ELSEVIER

Journal of Neuroscience Methods 58 (1995) 127–141

JOURNAL OF NEUROSCIENCE METHODS

A method for detecting the time course of correlation between single-unit activity and EMG during a behavioral task

Andrew B. Schwartz [a,*], James L. Adams [b]

[a] *Division of Neurobiology, Barrow Neurological Institute, 350 W. Thomas Road, Phoenix, AZ 85013-4496, USA*
[b] *Department of Physiology, Northwestern University Medical School, 303 E. Chicago Avenue, Chicago, IL 60611, USA*

Received 19 October 1993; revised 4 May 1994; accepted 14 September 1994

Abstract

The chance that a change in excitability of one neuron leads to a change in excitability of another is likely to vary within a single volitional act. This temporal variability in functional connectivity is impossible to assess with standard analytical procedures to accurately that measure the correlation between such elements. This reports describes a technique designed to overcome this limitation by expressing a correlation measure calculated repeatedly in short epochs throughout a behavioral trial. The activity of two elements, a motor cortical neuron and a shoulder muscle, that might take place during a drawing task was first simulated so that the correlation could be manipulated. Various correlation algorithms (standard cross-correlation, spike-triggered average, impulse–response function, impulse–response surface) were tested with these data. Spike trains from a monkey's motor cortex and rectified EMG from its posterior deltoid muscle were compared using the same techniques and shown to have a correlation that changed in a characteristic manner throughout a task that required the monkey to draw a sinusoid.

Keywords: Spike-triggered average; Time-varying correlation; Effective connectivity; Motor cortex

1. Introduction

Correlation techniques are commonly used to deduce functional connectivity within neuronal circuits and between neurons and muscles. If a correlation between two elements exists then, depending on the timing relation between them, their possible location in a circuit and their excitation or inhibition upon each other can be obtained (Perkel et al., 1967b; Knox, 1974; Kirkwood, 1979; Aertsen and Gerstein, 1985; Houk et al., 1987). However such conclusions are subject to error because a consistent timing between elements may be mediated by different direct and indirect circuits. At every synapse there is a large convergence of presynaptic input. The probability that a single presynaptic impulse will elicit a post-synaptic element is dependent on the excitability of the other presynaptic elements. Because the threshold of membrane voltage at which an action potential is elicited is non-linear, an anatomic connection does not guarantee functional connectivity. This functional connectivity is time dependent since the activity of the population of presynaptic inputs is likely to vary within the observation. Standard techniques assume that data collected at one point in time during the observation is inter-related the same way at each point in time. The new technique reported here makes it possible to study correlation between variables that varies during the observation period.

This problem has been approached in previous work using joint–post-stimulus time histograms (JPSTH) (Palm et al., 1988). Coincidence in the time of action potentials recorded from different pairs of cells was measured as sensory stimuli were applied. This technique resulted in a 3-dimensional surface from which the time variability of the correlation between spike trains could be deduced. Various normalization procedures can be applied to separate the coincidence due to joint stimulus-related modulation from the 'intrinsic neuronal dependencies' representing 'effective connectivity'. These two contributions to coincidence have

* Corresponding author. Tel.: (602) 406-3389; Fax: (602) 406-4172.

SSDI 0165-0270(94)00167-7

been referred to as the 'secondary and primary effect' elsewhere (Kirkwood, 1979). This analysis depends on the fact that spike trains are point processes – within each bin there either is or is not a spike occurrence. This technique cannot be applied easily to the comparison of a spike train to a continuous signal such as that encountered when recording whole-muscle EMG.

We wished to examine the relation between single-unit activity in the primary motor cortex to EMG activity of the proximal arm muscles as a monkey performed drawing tasks (Schwartz, 1992). Specifically, we wanted to know whether a given neuron–muscle correlation would vary in time. It is known in non-human primates (Landgren et al., 1962; Kuypers, 1981), that although both a multi- and monosynaptic pathway from motor cortex to the motoneurons of the distal extremities exist, the pathway to the proximal arm muscles is solely multisynaptic. Since the question we are asking is primarily functional instead of anatomical, it was necessary to develop a technique that was sensitive enough to detect multisynaptic changes in activity that consistently covaried, to separate the covariation due to common modulation from that related to the interaction between the recorded neuron–muscle pair and to observe this covariation in time. This technique measures the temporal association between excitability changes in two elements and the results of the analysis could yield insight as to how one element contributes to the changes in excitability of the other.

One of the problems in detecting correlation between two time-varying neurological signals is that any periodic variation within the individual signals will tend to contribute to the correlation between the signals if the period of modulation is on the same order of magnitude as the analysis epoch. This type of common modulation was especially important in our experiments since the monkeys performed sinusoidal drawing movements with periods of 0.3–1 s. Both the cortical cells and the proximal arm muscles tended to increase their activity once per cycle. This periodic modulation will be manifest as a prominent feature in the result of standard correlations. This broad effect may obscure other structure in the cross-correlation ('secondary effect') (Knox, 1974; Kirkwood, 1979).

Periodic contribution to the cross-correlation can be removed by using the *impulse–response function* (IRF) to describe the interaction between two signals. This function was originally developed in linear systems analysis to describe transfer functions. The impulse–response is the idealized response of a system to a pulse of constant amplitude that has an infinitely short duration. Such a system was used by Humphrey (1972) for relating motor cortex activity to the generation of torque. The technique was applied to neuronal activity by Soechting et al. (1978) as a measure of correlation between an input (interpositus activity) and an output (red nucleus activity) and reviewed by Houk et al. (1987). Although this function was developed for studying linear systems, we use it here as a way to reduce periodic changes in the input signal that might contribute to the comparison between input and output.

In order to visualize the changes between motor cortical activity and proximal-arm EMG that might take place, we calculated the impulse–response function in overlapping 350-ms windows throughout the task. This resulted in a surface we termed the 'impulse–response surface' that showed the evolution of the impulse–response function in absolute time through the drawing task. This technique as well as standard cross-correlation and spike-triggered averaging were compared and tested with simulated data. In the simulation, the correlation between spike and EMG could be specified for particular epochs.

2. Methods

2.1. Simulations

A 2-s spike train was simulated in which firing frequency was modulated sinusoidally. A probability–density function, consisting of 300 values ($0.5 < P < 0.9$), was first mapped to the 2-cycle sinusoid. A random number between 0 and 1 was chosen for each of the 300 bins. If the generated probability for the bin exceeded this value, a spike was placed in the bin. This resulted in a sinusoidally modulated spike train where the instantaneous discharge frequency varied between 4 and 167 Hz. EMG activity was then simulated by associating a TTL pulse (1-ms duration) with each spike. In one condition the pulse was concurrent with the spike. In other conditions, a gaussian distribution determined the placement of the EMG pulse relative to the spike. The standard deviation of the distribution could be varied to control the uncertainty of the time span between the spike and the EMG trigger. The pulse was used as an input to the recording amplifier and filters used to record the actual data. The simulated data made it possible to test the analytical techniques with conditions where the correlation between spike and EMG varied in a known manner.

2.2. Behavioral task and data acquisition

A rhesus monkey was seated in front of a graphics monitor with its head fixed and one arm restrained. With its index finger on a touch screen placed over a graphics monitor, the animal traced sinusoids that spanned the entire horizontal extent of the screen oriented in the frontal plane. The sinusoids had peak-to-peak amplitudes that ranged from 3 to 12 cm and spatial frequencies between 0.053 and 0.21 cycles/cm.

