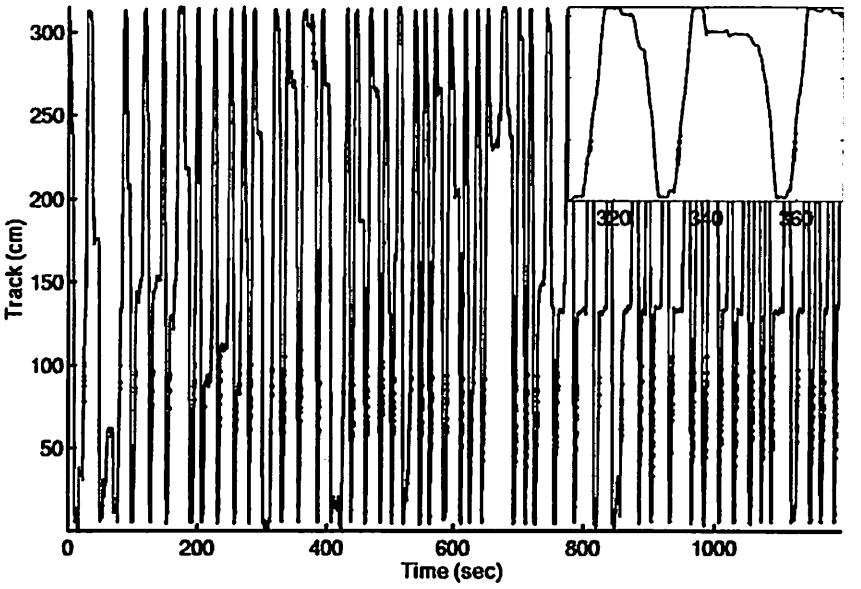

FIGURE 2. Place-specific firing dynamics of an actual CA1 place cell recorded from a rat running back and forth on a 300-cm U-shaped track for 1200 sec. The track was linearized to display the entire experiment in a single graph. The vertical lines show the animal's position and the red dots indicate the time at which a spike was recorded. The inset is an enlargement of the display from 320 to 360 sec to show the cell's unidirectional firing, i.e., spiking only when the animal runs from the bottom to the top of the track. (Reprinted, with permission, from Brown et al. 2001 [©National Academy of Sciences, U.S.A.].)

(Fahrmeir and Tutz 2001; Smith and Brown 2003). In the analyses we present, the state process represents the biological signal or the parameters of the neural receptive field. For an ensemble of rat hippocampal neurons, the state process would be the position of the animal in its environment (Barbieri et al. 2004). In the case of a hippocampal neuron whose receptive field is changing with time, the state equation defines the dynamics of the parameters in the model describing the receptive field (Brown et al. 2001; Eden et al. 2004). For other examples, for an ensemble of cat lateral geniculate neurons, the state equation could be the natural image being projected onto the animal's visual field (Stanley et al. 1999), whereas for monkey M1 neurons, it could be the velocity profile of an intended hand movement (Wu et al. 2004).

The Observation Equation

The observation equation completes the state-space model paradigm and defines how the observed data relate to the unobservable state process. In all the cases that we consider, the observed data from the neurophysiological experiments are neural spike trains. For these analyses, we consider the neural spike trains as point processes, time series that take on the values of 0 or 1 in continuous time (Brown et al. 2003; Daley and Vere-Jones

2003; Brown 2005). For a neural spike train modeled as a point process, the relationship between the spiking activity and the state process is defined by the conditional intensity function (Brown et al. 2003; Daley and Vere-Jones 2003; Brown 2005). The conditional intensity function is the rate function of each neuron. Like the rate function of an inhomogeneous Poisson process, the conditional intensity function can depend on a time-varying process that in our analyses will be the state process. For any small time interval, the conditional intensity function defines for each neuron the probability of a spike occurring in that small time interval. The conditional intensity function generalizes the concept of a Poisson rate function in that it can also be dependent on the history of the neurons' spiking activity and the history of the state process.

We describe the state-space model in terms of the decoding problem for hippocampal place cell neurons. In the first part of the analysis, termed the encoding analysis, the receptive fields of the neurons are characterized (Fig. 1A). In the second part, the decoding algorithm is used to analyze how the representation of the animal's position in its environment is carried in the ensemble spiking activity of the pyramidal neurons (Fig. 1B).

FORMULATING THE NEURAL SPIKE TRAIN DECODING ALGORITHM

We denote the decoding interval as $(0,T]$. We define the updating lattice by choosing K large and dividing $(0,T]$ into K intervals of equal width $\Delta = K^{-1}T$. In the decoding analysis, we choose $\Delta = 3.3$ msec, whereas in the neural receptive field analysis, we choose $\Delta = 1$ msec.

State and Observation Equations

To define the state-space model, we assume that the position of the rat during foraging obeys a bivariate first-order linear Gaussian autoregressive *AR(1)* model defined as

$$x_k = \mu_x + Fx_{k-1} + \epsilon_k \tag{1}$$

where $x_k = (x_1(k\Delta), x_2(k\Delta))$ is the animal's position at time $k\Delta$, F is a 2 x 2 matrix of system parameters, μ_x is a 2 x 1 vector of mean parameters, and ϵ_k is a 2 x 1 Gaussian random variable with mean zero and a 2 x 2 covariance matrix W_ϵ.

The observation model for the ensemble neural spike train is the discretized version of the point process joint probability density defined as (Brown 2005)

$$p(n_k \mid x_k) = \exp\left(\sum_{c=1}^{C}[n_k^c \log \lambda_k^c - \lambda_k^c \Delta]\right) \tag{2}$$

where n_k^c is the number of spikes in the interval $((k-1)\Delta, k\Delta]$, from neurons $c = 1,\ldots,C$, λ_k^c is the conditional intensity function of neuron c at time $k\Delta$, and $n_k = (n_k^1,\ldots,n_k^C)$. For each problem, we give the explicit form of the conditional intensity function. We let $N_k = [n_1, \ldots, n_k]$ be the sequences of spike train observations in the interval $(0, k\Delta]$ for $k = 1,\ldots,K$. Equation (2) assumes that the spiking activities of the individual neurons are independent given the value of the position at time $k\Delta$. In the subsequent discussions, we denote time $k\Delta$ as time k.

