cells we analyzed here. The mean improvement was 7.5%, and in three cells the thresholds dropped by roughly a factor of 2. In other words, the "Caltech" ideal observer who distinguishes spikes from bursts can—for these cells—determine the correct direction of motion (using a statistical criterion) at a lower level of coherence c than the "Stanford" observer. $\alpha = 0.5$ and 2 gave smaller improvements (as did $\alpha = 0.75$ and 1.5; not shown). As α is made larger ($\alpha \rightarrow \infty$), bursts are more and more emphasized at the expense of single isolated spikes and the thresholds increase by 53%. The threshold also increases when bursts are weighted by the square of the number of spikes in each burst ($\beta = 2$). Weighting events by the square root of the number of spikes per event ($\beta = 0.5$), on the other hand, decreases thresholds (Fig. 9), since it decreases the variability contributed to the final signal.

From the point of view of our fictitious pair of ideal observers, the Caltech observer does better than his Stanford counterpart by replacing each occurrence of a burst of spikes by a single spike. In this sense it can be said that a crude measure of the temporal organization of spike trains does better in terms of signal detection theory than a simple spike count. This is not to say, however, that more sophisticated measures of temporal organization, possibly taking account of the simultaneous activity of many neurons, cannot do better yet (e.g., Aertsen et al., 1989; Richmond and Optican, 1992; Singer, 1994).

We do not know at this point the code that neurons postsynaptic to MT cells use to decide whether the stimuli move in one or the other direction. The fact that the neuronal threshold of many cells is frequently lower than the psychophysical threshold of the entire animal (Newsome et al., 1989a) requires an explanation as to why the animal does not do better than it does (invoking correlated activity among cells and population coding; Britten et al., 1992). Our results point to an additional explanation: if bursts are substantially more efficient in elevating postsynaptic firing rate than isolated spikes, corresponding to $\alpha \gg 1$, thresholds would increase and averaging over many cells would be required in order to mimic the psychophysical thresholds. It is important at some point that this question be resolved experimentally, possibly using a combination of in vivo slice techniques with behavioral studies.

Functional considerations

What is the function of bursts? Why should cortex have two types of long-range projection cells, one signaling isolated spikes and the other responding frequently with bursts of spikes? It has been argued (Koch and Crick, 1994) that bursting neurons are much more efficient at accumulating calcium in their axonal terminals than cells that fire isolated spikes (i.e., four spikes within a 10 msec interval cause a much larger increase in intracellular calcium at the end of the last spike than four spikes within a 40 msec interval). Because intracellular calcium accumulation in the presynaptic terminal is thought to be mainly responsible for various forms of short-term potentiation (in particular, facilitation and augmentation; Magleby, 1987), it may well be that the primary function of bursting neurons is to induce this non-Hebbian (i.e., nonassociative) type of synaptic plasticity at its postsynaptic targets outside of the cortical system. In essence, the burst of spikes could turn on short-term memory, which would then decay over several seconds (see also Crick, 1984). One might then expect there to be a relationship between bursting and short-term learning.

It is important to know whether our "bursting" cells corre-

spond to the "intrinsically bursting" cells identified by intracellular current injection. The latter cells appear to be confined (at least in rat and guinea pig slice) to layer V (Agmon and Connors, 1992). In rat area 17, these cells have been shown to project outside cortex, in this case to the ipsilateral superior colliculus, while the remaining pyramidal cells in layer V project to the contralateral striate cortex (Kasper et al., 1991). Recent in vivo recordings in awake cat motor cortex have revealed that cells at or below a depth of 800 μm (corresponding to layer 5) show strong bursting activity (Baranyi et al., 1993). It is not known to what extent such cells exist or are localized to particular layers in primate cortex.

Correlation to behavior

Finally, we return to a question that provided primary motivation for this study. Is the animal's perception of the stimulus, as evidenced by performance, influenced by temporal structure in the spike trains, particularly with respect to the peak in the power spectrum near 40 Hz? We correlated the presence and strength of the peak in the spectrum to both the stimulus and the behavior of the monkey. As seen in Figures 5 and 6, we found no significant correlation between the fraction of dots moving in one or the other direction and P, the measure of the peak in the power spectrum, for most cells. We repeated this measurement using the integral of power in the 40 Hz band of $S'(f)$ with similar results. Furthermore, if the monkey is not forced to respond to the stimulus or even in the absence of the motion stimulus, the basic propensity of a cell to show this peak remains. This is related, of course, to our earlier result that bursting in these cells does not depend on stimulus conditions.

We find a similar lack of correlation between P and the various measures of behavior used in a previous comparison of neuronal responses and psychophysical performance (Newsome et al., 1989a; Britten et al., 1992). For instance, we tested for significant changes in the distribution of P when the monkey made correct versus incorrect decisions at near-threshold coherence levels and when the monkey made preferred versus null guesses for zero coherence motion, but we found no significant (paired t test, $p > 0.05$) correlations.

Figure 7 shows another attempt at studying the relationship between the peak in the spectrum and the behavior of the monkey. As is evident, no correlation exists between P and the level of the neuronal threshold, c_{cell}, that is, the fraction of dots moving in the cells preferred direction at which the cell can "decide" the correct direction of motion (using an ROC criterion; Newsome et al., 1989a; Britten et al., 1992). A similar lack of correlation exists between P and the ratio of single-cell thresholds to the threshold of the animal (c_{cell}/c_{system}) and between P and the decision related probability of each neuron (not shown). Thus, for our stimulus conditions, the presence or strength of a peak in the power spectrum of well-isolated units does not tell us anything about the behavior of these animals.

When analyzing our data set, it should be kept in mind that the three monkeys from which the MT cells were recorded were extensively trained using operant conditioning techniques. In order to perform correctly the motion discrimination task at threshold levels, up to 6 months of training was required (Britten et al., 1992). We analyzed in a preliminary manner data from MT cells from one naive monkey who was only trained to fixate (E. Zohary and W. Bair, personal communication), and we found no significant difference in the distribution of burst and nonburst cells or in the shape of the associated power spectra. However,