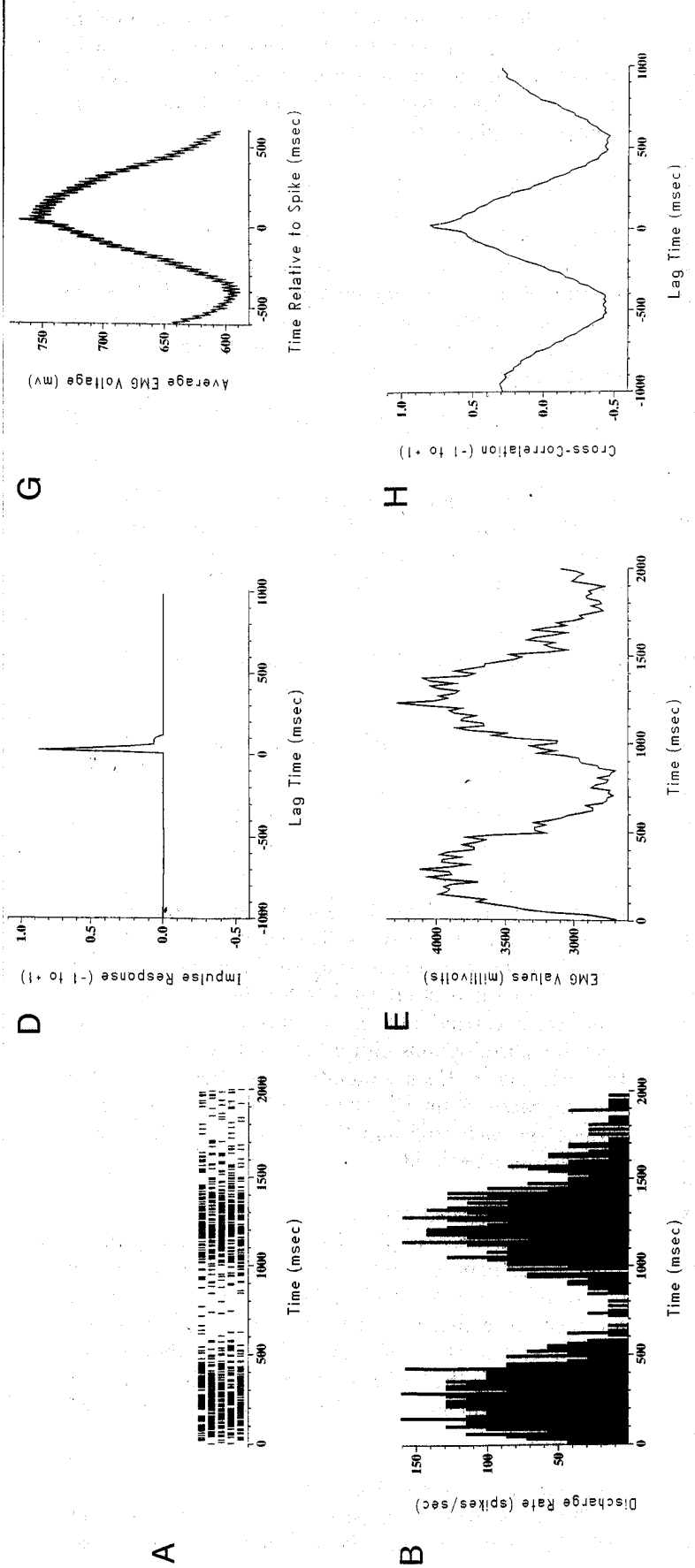


Fig. 1. Generated spike train. Spike trains were generated as described in the text from random numbers and a sinusoidal change in probability. The sine function had a 1-s period and five 2-s trials were generated. A raster representation of each trial is displayed in (A) with a tick mark indicating the position of each impulse in the trial. This simulation is represented in a histogram with 14-ms bins in (B). The units on the ordinate are in impulses or spikes per second. The autocorrelation of these impulses is shown in (C) and shows that the predominant period of modulation is 1000 ms. Square-wave pulses generated electronically and placed after each simulated spike were filtered and used to simulate EMG. The impulse–response to a single pulse is shown in (D) and is identical to the analog display of the EMG amplifier and filter. Five trials of the 2-s sequence were accumulated and averaged in (E). The autocorrelation structure (F) of the EMG signal has a period of 2 s. The simulated spike and EMG data were compared in (G–I). G: the EMG occurring 600 ms before and after each spike was averaged to a form a STA that peaked at a lag (τ) of 30 ms. The overall structure has a period of 1 s. H. the cross-correlation between the spike and EMG data has a peak at $\tau = 30$ ms and the low-frequency component has a 1-s period. I: the low-frequency component is removed from the correlogram when using the impulse–response function, leaving a single peak at $\tau = 30$ ms.

Each of 5 different waveforms were traced from both the right and left sides of the screen. The sinusoids were presented in a random order within blocks where each of the sinusoids was presented once. Each sinusoid pattern was repeated 5 times in this randomized block design. The experimental design and recording methods are described in more detail in an earlier paper (Schwartz, 1992).

Single-unit motor cortical activity was recorded with an extracellular microelectrode and the time of spike occurrence relative to the start of the trial was recorded using an amplitude discriminator and timer/buffer interface. All data were analyzed on a trial-by-trial basis. For the auto- and cross-correlations the spike counts were binned. A method used to convert the spike count per bin to a continuous spike density function (Richmond et al., 1987; Schwartz, 1992) was compared to a standard method of dividing the bin count by the bin width. Since there was no difference between the two methods for these data, the pattern of neuronal activity was converted to a series of instantaneous frequencies by binning the spikes and dividing the number of spikes per bin by the bin duration. This simple method may be imprecise for low spike rates when long interspike intervals span the boundaries of the bins. There are other problems associated with converting a point process into a continuous signal (French and Holden, 1971). These are due primarily to the binning process which can lead to discontinuities in frequency if the bin width is too small and to high frequency attenuation if too wide. However, for the time scale examined in this study, these issues are not problematic.

EMG activity recorded with bipolar electrodes was amplified, rectified, and integrated with a Paynter filter ($\tau = 50$ ms) and recorded at 10-ms intervals. This filter, in addition to removing high frequency components from the signal, produces a delay. Its response to an impulse (1-ms TTL pulse) begins 6 ms after the pulse, reaches a peak in 20 ms, decays rapidly in the next 20 ms and then to its baseline value with a slower time constant over the next 60 ms (Fig. 1D). Compared to an ideal filter simulated with a 10-ms square wave, the Paynter filter shifted the data by approximately 25 ms and reduced the high-frequency components in the correlations. Activity from the following muscles was recorded: anterior, medial and posterior deltoids, long head of triceps and biceps.

The fast-Fourier transforms (FFT) that were employed later in the analysis were most efficient if the input vector length was a product of small prime numbers. In the simulation we chose to divide the data into 221 bins. The real data were divided into 143 bins. The mean duration over the 5 repetitions of each condition was found and the individual trials were normalized in time using a spline function to this mean duration. This resulted in bin widths of 14 ms for the simulated data and 10 ms for the real spike data. The EMG data although collected at 10-ms intervals, was normalized in the simulations to the same number of points as the number of spike bins.

2.3. Analyses

2.3.1. Correlations

Auto- and cross-correlations were calculated conventionally using the binned spike count of the cortical unit and the rectified-filtered waveform of the EMG. The cross-correlations were estimated using the following formula:

$$r_{xy}(\tau) = c_{xy}(\tau)/\left[c_{xx}(0)c_{yy}(0)\right]^{1/2} \tau = 0, \pm 1, \ldots, \pm T \quad (1)$$

where τ is the lag represented as some time interval or bin. The terms $c_{xx}(0)$ and $c_{yy}(0)$ are the autocovariance for x and y (equivalent to the sample variance of each term) at zero lag. The cross-covariance function

$$c_{xy}(\tau)$$

was estimated by the following equations, according to the positive and negative lag values:

$$c_{xy}(\tau) = \frac{1}{n} \sum_{t=1}^{n-\tau} (x_t - \bar{x})(y_{t+\tau} - \bar{y}) \quad \tau = 0, 1, \ldots, T \quad (2)$$

$$c_{xy}(\tau) = 1/n \sum_{t=1-\tau}^{n} (x_t - \bar{x})(y_{t+\tau} - \bar{y})$$

$$\tau = -1, -2, \ldots, T \quad (3)$$

T is the maximum lag value and $\bar{x}$ and $\bar{y}$ are the mean x and y values, respectively. The estimated autocovariance function was found using the same equations by substituting the desired variable for the other. Correlations were carried out on a trial-by-trial basis. The resulting correlograms plotted the different values of r for each lag τ. Each simulation and behavioral trial was repeated 5 times. These calculations were performed on each trial and the results were then averaged across repetitions.

2.3.2. Spike-triggered average

The occurrence of a spike served as a trigger to align a segment of EMG data ranging from δ ms before to δ ms after each spike. This was carried out for each spike that was at least δ ms from the beginning or end of the trial. An average was then calculated by adding the values of corresponding bins for all the EMG segments and dividing by the number of segments.