The Bayes' Rule and Chapman-Kolmogorov Equations

The objective of our state-space modeling analysis is to construct a recursive filter algorithm to estimate the state x_k at time k from the spiking activity N_k. The standard approach to deriving such a filter is to recursively define the probability density of the state given the observations. For the state model defined in Eq. (1) and the observation process defined in Eq. (2), the Bayes' rule formula for the posterior probability density of the state-given spike train observations is (Brown et al. 1998; Barbieri et al. 2004; Eden et al. 2004)

$$p(x_k \mid N_k) = \frac{p(x_k \mid N_{k-1})p(n_k \mid x_k)}{p(n_k \mid N_{k-1})} \tag{3}$$

and the associated one-step prediction probability density or Chapman-Kolmogorov equation is

$$p(x_k \mid N_{k-1}) = \int p(x_{k-1} \mid N_{k-1})p(x_k \mid x_{k-1})dx_{k-1} \tag{4}$$

Before presenting the explicit form of the recursive filter algorithm, we explain the logic behind Eqs. (3) and (4) (Fig. 3). The first term on the right side of Eq. (3), $p(x_k \mid N_{k-1})$, is the one-step prediction probability density from Eq. (4). Before recording the spike trains at time k, it defines the predictions of the most likely values of the state at time k given the observations up through time $k-1$. Equation (4) predicts the most likely values of the state at time k by integrating ("averaging") over the most likely values of the state at time $k-1$ given the data up through $k-1$ defined by $p(x_{k-1} \mid N_{k-1})$ and the most likely set of changes in state between times $k-1$ to k defined by $p(x_k \mid x_{k-1})$. The most likely value of the state at time $k-1$ given the data up through $k-1$, $p(x_{k-1} \mid N(x_{k-1})$ is the first term of the integrand in Eq. (4) and this value defines the posterior probability density at $k-1$. The most likely change in state from time $k-1$ to k, $p(x_k \mid x_{k-1})$, is defined

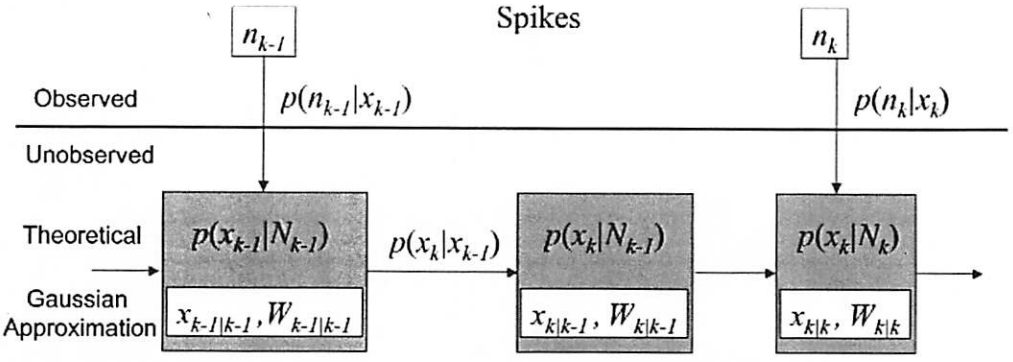

FIGURE 3. Schematic representation of the state-space paradigm showing the relationship between the observation process $p(n_k \mid x_k)$ (Eq. [2]), the state process $p(x_k \mid x_{k-1})$ defined by Eq. (1), the one-step prediction $p(x_k \mid N_{k-1})$ (Eq. [4]), its Gaussian approximation from the point process filter algorithm with mean $x_{k|k-1}$ (Eq. [5]) and variance $W_{k|k-1}$ (Eq. [6]), and the posterior probability density $p(x_k \mid N_k)$ (Eq. [3]), and its Gaussian approximation with mean $x_{k|k}$ (Eq. [7]) and variance $W_{k|k}$ (Eq. [8]) from the point process filter algorithm.

by the autoregressive probability model for the state in Eq. (1). The second term on the right-hand side of Eq. (3), $p(n_k \mid x_k)$, is the point process probability mass function at time k given the state x_k defined in Eq. (2). The probability density $p(n_k \mid N_{k-1})$ in the denominator of Eq. (3) is the integral of the numerator and defines the normalizing constant, which ensures that the posterior probability density integrates to 1.

Together, Eqs. (3) and (4) define a recursion that can be used to compute the probability of the state given the observations (Fig. 3). The formulae in Eqs. (3) and (4) are recursive because Eq. (4) uses the posterior probability density at time $k-1$, $p(x_{k-1} \mid N_{k-1})$, to generate the one-step prediction probability density at time k, $p(x_k \mid N_{k-1})$, which in turn, makes it possible to compute the new posterior probability density at time k, $p(x_k \mid N_k)$, given in Eq. (3). In other words, at each time k, the recursion relations fuse information from two sources (Fig. 3). The first is the prediction of the state at time k given the model, N_{k-1}, and the observations up through time $k-1$, as given by Eq. (4), and the second is the spike train's observations recorded in the interval $((k-1)\Delta, k\Delta]$ as given by Eq. (2).

Evaluating the Bayes' Rule and Chapman-Kolmogorov Equations

Given the state model in Eq. (1) and the observation process in Eq. (2), the challenge of estimating the unobserved states from the observed neural spike trains is simply the computational problem of evaluating Eqs. (3) and (4). For systems with low-dimensional state and observation models, Eqs. (3) and (4) can be evaluated numerically (Kitagawa and Gersh 1996). As these dimensions increase, numerical computation becomes less feasible. A standard approach, and the one we apply here, is to compute Gaussian approximations to Eqs. (3) and (4) (Brown et al. 1998; Barbieri et al. 2004; Eden et al. 2004). This approach, also termed *maximum a posteriori* estimation, amounts to finding the maximum and the second derivative of Eq. (3) as a function of the state x_k. The maximum corresponds to the mode or mean of the posterior density (Eq. [3]) and the curvature defines its covariance matrix. These two quantities are sufficient to compute the Gaussian approximation to Eq. (3) because a Gaussian probability density is defined completely by its mean and covariance matrix (Fig. 3).

Assuming that a Gaussian approximation to Eq. (3) was computed at time $k-1$, then the integral in Eq. (4) can be computed analytically using standard formulae for computing the mean and covariance of the sum of two Gaussian random variables (Barbieri et al. 2004). Under the Gaussian approximation, the recursion defined by the probability densities in Eqs. (3) and (4) becomes a recursive filter because it simplifies to computing recursively just the means and covariance of these probability densities. In the special case that the state process is a linear Gaussian system and the observation model is a linear Gaussian function of the state process, this recursive computation of the means and covariances is the well-known Kalman filter algorithm (Kitagawa and Gersh 1996).

Neural Spike Train Decoding Algorithm

Under the assumption that the neurons are conditionally independent, the solutions to these equations using Gaussian approximations to Eqs. (3) and (4) for the state model in

Eq. (1) and the point process observation model in Eq. (2) yield the following recursive filter algorithm (Barbieri et al. 2004):

$$\textit{One-step prediction: } x_{k|k-1} = \mu_x + Fx_{k-1|k-1} \tag{5}$$

$$\textit{One-step prediction variance: } W_{k|k-1} = FW_{k-1|k-1}F' + W_\epsilon \tag{6}$$

$$\textit{Posterior mode: } x_{k|k} = x_{k|k-1} + W_{k|k-1} \sum_{c=1}^{C} \nabla \log \lambda_k^c [n_k^c - \lambda_k^c \Delta] \tag{7}$$

$$\textit{Posterior variance: } W_{k|k-1}^{-1} = \left[W_{k|k-1}^{-1} - \sum_{c=1}^{C} \left[\nabla^2 \log \lambda_k^c [n_k^c - \lambda_k^c \Delta] - \nabla \log \lambda_k^c [\nabla \lambda_k^c \Delta]^2 \right] \right] \tag{8}$$

for $k = 1,\ldots,K$, where λ_k^c is the conditional intensity function for neuron c; F and W_ϵ are, respectively, the transition matrix and the white noise covariance matrix for the $AR(1)$ model in Eq. (1); ∇ (∇^2) denotes the first (second) derivative of the indicated function with respect to x_k; and the notation $j|k$ denotes the estimate at time j given the spiking activity in the interval $(0, k\Delta]$.

We termed the algorithm in Eqs. (5)–(8) a point process filter because it takes point process observations as inputs at each time k, which, in this case, are the neural spike train observations n_k, and it provides a recursive system of Gaussian approximations ~~with which~~ to compute Eqs. (3) and (4) (Fig. 3). Equations (7) and (8) give, respectively, the mean and covariance matrix for the Gaussian approximation to $p(x_k|N_k)$ (Eq. [3]), whereas Eqs. (5) and (6) give, respectively, the mean and covariance matrix for the Gaussian approximation to $p(x_k|N_{k-1})$ (Eq. [4]). As mentioned above, because we compute Gaussian approximations to $p(x_k|N_k)$ and $p(x_k|N_{k-1})$, it suffices to compute recursively only the respective means and covariance matrices (Fig. 3).

Properties of the Algorithm

The position update $x_{k|k}$ combines the one-step prediction $x_{k|k-1}$, on the basis of the ensemble spiking activity from $(0,(k-1)\Delta]$, with the weighted sum of $[n_k^c - \lambda_k^c\Delta]$, the innovation or error signal from neuron c for $c = 1,\ldots,C$ multiplied by the one-step prediction variance (Barbieri et al. 2004). The innovations are the new information from the ensemble spiking activity in $((k-1)\Delta, k\Delta]$. As is true for a Poisson process, $\lambda_k^c\Delta$ defines for a general point process model the probability of a spike in $((k-1)\Delta, k\Delta]$ (Brown et al. 2003; Daley and Vere-Jones 2003; Brown 2005). Each innovation compares the probability of a spike, $\lambda_k^c\Delta$, in $((k-1)\Delta, k\Delta]$ with n_k^c, which, for small Δ, is 1 if a spike is observed from neuron c in $((k-1)\Delta, k\Delta]$ and 0 otherwise. Thus, for small Δ, the innovation gives a weight in the interval $(-1,1)$. A large positive weight results if a spike occurs when a neuron has a low probability of spiking in $((k-1)\Delta, k\Delta]$, whereas a large negative weight arises if no spike occurs in $((k-1)\Delta, k\Delta]$ when the probability of a spike is high. In this way, the algorithm makes use of the times when a neuron either does or does not fire. The algorithm in Eqs. (7) and (8) is nonlinear and must be solved using Newton's method because the right- and left-hand sides of Eq. (7) depend on $x_k|_k$.

DYNAMIC ANALYSIS OF POSITION REPRESENTATION IN ENSEMBLE HIPPOCAMPAL SPIKING ACTIVITY

Experimental Protocol

We applied the recursive filter paradigm to place cell spike train and position data recorded from a Long-Evans rat freely foraging in a circular environment 70 cm in diameter with 30-cm-high walls and a visual cue maintained in a fixed location. A multiunit electrode array was implanted into the CA1 region of the animal's hippocampus. The simultaneous activity of 34 place cells was recorded from the electrode array while the animal foraged in the environment for 25 minutes. Simultaneously with the recording of the place cell activity, the position of the animal was measured at 30 Hz by a camera tracking the location of two infrared diodes mounted on the animal's headstage (Brown et al. 1998; Barbieri et al. 2004).