2.3.3. Impulse–response function

The cross-correlation between the spike train and the EMG data and the autocorrelation of the spike

train were converted from the time to the frequency domain using a FFT to yield the corresponding cross-spectrum and autospectrum (Eq. 4). This was done by applying a Hamming window to the correlations to remove extraneous frequencies ('leakage' (Ramirez, 1985)) encountered by using a non-infinite signal. The impulse–response function is obtained by dividing the cross-spectrum by the autospectrum of the spike train (Eq. 5) and taking the inverse transform (Eq. 6) of the result to convert it to the time domain (Soechting et al., 1978). It should be noted that the correlograms in the time domain begin with negative lag values and since the FFT is constructed to operate on an input with a range of $0–T$, it is necessary to exchange the negative and positive halves of the time series after taking the inverse transform.

$$c_{xy}(\tau) \leftrightarrow c_{xy}(f);\ c_{xx}(\tau) \leftrightarrow c_{xx}(f) \tag{4}$$

$$h(f) = \frac{c_{xy}(f)}{c_{xx}(f)} \tag{5}$$

$$h(f) \leftrightarrow h(\tau) \tag{6}$$

3. Results

3.1. Simulations

Five repetitions of a 2-cycle sinusoidally modulated spike train are simulated in Fig. 1. The raster in Fig. 1A shows the time of occurrence of each spike relative to the beginning of the trial with a vertical tic mark. Each repetition is shown on a different horizontal line. The pattern is similar between trials, but the exact placement of the spike in time varies randomly.

These data are collapsed into a histogram in Fig. 1B and plotted with the same time scale as Fig. 1A. This function was calculated with 14-ms bins and shows the probability of discharge (in units of firing rate) at a given instant in the task. The simulated cell was most likely to fire at 250 and 1250 ms after the beginning of the trial.

An autocorrelation of this spike train was calculated for each repetition and the average across repetitions is displayed in Fig 1C. This calculation expresses the likelihood that given a spike at time 0, another spike would occur at a latency of τ ms. The abscissa is lag time and the peak at $\tau = 0$ results from the certainty that a spike always occurs at $\tau = 0$. The structure in this correlation histogram (correlogram) is primarily the result of the sinusoidal modulation of the spike train, which has a period of 1 s. There are also smaller peaks near the origin showing that there are other inherent periodicities in the signal. The first peak is 60, the second 150, and the third 210 ms from the origin. Evident in the autocorrelations of the individual repetitions are sharper peaks about 30 ms apart. This suggests that there is a 30-ms periodicity in the spike train.

The simulated EMG activity resulting from this simulated spike is shown in Fig. 1D. A TTL trigger pulse generated after each spike was passed through the recording amplifier and Paynter filter. This analog signal was collected with the laboratory interface and stored as simulated EMG. The response of the filter to the TTL pulse has a rapid onset with a slightly prolonged initial decay followed by a longer decay when the response was within 10% of its peak amplitude. All 5 repetitions of the spike train were used in this simulation, with the result displayed in Fig. 1E as the average. The autocorrelation of this EMG is shown in Fig. 1F. It has the same low-frequency component as the spike train autocorrelation and corresponds to the rhythmic signal in Fig. 1B.

Three different methods of displaying the relationship between these signals are shown in Fig. 1G–I. The spike-triggered average (STA) of this simulation is shown in Fig. 1G. EMG segments beginning 600 ms before and ending 600 ms after each spike were used to construct the average. In the average there is a prominent peak at approximately $\tau = 30$ ms and the low-frequency component of the signal is the same as that seen in Fig. 1A–C, E–F. The STA is built from individual spike occurrences in contrast to the cross-correlation which is based on binned data. Therefore the features in the STA will tend to be less smooth than those in the corresponding cross-correlogram. The sharp, higher-frequency peaks superimposed on the low-frequency component may be due to the 30-ms component of rhythmic modulation in the spike train seen in its autocorrelation (Fig. 1C) and to the 25–30 ms delay of the EMG pulse to the simulated spike. The cross-correlation (Fig. 1H) was calculated trial-by-trial using the instantaneous frequency of the spike train and the simulated EMG data. Since the spike data were binned at 10-ms intervals and the analog EMG was also sampled at this interval, both signals could be considered sampled at a 100 Hz rate. The peak of the correlogram occurs at $\tau = 30$ ms. This is the primary effect due to this deterministic simulation in which the same pulse of EMG follows every spike at a fixed latency. The position of the peak suggests that, on average, the EMG pulse peaks 30 ms after the spike. The periodicity seen in the autocorrelation is also evident here as the secondary effect due to the periodic modulation of each signal. In Fig. 1I the impulse–response function for these data is shown. Again the prominent peak at $\tau = 30$ ms is apparent. Notice there is no periodic component to the impulse–response correlogram. Strictly speaking, the relationship depicted by the impulse–response func-

tion is not a correlation, but we will use a looser interpretation of the term to the result of this calculation. A single peak is clearly discernable and represents the 'effective connectivity' between the simulated cell and muscle. This is much more useful in this regard than either the STA or conventional cross-correlation. The ability of the IRF to recover the idealized impulse can be ascertained by comparing the response shown in Fig. 1I to that of Fig. 1D. The impulse–response in Fig. 1D is the output of the rectifier-Paynter filter system to a single 1-ms TTL impulse.

The way the correlation changes in time is illustrated in the impulse–response surface of Fig. 2. The EMG and spike data were divided into overlapping 350-ms segments. The beginning of each successive segment was offset from its predecessor by 10 ms. The impulse–response function was calculated in each data segment and this correlogram was plotted perpendicular to the time axis at the point in time that the data segment began. This dimension along the correlogram, perpendicular to the time axis, has units of lag time τ and ranges from ± 175 ms. The time axis shows absolute time T, through the trial. The peak in the impulse–response again occurred at $\tau = 30$ ms lag and was nearly constant throughout the trial, forming a ridge in the surface. The upper trace in the panel behind the surface displays the spike histogram and the lower trace of the panel is the average EMG.

Simulated EMG was again derived from filtered TTL pulses using the spike train shown in Fig. 1A. However the EMG trigger pulses were placed relative to each spike using a gaussian distribution centered 20 ms after each spike with a SD of 60 ms. The shape and amplitude of all the EMG pulses were identical. This condition was modeled to represent a multisynaptic linkage between the spike and muscular activity. The resulting EMG pattern is very similar to that of Fig. 1E. The STA of these data is shown in Fig. 3A which shows that the distinction between the secondary modulation and the primary effect is not distinct when the correlation between the two signals is weaker. The cross-correlation between the spike train and EMG activity is shown in Fig. 3B. The primary correlation, which is now of smaller amplitude, is obscured by the periodic secondary effect. The utility of the impulse–response function is demonstrated in Fig. 3C where the primary correlation is clear and the secondary effect is absent.

The impulse–response surface calculated with the data of Fig. 3A–C is shown in Fig. 3D. The peak is wider and more variable than that seen in Fig. 2.

Intermittent correlation between spike and muscle

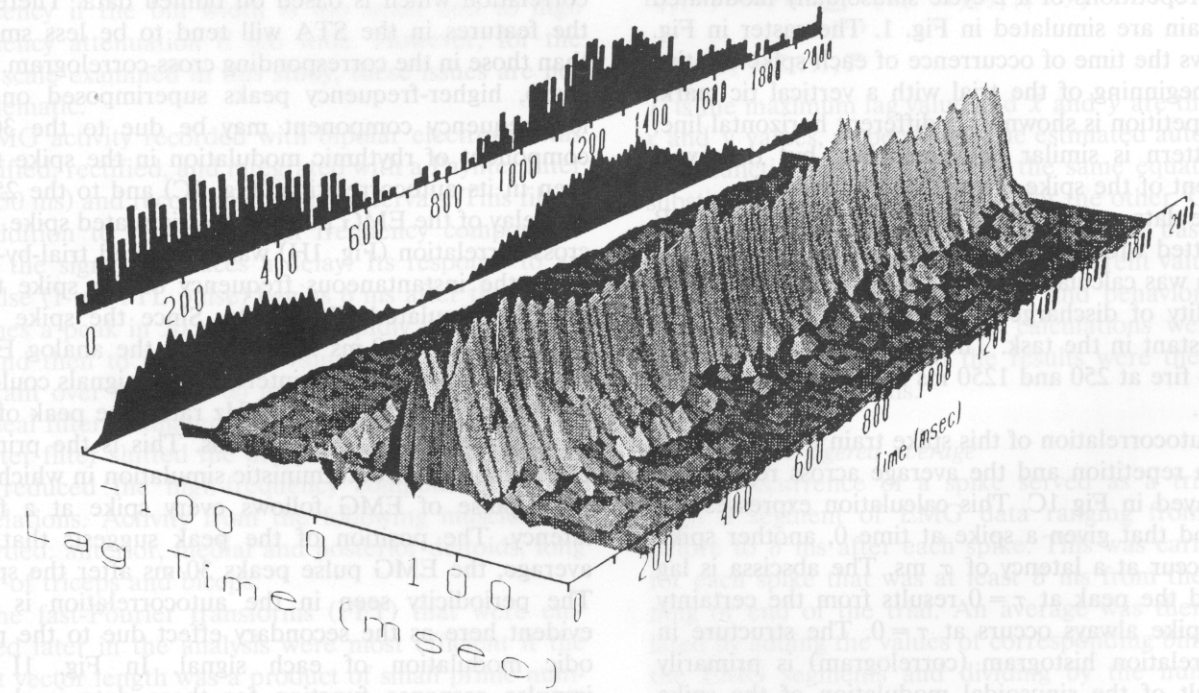


Fig. 2. Impulse–response surface. The top trace on the back panel is a histogram of the simulated spike train. The lower trace is an average of the resulting EMG signal. The time axis of the surface spans the duration of the spike sequence, ranging from $T = 0$ to 2000 ms. The surface itself was calculated from $T = 175 - 1825$ ms to allow for a maximum lag of 175 ms between spike and EMG. The surface was produced by calculating the impulse–response function in successive 350-ms sliding windows. A ridge in the surface of fairly constant amplitude and constant lag of 30 ms, showed that the simulation resulted in a stationary relation between the spikes and EMG.

activity was simulated in Fig. 4. The spike activity and EMG are completely correlated in a window from $T = 1000$ to $T = 1500$ ms. At all other times through the trial the two signals are related by the same gaussian function used in Fig. 3. The STA in Fig. 4A is very similar to that of Fig. 3A except for a small narrow component at $\tau = 40$ ms riding on the large low-frequency waveform. Any component resulting from primary interaction is obscured by the large secondary modulation in the cross-correlation of Fig. 4B. The primary interaction is evident in the impulse–response shown in Fig. 4C. There is little difference between these waveforms and those of Fig. 3 where there was no window of high correlation. These waveforms (A–C) are more similar to those of Fig. 3 than those displayed in Fig. 2 because in 3/4 of the trial the interaction between the spike train and the EMG signal was governed by the wide gaussian distribution.

The temporal evolution of the spike–EMG relationship and the change in spike–EMG correlation is shown in Fig. 4D,E. This surface shows the epoch in the trial when the two signals are more correlated. There is a

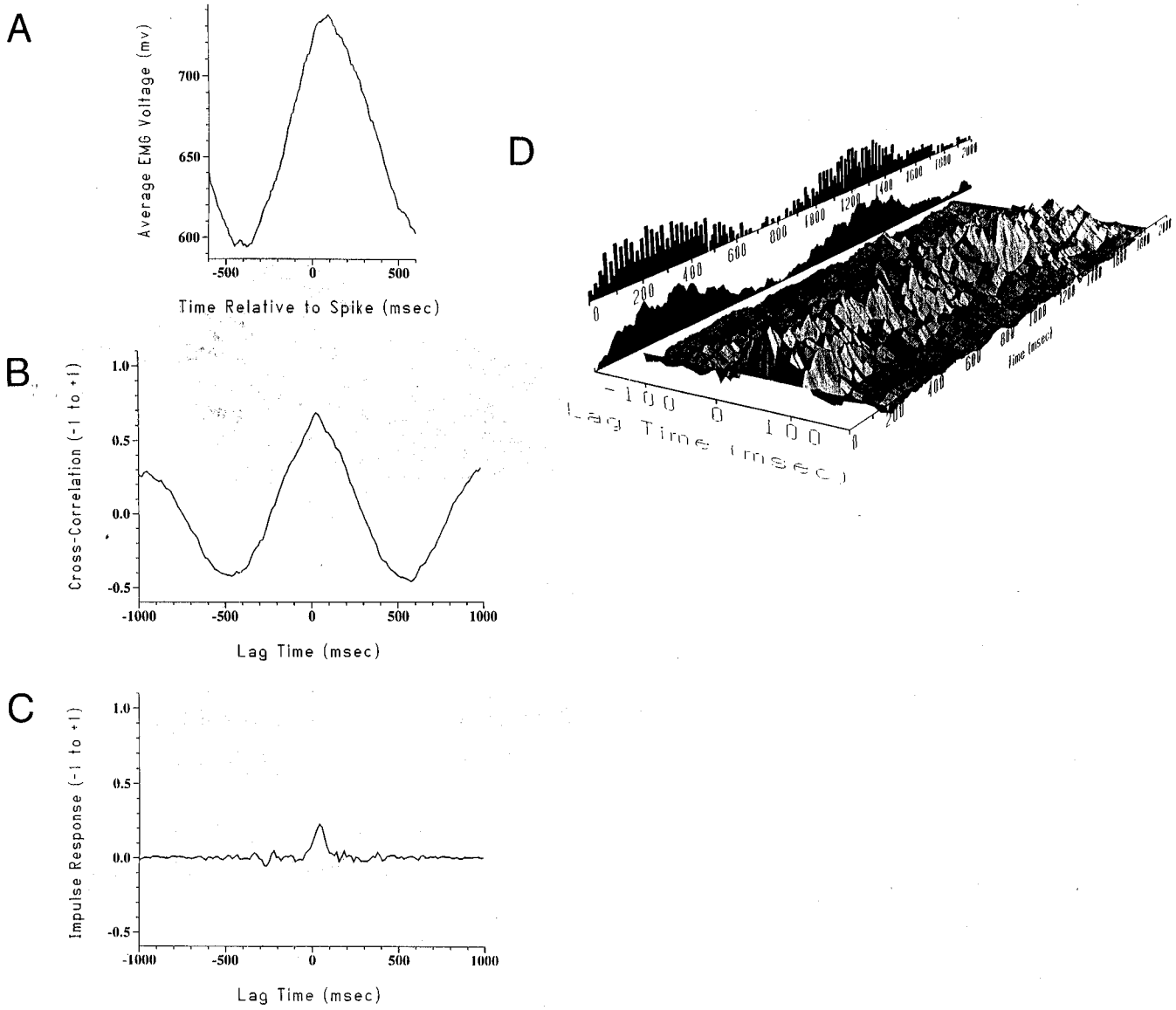


Fig. 3. Comparison of simulated spike train to EMG pulsed generated with a gaussian distribution. A pulse of EMG was triggered after each spike according to a gaussian distribution centered 20 ms after each spike. The distribution had a standard deviation of 60 ms. A: the STA of this simulation shows little evidence of a primary effect. The low-frequency component of the average representing the secondary effect has a 1-s period. B: there is also no evidence of primary correlation in the cross-correlogram. C: the impulse–response function shows a peak of positive correlation at $\tau = 40$ ms that is somewhat wider and smaller than that seen in Fig. 1. The impulse–response surface of this gaussian simulation is shown in (D). The ridge in the surface structure is still evident, but is wider, of smaller amplitude and is more variable for the less correlated result of this simulation.

ridge that begins at $T = 850$ ms, peaks at $T = 1250$ and ends at approximately $T = 1650$. The epoch of increased correlation in the simulation began at $T = 1000$ ms and ended at $T = 1500$ ms. Since the sliding window in which the impulse function was calculated was ± 175 ms, the boundary between the more correlated and less correlated signals was stretched in time. When the sliding window is completely contained in the epoch of increased correlation, the value of the impulse–response will be the greatest. As the window advances beyond this epoch, the impulse–response function decreases gradually. The use of a shorter duration window would make it possible to detect the onset of a correlated epoch more precisely (see Fig. 7).

3.2. Empirical data

The raster and histogram of activity from a motor cortical cell recorded while a monkey traced the sinusoid in Fig. 5A are shown in B and C. The sinusoid was traced from left to right. This response encodes the direction of arm movement and is typical of many cells found in the motor cortex. An analytical approach has been developed (Schwartz, 1992) which shows that