Encoding Analysis

We modeled the spatial receptive field of each neuron by representing the conditional intensity function as a linear combination of Zernike polynomials defined as

$$\lambda_k^c = \exp\left\{\sum_{\ell=0}^{L}\sum_{m=-\ell}^{\ell}\theta_{\ell,m}^c z_\ell^m[\rho[k\Delta],\phi[k\Delta]]\right\} \tag{9}$$

where $z_{\ell,m}$ is the mth component of the ℓth-order Zernike polynomial, $\theta_{\ell,m}^c$ is the associated coefficient, $\rho(k\Delta) = r^{-1}[(x_1(k\Delta) - \eta_1)^2 + (x_2(k\Delta) - \eta_2)^2]^{1/2}$, $\phi(k\Delta) = \tan^{-1}[(x_2(k\Delta) - \eta_2)(x_1(k\Delta) - \eta_1)^{-1}]$, (η_1, η_2) are the coordinates of the center of the circular environment, r is the radius of the circular environment, $\eta_1 = \eta_2 = r = 35$ cm, and $L = 3$. The Zernike polynomials are orthogonal polynomials on the disk and provide a mathematically efficient way to represent the range of shapes taken on by the spatial receptive fields. The order-3 model has ten nonzero coefficients. We estimated the spatial receptive field of each neuron by fitting the Zernike model to the spike train data of each neuron from the first 15 minutes of the experiment. That is, we used maximum likelihood to estimate the parameters $\theta_{\ell,m}^c$ for each neuron. The Zernike place field estimates, shown in Fig. 4, covered nearly the entire environment.

Decoding Analysis

We estimated the F and W_ϵ matrices (Eq. [1]) by maximum likelihood for the first 15 minutes of the path data from the encoding analysis. We decoded position from the ensemble spiking activity of the 34 neurons from the last 10 minutes of foraging with updates computed at $\Delta = 3.3$ msec. A representative example of the algorithm's performance is shown in Figure 5. In each of the four continuous 15-second segments, the estimated trajectories (Eq. [7]) (*red*) closely resembled the true (*black*) trajectories. The instantaneous 95% confidence ellipses were computed from Eq. (7) and the covariance matrix in Eq.

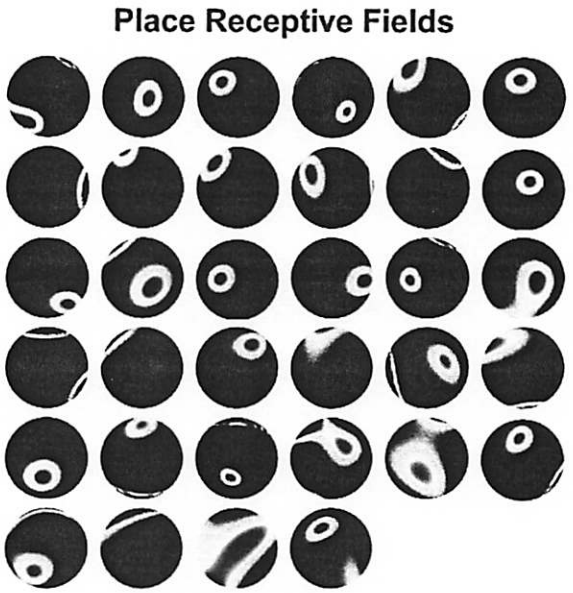

FIGURE 4. Pseudocolor maps of the spatial receptive field estimates from the Zernike models for the 34 place cells. (Adapted, with permission, from Barbieri et al. 2004 [©Massachusetts Institute of Technology].)

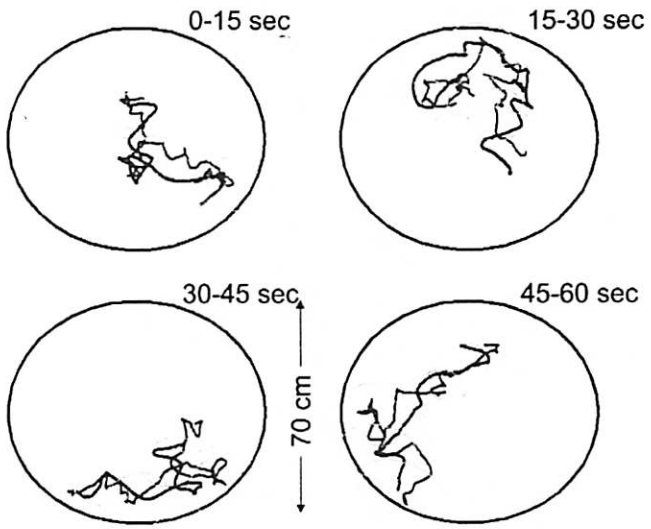

FIGURE 5. Continuous 60-sec segment of the true path (*black*) displayed in four 15-sec intervals with the estimated path (*red*) from the Zernike model.

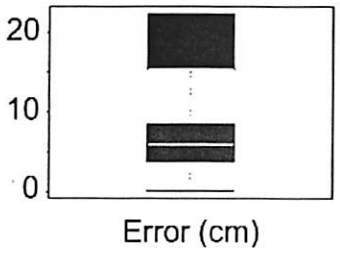

FIGURE 6. Box plot summary of decoding error for the Zernike model distribution (error = true position – estimated position). The lower border of the box is the 25th percentile of the distribution and the upper border is the 75th percentile. The white bar within the box is the median of distribution. The distance between the 25th and 75th percentiles is the interquartile range (IQR). The lower (upper) whisker is at 1.5 x the IQR below (above) the 25th (75th) percentile. All of the black bars above the whiskers are far outliers. (Adapted, with permission, from Barbieri et al. 2004 [©Massachusetts Institute of Technology].)

(8). The performance of the algorithm in decoding the spike train data can be best appreciated in the videos of this analysis on our Web site http://neurostat.mgh.harvard.edu/.

To quantify further the performance of the decoding algorithm, we computed the decoding error as the difference between the true and estimated path at each 33-msec time step for the 10 minutes of decoding. The box plot summaries of this error distribution are shown in Figure 6. The median decoding error for the Zernike model was 5.9 cm, with a 25th percentile of 3.9 cm, a 75th percentile of 8.6 cm, a minimum of 0.8 cm, and a maximum of 22.3 cm. The Zernike model was highly accurate near the borders of the environment (Fig. 6) because, by Eq. (5), the support of the Zernike model is restricted to the circle. The median error of 5.9 cm is an improvement of 2 cm from the 7.9 cm using spatial Gaussian functions to model the place receptive fields (Brown et al. 1998). Of this improvement, 1.2 cm was attributable to optimizing the learning rate, whereas 0.7 cm was because of the switch from the spatial Gaussian to the Zernike model. These results show that the CA1 region of the hippocampus maintains a dynamic representation of the animal's position in its environment. They also illustrate how the state-space modeling paradigm may be used to analyze how well a rat's internal representation of its position in its environment may be recovered from ensemble spiking activity of approximately 30 hippocampal neurons.