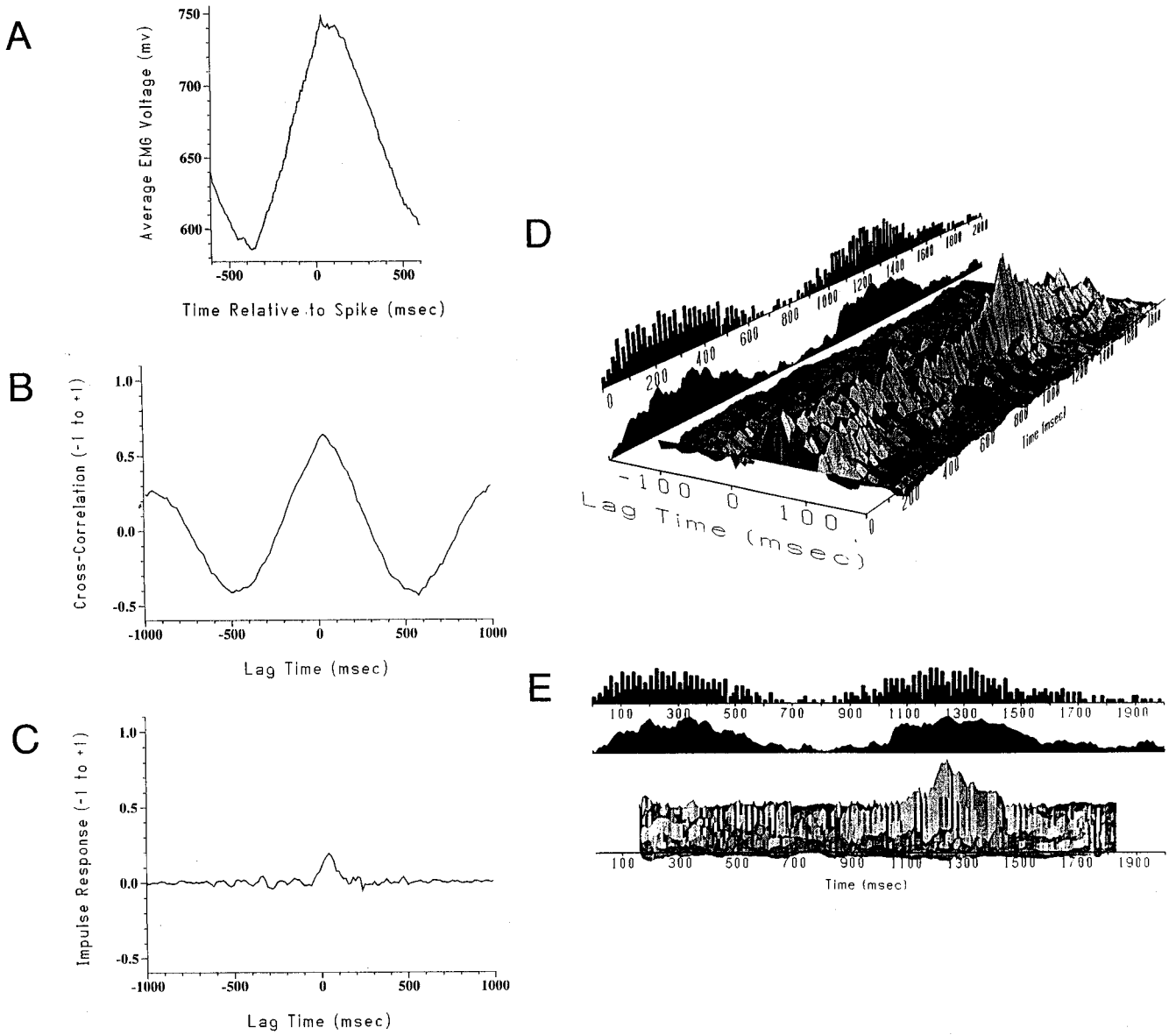


Fig. 4. Simulated transient correlation. Complete correlation was simulated during each trial in the epoch form $T = 1000$ to $T = 1500$ ms. At all other times the gaussian function was used to place the EMG pulses. A: STA of intermittent correlation. B: cross-correlation of intermittent correlation. C: impulse–response function of intermittent correlation. In each case the result is similar to the gaussian simulation illustrated in Fig. 3. The impulse–response surface of this intermittent correlation is shown in (D) and (E). The two views of the surface show that the correlation between simulated spike and EMG activity increased from approximately $T = 1100$ to $T = 1400$ ms.

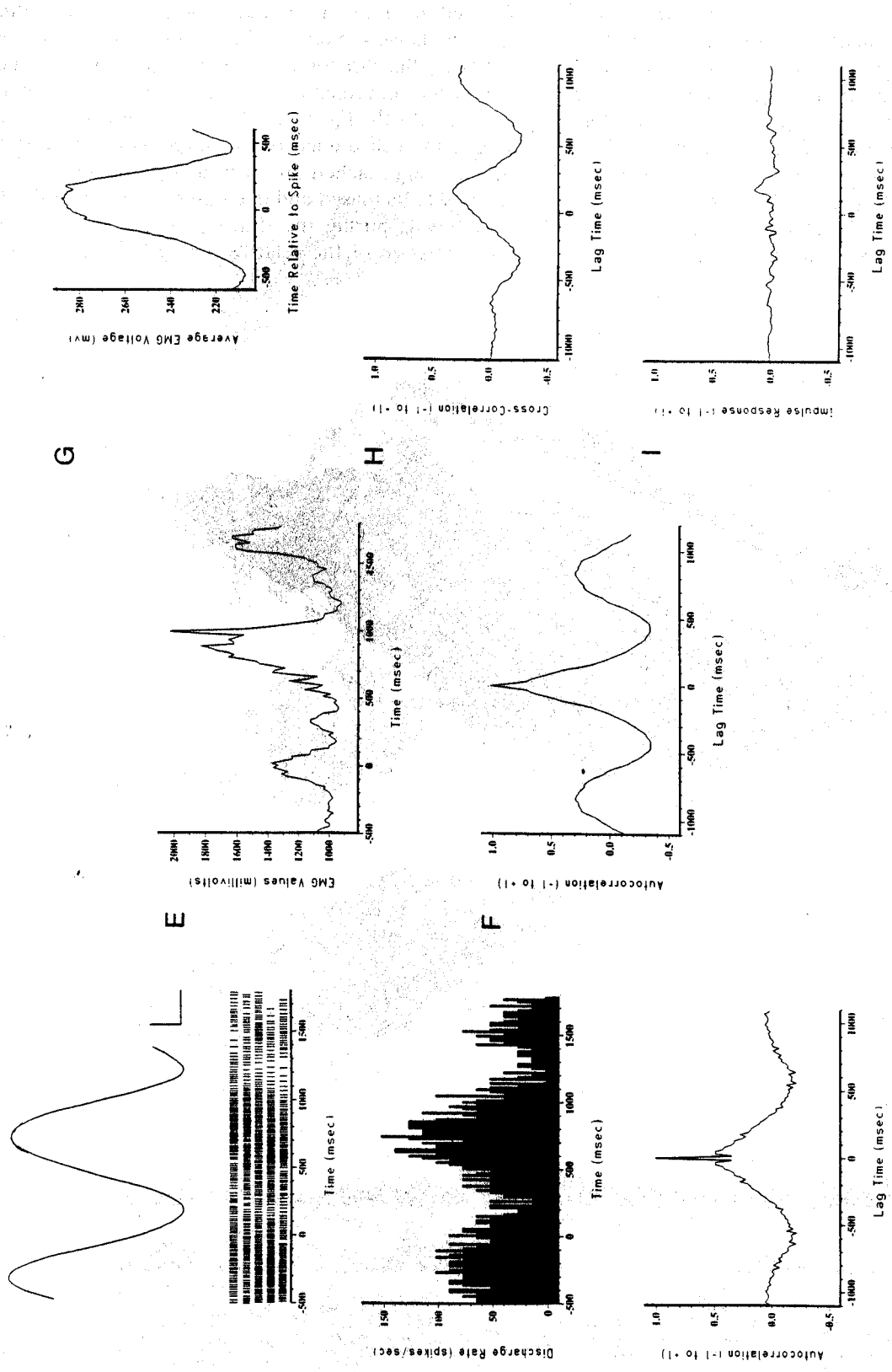
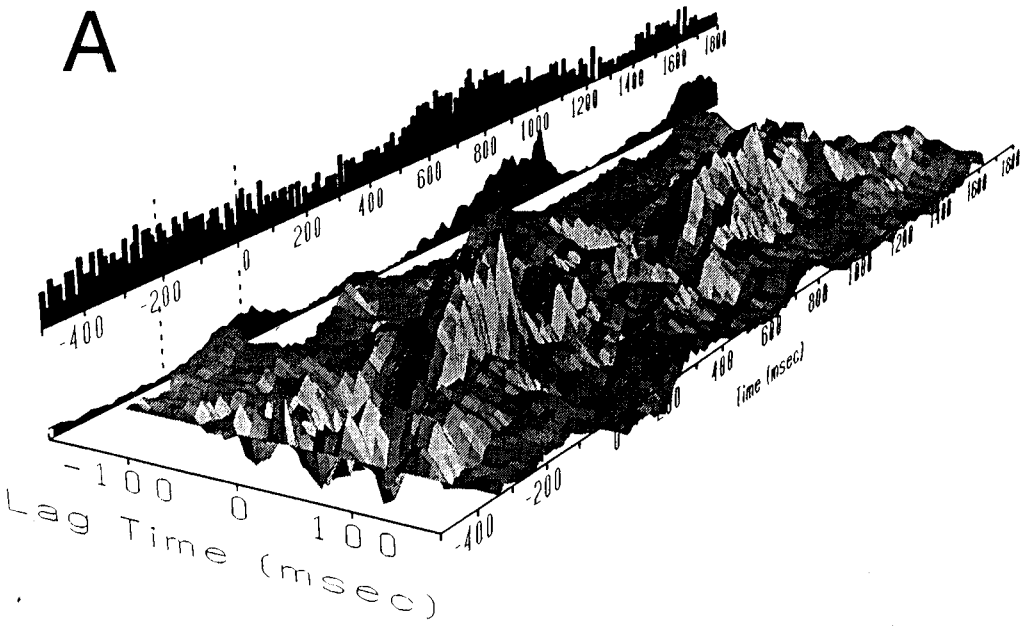


Fig. 5. Empirical results. A: the finger trajectory was recorded in $x - y$ coordinates and averaged over 5 repetitions of the task. The calibration bars are 2 cm. B: the raster was constructed from the response of a motor cortical cell during the task. Each repetition was aligned to movement onset. The dotted line in each trial marks the time that the sinusoid was presented. C: the histogram of the cell's response during the task has a bin width of 10 ms. D: the autocorrelation of the cell's activity has a short-term component related to the refractory period of the cell. The overall period of the signal is approximately 1 s. The EMG shown in (E) is the rectified-integrated response of the posterior deltoid muscle averaged over the 5 repetitions. In (F) it can be seen that the autocorrelation of the response has a period of 800 ms. A comparison of empirical spike and EMG responses are displayed in (G-I). G: the STA of the spike and posterior deltoid response shows a small primary component at $\tau = 180$ ms riding on a much larger lower frequency component with a period of about 900 ms. H: there is no evidence of a primary correlation component in the cross-correlogram. The secondary component of the correlogram has a peak at $\tau = 180$ ms and a duration of 1 s. This suggests a positive correlation between the cell and EMG activity from 100 ms before to 350 ms after each spike. I: there is a single peak in the impulse–response function suggesting that the EMG tends to follow the spike with a latency of 180 ms.

these cells tend to fire fastest in their 'preferred' direction of arm movement. The preferred direction for this cell's activity was at 50° from the horizontal (when the monkey moved its finger up and to the right). The EMG of the posterior deltoid muscle (Fig. 5E) was analyzed in a similar fashion. The peak activity of this muscle occurred when the direction of movement was 157° (up slightly and to the left).

An analysis (Schwartz, 1992) of the directional information in the activity pattern showed that this cell changed its rate of discharge about 160 ms before movements were made in the corresponding direction. Similarly, the direction-related muscle activity tended to precede movements in the corresponding direction by 70 ms. As the figure was traced rightward across the screen, the cell tended to discharge as the movement direction approached 50°. The activity of the cell and the muscle increased and decreased 3 times at 800–900 ms intervals during the drawing of the sinusoid. An approximation of the relative timing between the two



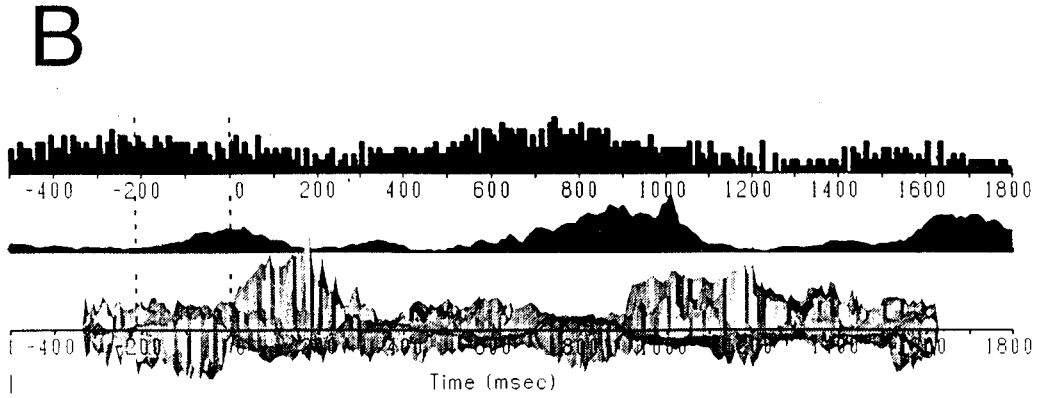


Fig. 6. Impulse–response surface: empirical data. Two views (A and B) of the surface are presented. The dotted line at $T = -200$ ms represents the average time of the sinusoid presentation. The data are aligned to movement onset. The top trace in the back panel is the average EMG, the lower is a histogram of the spike activity. There are two epochs of positive correlation evident in the surface. One epoch of correlation is located in time at $T = 0$–300 ms, the other ranges from 900–1200 ms.

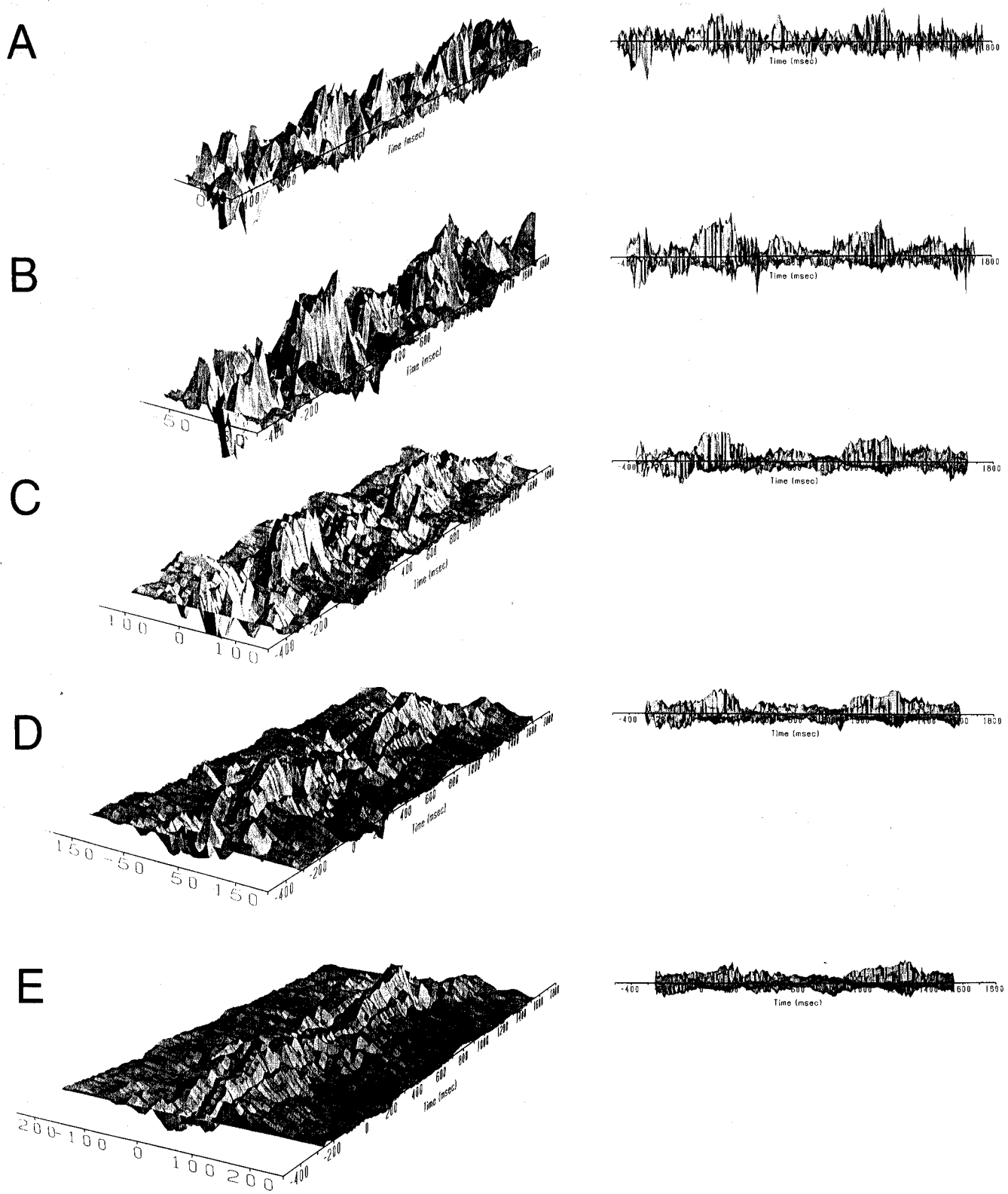


Fig. 7. Effect of window size. A series of different duration windows were used in the calculation of each surface. The window in (A) was 110 ms, (B) 210 ms, (C) 310 ms, (D) 410 ms and (E) 510 ms. The surface features became evident in (B) and were gradually washed out as the duration of the window increased.

can be ascertained by noting that the middle cycle began to increase in neuronal activity about 250 ms after movement onset and at about 450 ms for the muscle (Fig. 5E). Based on the histograms, the cell activity would be expected to lead that of the muscle by approximately 200 ms.