POINT PROCESS FILTER ALGORITHM FOR TRACKING HIPPOCAMPAL PLACE RECEPTIVE FIELD DYNAMICS

Steepest-descent Point Process Filter Algorithm

We can use the state-space paradigm to devise an algorithm to track the dynamics of a neuron's receptive field. The general formulation of this state-space algorithm for tracking neural plasticity is given in Eden et al. (2004). Here, we illustrate the idea with the special case of the steepest-descent point process filter (Brown et al. 2001). To derive the steepest-descent algorithm from the point process filter algorithm in Eqs. (5)–(8), we take the $C = 1$ neuron and set $F = I$, $W_{\epsilon} = 0$, and $W_{k|k-1} = D(\epsilon)$, where $D(\epsilon)$ is a diagonal matrix with possibly distinct constant value ε_i on each ith diagonal. In this case, the point process filter algorithm simplifies to a single formula in which Eq. (7) becomes

$$\theta_{k|k} = \theta_{k-1|k-1} + D(\epsilon)\nabla \log \lambda_{k-1}[n_k - \lambda_{k-1}\,\Delta] \tag{10}$$

where θ is the parameter of the conditional intensity function that is evolving in time. In the previous example, θ was the vector of coefficients in the Zernike polynomials. Because in our paradigm the conditional intensity function defines the receptive field of the neuron, and because θ parametrizes the conditional intensity function, tracking the evolution of θ through time allows us to track the temporal evolution of the receptive field. We term Eq. (10) the steepest-descent point process adaptive filter because it can also be derived using a steepest-descent argument applied to an appropriate local criterion function as is true for continuous-valued measurements (Brown et al. 2001). Unlike Eq. (7), in Eq. (10), the conditional intensity function is evaluated at time $k-1$ instead of time k, making Eq. (10) a linear algorithm.

To illustrate the steepest-descent point process filter algorithm, we consider a spike train from a pyramidal cell in the CA1 region of the rat hippocampus recorded while the animal ran back and forth on a U-shaped track (Fig. 2). As stated above, the spatial receptive fields of hippocampal neurons have well-documented dynamic properties when an animal executes either a familiar or novel behavioral task (Mehta et al. 2000). In our analysis, we consider only the dynamic properties of place field position and shape, not other factors such as theta rhythm oscillations, and the interaction between the theta rhythm and position called phase precession (O'Keefe and Reece 1993). If $k(k\Delta)$ is the animal's position at time k, then we define the conditional intensity function for the place field model as

$$\lambda(k\Delta|\theta = \exp\{\alpha - (2\sigma^2)^{-1}\,(x(k\Delta) - \mu)^2\} \tag{11}$$

where μ is the center of the spatial receptive field, σ is a scale factor, and $\exp\{\alpha\}$ is the neuron's maximum firing rate that occurs at the center of the receptive field. For this model, $\theta = (\alpha,\sigma,\mu)'$ is the three-dimensional parameter vector. From Eqs. (10) and (11), the steepest-descent point process filter algorithm at time k is

$$\theta_{k|k} = \theta_{k-1|k-1} - \left[\begin{matrix} \epsilon_\alpha \\ \epsilon_\sigma(\sigma^{-3}_{k-1|k-1})(x(k\Delta) - \mu_{k-1|k-1})^2 \\ \epsilon_\mu\sigma^{-2}_{k-1|k-1}(x(k\Delta) - \mu_{k-1|k-1}) \end{matrix}\right]\left[n_k - \lambda((k-1)\Delta|\theta_{k-1|k-1})\right] \tag{12}$$

where ϵ_α, ϵ_σ, and ϵ_μ are, respectively, the diagonal elements of $D(\epsilon)$ for α, σ, and μ, and $k = 1,\ldots,K$. The elements of $D(\epsilon)$ are the learning rate parameters and control how much the innovations, i.e., the new spiking information, are weighted.

Experimental Protocol

We applied the steepest-descent point process filter algorithm to an actual place cell spike train recorded from a Long-Evans rat running back and forth for 1200 sec on a 300-cm U-shaped track (Fig. 2) (Brown et al. 2001). Multiple single units as well as the position of the animal on the track were recorded as described in the previous example. The actual trajectory was at times irregular because the animal started and stopped several times, and in two instances (50 and 650 sec), turned around shortly after it initiated its run. On several of the upward passes, particularly in the latter part of the experiment, the animal slowed as it approached the curve in the U-shaped track at approximately 150 cm. The strong place-specific firing of the neuron is easily visible because the spiking

activity occurred almost exclusively between 50 and 100 cm. The spiking activity of the neuron was entirely unidirectional because the cell discharges only as the animal ran up and not down the track (Fig. 2, inset).

Tracking Hippocampal Spatial Receptive Field Dynamics

We applied the steepest-descent point process filter algorithm (Eq. [12]) to these spike train data, updating the estimates every 1 msec. We used the learning rate parameters, $D(\epsilon)$ chosen in Brown et al. (2001). The starting estimates for the parameter θ were the maximum likelihood estimates computed from the first 50 spikes (~200 sec). The time course of $\exp(\alpha)$, the maximum spike rate, showed a steady increase from 3 to almost 30 spikes/sec during the 1200 sec of the experiment (Fig. 7A). The increase was apparent in the spike train data in Figure 2. The scale parameter σ showed the greatest fluctuations during the experiment; it rose during the first 500 sec from 10 to 16 cm, and fluctuated between 15 and 16 cm for the balance of the experiment (Fig. 7B). This fluctuation in scale was also easily visible in the spike train data in Fig. 2. The place cell center migrated during the first 700 sec from 85 to 65 cm and stayed near 65 cm for the remainder of the experiment (Fig. 7A) 7C