If the neuron and muscle possessed the same preferred direction, the latency between their two activity patterns would be expected to be 90 ms. However, since the direction of movement associated with maximal muscle activity and that associated with maximal neuronal activity occur sequentially, an additional latency must be considered. These considerations as well as the pitfalls of using histograms to describe relative activity illustrate the utility of a method that allows the relationship between spike and muscle activity to be directly visualized in absolute time.

The autocorrelation (Fig. 5D) of the spike activity reflects the 1-s period of modulation in discharge rate. The EMG activity had a period of 800 ms (Fig. 5F). STA (Fig. 5G) of these signals suggests that the cell tends to fire at the peak of the cyclic muscle activity. The cross-correlation (Fig. 5H) shows positive correlation for the time interval from 100 before to 350 ms after each spike with a peak at 180 ms and a cyclic period of 1 s. This measure of correlation shows that the EMG activity tended to follow the spike by 180 ms (ignoring filter characteristics). This timing relation can be more clearly recognized in the impulse–response function (Fig. 5I) in the absence of the secondary effects caused by the periodic modulation of the signals. Now there is only a single peak at $\tau = 180$ ms.

The plots in Fig. 5G–I show the average correlation of these signals throughout the task. In order to visualize the changes in correlation within the task it is necessary to plot this correlation on an absolute as well as a relative time axis. The surface shown in Fig. 6 expresses the result of this analysis. Behind the surface the spike and EMG histograms are plotted in absolute time. The origin on the time axis coincides with movement onset. The dotted vertical line at approximately $T = -200$ shows the average time in the trial when the sinusoid was presented. There were two prominent epochs of correlation represented by this surface as elevated ridges at $T = 0$–300 ms and at $T = 900$–1200 ms. These areas are similar in that they both begin approximately at $\tau = 0$ and have a duration of 60–100 ms. The relation of this structure to the spike and EMG pattern can be seen better with the orientation shown in Fig. 6B. The peaks in correlation are not simply related to excitability; the epochs of positive correlation span transitions where both the cell and muscle go from an active to inactive state. This increase in correlation occurs in the same part of the traced sinusoid starting before and ending after the maximum vertical portion of the figure. This analysis shows that it is likely that EMG and spike activity are associated only when the finger is at the peak of each cycle in the sinusoid.

The size of the window used to calculate the impulse–response surface may affect the shape of the surface. If there is a punctate event where the correlation is high surrounded by times where the correlation is low, the window size can either smooth and reduce the size of the surface feature if the window is too large or pick up noisy correlations if too few data points are used in a small window. A small window will also miss correlated events with large τs. Some of these effects are illustrated with the series of surfaces shown in Fig. 7. These are derived from the same data used in Fig. 5–6. The window size was varied from 110 to 510 ms in increments of 100 ms. In A the correlation lag varied from -50 to $+50$ ms and incorporated 11 data points. The surface is rather noisy but 3 epochs of correlation may be distinguished at times 0–300, 400–600 and 900–1200. When the window was increased to 210 ms (Fig. 7B), the correlation epochs became more distinct. The major features occurred from 0–300 and from 900–1200 ms. As the window was increased in C–E, these features were smoothed out and became less distinct.

4. Discussion

Since we were primarily interested in the long-term interactions between the signals in a multisynaptic system, we used relatively wide bin widths (10–16 ms) over the entire task which typically had a 2-s duration. If shorter time scale interactions were of interest, the bin width could be reduced. By minimizing the secondary effects, it is more likely that the resulting structure in the impulse–response function is due to more direct interactions between the observed motor cortical cell and EMG activity. The intervening neuronal structures and connections cannot be determined with this technique and is not the purpose of this investigation. Rather, we are interested in the tendency of the EMG to change consistently (increase or decrease) with changes in spike rate.

The technique of applying correlation analysis to examine the relation between the occurrence of a spike and an increase or decrease in the excitability of a muscle was first developed by Mendell and Henneman (1971) in spinal cord studies and employed by Fetz et al. (1976) to study the effect of motor cortical activity on distal arm muscle activity. The analysis was performed by using the spike occurrence to align each sweep of rectified EMG activity before averaging. If correlation between the two active elements existed, the resulting wave form would then show a transient upward or downward deflection corresponding respec-

tively to post-spike facilitation or inhibition. This technique became known as 'spike-triggered averaging' (Watt et al., 1976). These initial investigations, as well as most of the subsequent work using this technique, examined linkages that were primarily mono- or disynaptic since the intermediate synapses increase the variability in timing between the neuronal spike and any related change in muscle activity. As with most other averaging techniques (Moore et al., 1966) the post-spike histogram can be interpreted as the probability of excitability in the recorded muscle at a time τ after the occurrence of the recorded spike.

The result of cross-correlation analyses (Box and Jenkins, 1976; Chatfield, 1980) is a probability of joint occurrence between two events allowing for some intervening time lag. In the case of continuous signals, the 'events' can be considered increases or decreases in the amplitudes of the signals. In the present case the correlation measured represents the probability that a unit will increase its discharge rate in relation to the probability that EMG amplitude will increase. A positive correlation also exists when both probabilities decrease in tandem. The same interpretation applies to STA, although instead of dealing with two continuous signals the spike train is handled as a point process. Thus there may be a slight difference between the techniques because the probability of discharge is derived from binned data which introduces smoothing and some temporal uncertainty. As the STA is calculated, a spike that occurs without a corresponding change in EMG amplitude will tend to erode any wave form that is accumulating as the measure of post-spike muscle excitability. The resulting average is also a joint probability. The results of our simulations show that the STA and cross-correlation techniques provide outputs that are nearly identical. The differences between the two techniques in a recent report (Miller et al., 1992) was attributed to the separate filtering and sampling used for the data passed through each analysis. Unfiltered data collected at 0.25-ms intervals were used in the STA analysis and data filtered with a 10-ms time constant were sampled at 200 Hz in the cross-correlation analysis. The conclusion of this study was that two types of interaction between rubrospinal cells and muscles were occurring. The first, detected with STA, was a short-latency interaction and the second, described by the cross-correlation, was present at long, variable lags. It should be emphasized that the large differences between the cross-correlation results and those of the STA were a result of the way the data were conditioned, not to inherent differences in the analytical techniques.

Perhaps the most serious issue of applying correlation techniques to active elements during a movement task is that of stationarity. A process is considered stationary if its mean and higher order statistical moments (i.e., cross-covariance) are fixed over the time period for which the observations are made (invariant in time) and if there are no periodic variations in the signal (Cox and Lewis, 1966; Perkel et al., 1967b; Perkel et al., 1967a; Box and Jenkins, 1976; Glaser and Ruchkin, 1976). An example of a stationary process would be the thermal noise encountered when measuring the impedance across an electrical resistor held at a constant temperature. It is unlikely, however, that task-related neuronal or EMG activity is stationary throughout a volitional movement. This is evidenced in the phasic activity of motor cortical neurons and the agonist–antagonist bursts of EMG throughout a reaching task. When the correlation between a pair of signals varies within the observation period, application of the standard correlation techniques will result in an average representation of the correlation between the signals during the observation due to the time integration in the analysis. For example, if there were a positive correlation between the signals at one point in the trial at a particular lag τ_1 and at some point later in the trial there were a negative correlation at the same lag τ_1, the two would tend to average out so that no structure would be evident in the correlogram. As illustrated in Fig. 4, a transient correlation will also be minimized by the standard correlation technique.

Methods that address this issue have been based on the JPST scatterplots that expressed coincidence between (delayed) spike occurrences (Gerstein and Perkel, 1972). An adaptation of this technique applied to the autocorrelation function for use on non-stationary neuronal data was developed by Ebner and Bloedel (1981). This 'generalized autocorrelation function' calculated the autocorrelation continuously in time for cerebellar Purkinje cells. The result of this analysis was plotted in 3 dimensions in such a way that absolute time was preserved. The autocorrelation could be viewed as it changed in time and had a characteristic structure demarcating a specific event, in this case climbing fiber discharge. The method we present in this paper is similar in that it also shows correlation trends over time.