We illustrate the evolution of the entire field in Figure 8 by plotting the instantaneous place field estimates at 300, 550, 800, and 1150 sec. The sequence of place field estimates showed the temporal evolution of the cell's spatial receptive field. In contrast, the

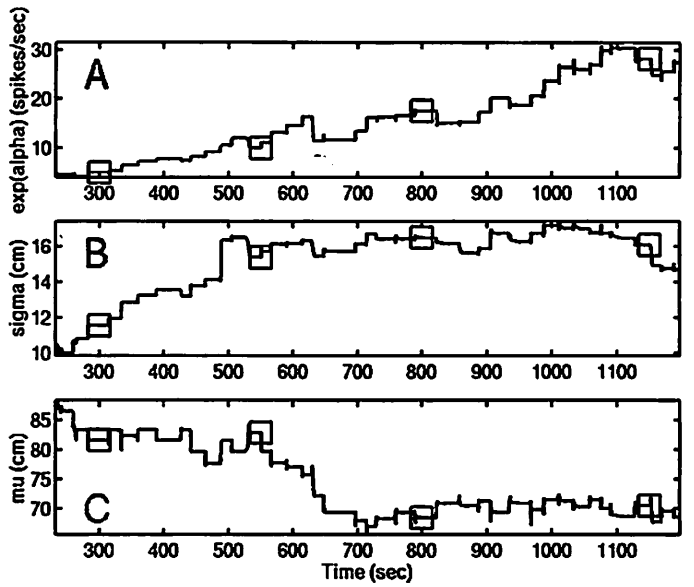

FIGURE 7. Steepest-descent point process filter estimates of the trajectories: (*A*) Maximum spike rate, $\exp(\alpha)$; (*B*) place field scale, σ; and (*C*) place field center, μ. The estimates were updated at 1-msec intervals. The squares in the panels at 300, 550, 800, and 1150 sec are the times at which the place fields are displayed in Fig. 8. The growth of the maximum spike rate (*A*), the variability of the place field scale (*B*), and the migration of the place field center (*C*) are all readily visible. (Reprinted, with permission, from Brown et al. 2001 [©National Academy of Sciences, U.S.A.].)

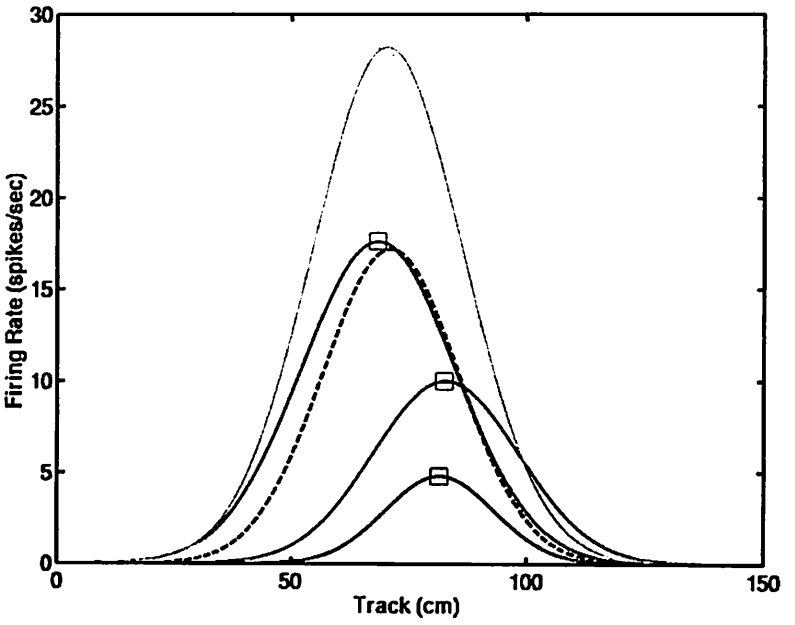

FIGURE 8. Estimated place fields at times 300 (*blue*), 550 (*green*), 800 (*red*), and 1150 (*aqua*) sec from the steepest-descent point process algorithm. The black dashed line is the maximum likelihood estimate of the place field obtained by using all of the spikes in the experiment. The maximum likelihood estimate ignores the temporal evolution of the place field. (Reprinted, with permission, from Brown et al. 2001 [©National Academy of Sciences, U.S.A.].)

maximum likelihood estimate based on the entire 1200 sec of data obscured the temporal dynamics by overestimating (underestimating) the place field's spatial extent, and underestimating (overestimating) its maximum firing rate at the experiment's beginning (end). This example illustrates that the dynamic of actual place cell receptive fields can be tracked instantaneously from recorded spiking activity. The migration of center as well as the growth in scale and in maximum firing rate were consistent with previous reports (Mehta et al. 2000). The current algorithm could not track the skewing of the place fields because the Gaussian model is symmetric. This feature as well as more complex spatial and temporal dynamics have been well captured by the spline extensions of this point process filter algorithm (Frank et al. 2002, 2004). Video presentations of our analyses can be found in the supplemental information at www.pnas.org along with analyses of the spatial receptive field dynamics of three other CA1 hippocampal neurons.

SUMMARY

We have presented a state-space modeling approach for analyzing neural ensemble representations of biological signals and for analyzing neural receptive field plasticity. These algorithms are part of the paradigm that we have been developing to conduct state estimation from point process measurements (Smith and Brown 2003; Barbieri et al. 2004; Eden et al. 2004). We previously compared our state-space paradigm with several other

approaches including the widely used reverse correlation algorithm (Brown et al. 1998; Barbieri et al. 2004). The reverse correlation algorithm, although simple to implement, performed significantly worse than our algorithm, because it made no use of the spatial and temporal structure in this decoding problem. This finding suggests that application of our algorithm to other decoding and neural prosthetic control problems where the reverse correlation methods have performed successfully (Bialek et al. 1991; Stanley et al. 1999; Wessberg et al. 2000; Serruya et al. 2002) could yield important improvements. Furthermore, our state-space approach to estimating receptive field dynamics offers a way of tracking their temporal evolution on a millisecond timescale (Brown et al. 2001; Frank et al. 2002, 2004; Eden et al. 2004).

GAPS AND OPPORTUNITIES

Further technical improvements in our algorithm can be readily incorporated into the current framework. These include combining the Zernike spatial model with models of known hippocampal temporal dynamics such as bursting (Quirk and Wilson 1999), theta phase modulation (Skaggs et al. 1996; Brown et al. 1998), and phase precession (O'Keefe and Reece 1993). Next, we can consider both higher-order linear state-space models (Shoham 2001; Smith and Brown 2003; Wu et al. 2004) and nonlinear state-space models to describe more accurately path dynamics (Kitagawa and Gersh 1996). Finally, we have developed Gaussian approximations to evaluate the Bayes' rule and Chapman-Kolmogorov equations in Eqs. (3) and (4). We can improve our algorithm and evaluate the accuracy of this approximation by applying non-Gaussian approximations (Pawitan 2001) and exact numerical integration methods (Kitagawa and Gersh 1996) and/or Monte Carlo techniques (Doucet et al. 2001; Shoham 2001) to evaluate these equations.

We foresee several future applications of the state-space modeling paradigm. First, we will continue to use this modeling approach to study ensemble neural representations in the hippocampus. More detailed models that take account of bursting, theta phase procession, and the spiking history of the ensemble should allow us to gain more insight into how information in these ensembles of neurons is being represented (Okatan et al. 2005). Second, as mentioned above, the state-space approach provides a natural framework for developing algorithms to control brain-machine interfaces and neural prosthetic devices. The appropriate application of these algorithms to these problems will depend critically on the nature of the encoding model used to describe the relationship between neural spiking activity and the signal being represented, as well as the kinematic model for the movement (Eden et al. 2004; Wu et al. 2004; Truccolo et al. 2005). Third, we plan to further develop our state-space models to study receptive field dynamics (Frank et al. 2002, 2004). In particular, our approach will allow us to study the specific dynamics of hippocampal place fields when they form for the first time as an animal learns about a novel environment. Finally, our state-space framework has led to a new approach for characterizing learning in behavioral experiments and for relating it to simultaneously recorded changes in neural activity (Wirth et al. 2003; Smith et al. 2004). This work has allowed us to identify specific neurons in the monkey hippocampus that change their spiking properties in close relation to the time course of the animal's learning of a new behavioral task (Wirth et al. 2003).

All of these applications have direct implications for the study of drug addiction, to the

extent that drugs alter neural circuits, and these drug-induced changes may affect the representation of information in the neural circuit and/or the normal plasticity of the neurons comprising the circuit. Depending on the nature of the neurophysiological investigation, these drug-induced impairments may be manifested in the activity of individual neurons, groups of neurons, or in the behavior of the organism. For example, the state-space methods we have presented could be used to analyze differences in the responses of neural circuits with and without a pharmacologic agent to help characterize such drug-induced effects. As a second illustration, we recently showed how our state-space paradigm applied to learning experiments could be used to dynamically quantify the differences in learning between a control group of rats and a second group treated with an NMDA receptor antagonist, MK801 (Smith et al. 2005). This latter application illustrates how our paradigm may be used to characterize dynamics on the timescale of an organism's behavior. For these reasons, we believe that the state-space paradigm offers a rich framework for extracting dynamic information from neurophysiological experimental investigations of both normal and pathological conditions.

ACKNOWLEDGMENTS

This work was supported in part by National Institute on Drug Abuse Grant No. DA-015644 and National Institute of Mental Health Grant Nos. MH-59733 and MH-61637. We are grateful to Julie Scott for technical assistance in preparing this chapter.

REFERENCES

Barbieri R., Frank L.M., Nguyen D.P., Quirk M.C., Solo V., Wilson M.A., and Brown E.N. 2004. Dynamic analyses of information encoding by neural ensembles. *Neural Comput.* **16:** 277–308.

Bialek W., Rieke F., de Ruyter van Stevenick R.R., and Warland D. 1991. Reading a neural code. *Science* **252:** 1854–1857.

Brockwell A.E., Rojas A.L., and Kass R.E. 2004. Recursive Bayesian encoding of motor cortical signals by particle filtering. *J. Neurophysiol.* **91:** 1899–1907.

Brown E.N. 2005. Theory of point processes for neural systems. In *Methods and models in neurophysics* (ed. D. Hansel). Elsevier, Paris. (In press.) **AU: Update; need pp. nos.**

Brown E.N., Kass R.E., and Mitra P.P. 2004. Multiple neural spike train data analysis: State-of-the-art and future challenges. *Nat. Neurosci.* **7:** 456–461.

Brown E.N., Barbieri R., Eden U.T., and Frank L.M. 2003. Likelihood methods for neural data analysis. In *Computational neuroscience: A comprehensive approach* (ed. J. Feng), pp. 253–286. CRC, London.

Brown E.N., Frank L.M., Tang D., Quirk M.C., and Wilson M.A. 1998. A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells. *J. Neurosci.* **18:** 7411–7425.

Brown E.N., Nguyen D.P., Frank L.M., Wilson M.A., and Solo V. 2001. An analysis of neural receptive field plasticity by point process adaptive filtering. *Proc. Natl. Acad. Sci.* **98:** 12261–12266.

Chapin J.K., Moxon K.A., Markowitz R.S., and Nicolelis M.A.L. 1999. Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex. *Nat. Neurosci.* **7:** 664–670.

Cohen N.J. and Eichenbaum H. 1993. *Memory, amnesia, and the hippocampal system.* MIT Press, Cambridge.

Daley D. and Vere-Jones D. 2003. *An introduction to the theory of point process,* 2nd edition. Springer-Verlag, New York.

Donoghue J.P. 2002. Connecting cortex to machines: Recent advances in brain interfaces. *Nat. Neurosci.* (suppl.) **5:** 1085–1088.

Doucet A., DeFreitas N., Gordon N., and Smith A. 2001. *Sequential Monte Carlo methods in practice.* Springer, New York.

Eden U.T., Frank L.M., Barbieri R., Solo V., and Brown E.N. 2004. Dynamic analyses of neural encoding by point process adaptive filtering. *Neural Comput.* **16:** 971–998.