Palm et al., 1988) binned the counts in the JPST to form a JPST histogram (JSPTH) which was also a 3D surface. Various methods to normalize the surface in order to remove the non-'stimulus-time-locked variations in near-coincident firing' were explored to isolate the 'genuine (rapid) modulations of activity and/or connectivity.' This method involved calculating the 'raw' PSTH that contained indirect and direct contributions to coincident firing and subtracting a 'predictor' of the indirect effect due to statistical trends in the individual spike trains. The 'residual' would then represent the direct, 'intrinsic neuronal dependencies' that one neuron could exert on another. The predictor was derived by multiplying the PSTs from each cell together. The

residual, when divided by the standard deviation of the predictor, yielded a number between 0 and 1 and was termed the 'normalized PSTH.' A significance measure was devised by comparing the joint standard deviation of the corresponding bins in the individual PSTHs to the experimental JSPTH.

The JSPTH relies on a comparison between two point processes. The bin widths were chosen so that each bin would contain, at most, one spike. This analysis was carried out on a short time scale (< 1 s). In contrast, we wished to compare a continuous EMG signal to a point process on a medium time scale (1–1000 s).

Our method of examining non-stationary data is to divide the input time series into smaller, epochs and carry out the analysis on each epoch (Priestly, 1965; Perkel et al., 1967b). Although the signals in each window may not be stationary, the division of data into windows and the intra-window comparison reveals temporal changes in correlation during the observation. The use of the impulse–response function is advantageous compared to the standard cross-correlation because it removes periodicities in the input signal from the calculation (compare Fig. 1H–I). If the correlation between signals remains constant through the trial, then using the impulse–response surface is not advantageous. This is the case for the simulations illustrated in Fig. 1 where the impulse–response function calculated in each window is constant. The standard impulse–response function averaged over time gives the same information as the impulse–response surface.

It is unlikely that this type of constant relation exists in real data on the time scale of the behavior examined here. We were interested in the EMG that occurred consistently within 100 ms of each cortical spike. The chosen bin width of 10 ms afforded the desired temporal resolution in the resulting correlogram. If higher resolution correlograms are required, for instance, to detect monosynaptic connections, then shorter bin widths would be required. Since a continuous signal is needed in this analysis, care must be taken in converting the spike train into instantaneous frequencies. As described above, bin widths that approach the mean interspike interval can be problematic. The intervals that span a bin boundary must be accurately attributed to each bin on a fractional basis with appropriate filtering (so as not to cause a large phase shift) applied to the intervals before conversion to spike frequency (Kuypers, 1981). EMG data should be filtered minimally for high-resolution analysis. Analog filters introduce delays and non-linearities to the data that complicate the interpretation of resulting analyses.

One issue not addressed in this study is significance testing. Usually, in correlation analyses, a confidence interval around the correlation coefficient is calculated. If this correlation interval encompasses a zero value (null hypothesis), the data is considered uncorrelated. The impulse–response surface could be tested in this fashion, using a bootstrap technique (Diaconis and Efron, 1983). One hundred surfaces could be calculated by randomly selecting (with replacement) n repetitions of n trials. The surfaces could then be rank-ordered according to their RMS values from the mean surface. The surface that was ranked 95 would be considered the 95th percent confidence interval. This surface plotted above and below the mean surface would give the confidence for each bin of the surface.

In dealing with a system as complex as that used to control volitional movement, correlation techniques can help identify general features of organization. In the absence of detailed anatomical knowledge, these techniques show whether one part of the system changes its activity with that of another portion. Correlation techniques in isolation cannot, however, prove causality. This may be of lesser concern as causal issues are becoming less important in neuroscience. It is clear through computational modeling and neurophysiological investigations of CNS systems that single events are unlikely to 'cause' particular effects within a distributed system. It is unusual for the functional linkage between single elements to be obligatory. As the technique described in this report shows, such functional linkage can be transient, occurring only during brief epochs within a movement. The inferences that can be drawn by relating these epochs of correlation to other time-related events may be useful in determining the cooperative interaction of different structures within the nervous system.

References

Aertsen, A.M.H.J. and Gerstein, G.L. (1985) Evaluation of neuronal circuitry: sensitivity of cross-correlation, Brain Res., 340: 341–354.

Box, G.E.P. and Jenkins, G.M. (1976) Time Series Analysis: Forecasting and Control, Holden-Day, Oakland, CA.

Chatfield, C. (1980) The Analysis of Time Series: an Introduction, 2nd edn., Chapman and Hall, London.

Cox, D.R. and Lewis, P.A.W. (1966) The Statistical Analysis of Series of Events, John Wiley, New York.

Diaconis, P. and Efron, B. (1983) Computer-intensive methods in statistics, Sci. Am., 248: 116–130.

Ebner, T.J. and Bloedel, J.R. (1981) Temporal patterning in the simple spike discharge of Purkinje cells and its relationship to climbing fiber activity, J. Neurophysiol., 45: 933–947.

Fetz, E.E., Cheney, P.D. and German, D.C. (1976) Corticomotoneuronal connections of precentral cells detected by post-spike averages of EMG activity in behaving monkeys, Brain Res., 114: 505–510.

French, A.S. and Holden, A.V. (1971) Alias-free sampling of neuronal spike trains, Kybernetics, 8: 165–171.

Gerstein, G.L. and Perkel, D.H. (1972) Mutual temporal relationships among neuronal spike trains: statistical techniques for display and analysis, Biophys. J., 12: 453–473.

Glaser, E.M. and Ruchkin, D.S. (1976) Principles of Neurobiological Signal Analysis, Academic Press, London.

Houk, J.C., Dessem, D.A., Miller, L.E. and Sybirska, E.H. (1987) Correlation and spectral analysis of relations between single unit discharge and muscle activities, J. Neurosci. Methods, 21: 201–224.

Humphrey, D.R. (1972) Rating motor cortex spike trains to measures of motor performance, Brain Res., 40: 7–18.

Kirkwood, P.A. (1979) On the use and interpretation of cross-correlation measurements in the mammalian central nervous system, J. Neurosci. Methods, 1: 107–132.

Knox, C.K. (1974) Cross-correlation functions for a neuronal model, Biophys. J., 14: 567–582.

Kuypers, H.G.J.M. (1981) Anatomy of the descending pathways. In: Handbook of Physiology. The Nervous System. II. American Physiological Society, Bethesda, MD, pp. 597–666.

Landgren, S., Phillips, C.G. and Porter, R. (1962) Cortical fields of origin of the monosynaptic pyramidal pathways to some alpha motoneurones of the baboon's hand and forearm, J. Physiol. (Lond.), 161: 112–125.

Mendell, L.M. and Henneman, E. (1971) Terminals of single Ia fibers: location, density and distribution within a pool of 300 homonymous motoneurons, J. Neurophysiol., 34: 171–187.

Miller, L.E., Sinkjaer, T., Andersen, T., Laporte, D.J. and Houk, J.C. (1992) Correlation analysis of relations between red nucleus discharge and limb muscle activity during reaching movements in space. In: R. Caminiti, P.B. Johnson and Y. Burnod (Eds.), Exp. Brain Res., Series 22, Control of Arm Movement in Space: Neurophysiological and Computational Approaches, Springer, Heidelberg, pp. 263–283.

Moore, G.P., Perkel, D.H. and Segundo, J.P. (1966) Statistical analysis and functional interpretation of neuronal spike train data, Annu. Rev. Neurosci., 28: 493–522.

Palm, G., Aertsen, A.M.H.J. and Gerstein, G.L. (1988) On the significance of correlations among neuronal spike trains, Biol. Cybern., 59: 1–11.

Perkel, D.H., Gerstein, G.L. and Moore, G.P. (1967a) Neuronal spike trains and stochastic point processes. I. the single spike train, Biophys. J., 7: 391–418.

Perkel, D.H., Gerstein, G.L. and Moore, G.P. (1967b) Neuronal spike trains and stochastic point processes. II. Simultaneous spike trains, Biophys. J., 7: 419–440.

Priestly, M.B. (1965) Evolutionary spectra and non-stationary processes, J. Roy. Statist. Soc. B, 27: 204–237.

Ramirez, R.W. (1985) The FFT: Fundamentals and Concepts, Prentice-Hall, Englewood Cliffs, NJ.

Richmond, B.J., Optican, L.M., Podell, M. and Spitzer, H. (1987) Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. I. Response characteristics, J. Neurophysiol., 57: 132–146.

Schwartz, A.B. (1992) Motor cortical activity during drawing movements. Single-unit activity during sinusoid tracing, J. Neurophysiol., 68: 528–541.

Soechting, J.F., Burton, J.E. and Onoda, N. (1978) Relationships between sensory input, motor output and unit activity in interpositus and red nuclei during intentional movement, Brain Res., 152: 65–79.

Watt, D.G.D., Stauffer, E.K., Taylor, A., Reinking, R.M. and Stuart, D.G. (1976) Analysis of muscle receptor connections by spike-triggered averaging. I. Spindle primary and tendon organ afferents, J. Neurophysiol., 39: 1375–1392.