Fahrmeir L. and Tutz G. 2001. *Multivariate statistical modeling based on generalized linear models,* 2nd edition. Springer-Verlag, New York.

Frank L.M., Stanley G.B., and Brown E.N. 2004. Hippocampal plasticity across multiple days of exposure to novel environments. *J. Neurosci.* **24:** 7681–7689.

Frank L.M., Eden U.T., Solo V., Wilson M.A., and Brown E.N. 2002. Contrasting patterns of receptive field plasticity in the hippocampus and the entorhinal cortex: An adaptive filtering approach. *J. Neurosci.* **22:** 3817–3830.

Gandolfo F., Li C., Benda B.J., Schioppa C.P., and Bizzi E. 2000. Cortical correlates of learning in monkeys adapting to a new dynamical environment. *Proc. Natl. Acad. Sci.* **97:** 2259–2263.

Georgopoulos A.P., Kettner R.E., and Schwartz A.B. 1986. Neuronal population coding of movement direction. *Science* **233:** 1416–1419.

Jog M.S., Kubota Y., Connolly C.I., Hillegaart V., and Graybiel A.M. 1999. Building neural representations of habits. *Science* **286:** 1745–1749.

Kaas J.H., Merzenich M.M., and Killackey H.P. 1983. The reorganization of somatosensory cortex following peripheral nerve damage in adult and developing mammals. *Annu. Rev. Neurosci.* **6:** 325–356.

Kitagawa G. and Gersh W. 1996. *Smoothness priors analysis of time series.* Springer-Verlag, New York.

Mehta M.R., Quirk M.C., and Wilson M.A. 2000. Experience-dependent asymmetric shape of hippocampal receptive fields. *Neuron* **25:** 707–715.

Merzenich M.M., Nelson R.J., Stryker M.P., Cyander M.S., Schoppmann A., and Zook J.M. 1984. Somatosensory cortical map changes following digit amputation in adult monkeys. *J. Comput. Neurol.* **224:** 591–605.

Mussallam S., Corneil B.D., Greger B., Scherberger H., and Andersen R.A. 2004. Cognitive control signals for neural prosthetics. *Science* **305:** 258–262.

Okatan M., Wilson M.A., and Brown E.N. 2005. Analyzing functional connectivity using a network likelihood model of ensemble neural spiking activity. *Neural Comput.* (in press). **AU: Update vol, pp. nos.**

O'Keefe J. and Dostrovsky J. 1971. The hippocampus as a spatial map: Preliminary evidence from unit activity in the freely-moving rat. *Brain Res.* **34:** 171–175.

O'Keefe J. and Reece M.L. 1993. Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus* **3:** 317–330.

Pawitan Y. 2001. *All likelihood: Statistical modelling and inference using likelihood.* Oxford University Press, London.

Quirk M.C. and Wilson M.A. 1999. Interaction between spike waveform classification and temporal sequence detection. *J. Neurosci. Methods* **94:** 41–52.

Rudy J.W. and Sutherland R.J. 1995. Configural association theory and the hippocampal formation: An appraisal and reconfiguration. *Hippocampus* **5:** 375–389.

Scoville W.B. and Milner B. 1957. Loss of recent memory after bilateral hippocampal lesions. *J. Neurol. Neurosurg. Psychiatry* **20:** 11–21.

Serruya M.D., Hatsopoulos N.G., Paninski L., Fellows M.R., and Donoghue J.P. 2002. Instant neural control of a movement signal. *Nature* **416:** 141–142.

Shoham S. 2001. "Advances towards an implantable motor cortical interface." Ph.D. thesis, University of Utah, Salt Lake City.

Skaggs W.E., McNaughton B.L., Wilson M.A., and Barnes C.A. 1996. Theta phase precession in hippocampal neuronal populations and the compression of temporal sequences. *Hippocampus* **6:** 149–172.

Smith A.C. and Brown E.N. 2003. Estimating a state-space model from point process observations. *Neural Comput.* **15:** 965–991.

Smith A.C., Stefani M.R., Moghaddam B., and Brown E.N. 2005. Analysis and design of behavioral experiments to characterize population learning. *J. Neurophysiol.* **93:** 1776–1792.

Smith A.C., Frank L.M., Wirth S., Yanike M., Hu D., Kubota Y., Graybiel A.M., Suzuki W.A., and Brown E.N. 2004. Dynamic analysis of learning in behavioral experiments. *J. Neurosci.* **24:** 447–461.

Squire L.R. 1982. The neuropsychology of human memory. *Annu. Rev. Neurosci.* **5:** 241–273.

Stanley G.B., Li F.F., and Dan Y. 1999. Reconstruction of natural scenes from ensemble responses in the lateral geniculate nucleus. *J. Neurosci.* **19:** 8036–8042.

Taylor D.M., Tillery S.I.H., and Schwartz A.B. 2002. Direct cortical control of 3D neuroprosthetic devices. *Science* **296:** 1829–1832.

Truccolo W., Eden U.T., Fellow M., Donoghue J.P., and Brown E.N. 2005. A point process framework for relating neural spiking activity to spiking history, neural ensemble and covariate effects. *J. Neurophysiol.* **93:** 1074–1089.

Wessberg J., Stambaugh C.R., Kralik J.D., Beck P.D., Laubach M., Chapin J.K., Kim J., Biggs S.J., Srinivasan M.A., and Nicolelis M.A. 2000. Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. *Nature* **408:** 361–365.

Wilson M.A. and McNaughton B.L. 1993. Dynamics of the hippocampal ensemble code for space. *Science* **261:** 1055–1058.

Wirth S., Yanike M., Frank L.M., Smith A.C., Brown E.N., and Suzuki W.A. 2003. Single neurons in the monkey hippocampus and learning of new associations. *Science* **300:** 1578–1581.

Wu W., Black M.J., Mumford D., Gao Y., Bienenstock E., and Donoghue J.P. 2004. Modeling and decoding motor cortical activity using a switching Kalman filter. *IEEE Trans. Biomed. Eng.* **51:** 933–942.