2000

Information Encoding by Individual Neurons and Groups of Neurons in the Primary Visual Cortex

Daniel Salo Reich

Information Encoding by Individual Neurons and Groups of Neurons in the Primary Visual Cortex

A thesis presented to the faculty of
The Rockefeller University
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy

by

Daniel Salo Reich

The Rockefeller University
New York City
October 31, 2000

UMI Number: 3014743

Bell & Howell Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

To Amy

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ANOVA	ANalysis Of VAriance
CRT	Cathode Ray Tube
IM	IntraMuscular
ISI	InterSpike Interval
ISIH	InterSpike Interval Histogram
IV	IntraVenous
LGN	Lateral Geniculate Nucleus
LL	Labeled Line
MT	Medial Temporal
NLIF	Noisy, Leaky Integrate-and-Fire
PSTH	Post- (*or* Peri-) Stimulus Time Histogram
RF	Receptive Field
RI	Redundancy Index
SMRP	Simply Modulated Renewal Process
SP	Summed Population
SS	Separate Sum
V1	primary visual cortex

Acknowledgments

There are many who deserve my gratitude.

Bruce Knight, my advisor at The Rockefeller University, has the rare gift by which a scientist can penetrate to the heart of a problem, cutting away the layers of dogma and wishful thinking that so often cloud scientific thinking. For the moments when such insight has benefited me—and there have been many over the course of the eight years I have known Bruce—I am eternally grateful.

Jonathan Victor, my advisor at Weill Medical College of Cornell University, is a perfect mentor. He is not only brilliant but overwhelmingly generous with his time—more generous than I could ever have wished for. He has provided endless guidance that reflects on all aspects of the work, from the design of experiments to the development of novel techniques for data analysis. He has taught me what it means to approach a problem scientifically, and how to evaluate the relative merits of taking one approach or another to solving it.

Several other scientists have also proved invaluable. Ferenc Mechler and I have shared an office and carried out experiments together for four years. Ferenc has taught me how to be a careful scientist and an incisive and critical thinker. He has been an ideal colleague and collaborator, and a good friend. Udi Kaplan introduced me to visual neurophysiology, instilled in me a love for scientific detail, showed me how to perform the experiments that are the heart of this thesis, and taught me how to learn from and appreciate the work of others. Keith Purpura has been a much appreciated source of scientific ideas, knowledge, and perspective.

For their support and assistance, I would like to acknowledge the members of my thesis committee, Albert Libchaber, Judith Hirsch, and Rob de Ruyter van Steveninck. I am indebted to colleagues at The Rockefeller University—Jose-Manuel Alonso, Lisa

Croner, David Heath, Pratik Mukherjee, Estela O'Brien Ogiste, Darren Orbach, Yoshi Ozaki, Clay Reid, Dan Samber, and Mavi Sanchez-Vives—and at Weill Medical College—Annemarie Canel, Mary Conte, Bob DeBellis, Steve Kalik, Michael Repucci, Niko Schiff, and Jeff Tsai—for help at many stages of the work. Rob de Ruyter van Steveninck and Markus Meister have given me useful comments and criticism on portions of the manuscript. Michale Fee made helpful suggestions on adapting his spike-sorting algorithm for use with my data. John Rinzel has encouraged me since I first became interested in systems neuroscience, and it is because of a comment he made that I first became aware of the Rockefeller group. David Reich, my brother, and Jason Eisner, my brother-in-law, have doubled as first-rate scientific advisors.

I would like to thank a number of institutions, including the Tri-Institutional M.D.-Ph.D. program; the Office of Graduate Studies at Rockefeller (especially Marjorie Russel); and the Marine Biological Laboratory at Woods Hole. The thesis work was supported by NIH grants GM07739 (to the M.D.-Ph.D. program), EY07138 (to the Tri-Institutional Vision Training Program), and EY9314 (to Jonathan Victor).

I am deeply indebted to some friends and colleagues, among them Ben Aldrich-Moodie, Jeffrey Degen, Nicole Kahn, David Kaufman, Johannes Schlöndorff, and especially Liat Kaplan. The extended Eisner family has encouraged and nourished me, and I am grateful to them.

My most profound thanks go to my family—to my sister Rebecca, my brother David, my parents Tova and Walter, and my dear wife Amy. They deserve much of the credit for this work, and I happily grant it to them. Their love and support are my most precious possessions.

Abstract

How is information about visual stimuli encoded into the responses of neurons in the cerebral cortex? In this thesis, I describe the analysis of data recorded simultaneously from groups of up to eight nearby neurons in the primary visual cortices of anesthetized macaque monkeys. The goal is to examine the degree to which visual information is encoded into the times of action potentials in those responses (as opposed to the overall rate), and also into the identity of the neuron that fires each action potential (as opposed to the average activity across a group of nearby neurons). The data are examined with techniques modified from systems analysis, statistics, and information theory. The results are compared with expectations from simple statistical models of action-potential firing and from models that are more physiologically realistic. The major findings are: (1) that cortical responses are not renewal processes with time-varying firing rates, which means that information can indeed be encoded in the detailed timing of action potentials; (2) that these neurons encode the contrast of visual stimuli primarily into the time difference between stimulus and response onset, which is known as the *latency*; (3) that this so-called *temporal coding* serves as a mechanism by which the brain might discriminate among stimuli that evoke similar firing rates; (4) that action potentials preceded by interspike intervals of different durations can encode different features of a stimulus; (5) that the rate of overall information transmission can depend on the type of stimulus in a manner that differs from one neuron to the next; (6) that the rate at which information is transmitted specifically about stimulus contrast depends little on stimulus type; (7) that a substantial fraction of the information rate can be *confounded* among multiple stimulus attributes; and, most importantly, (8) that averaging together the responses of multiple nearby neurons leads to a significant loss of information that increases as more neurons are considered. These results should serve as a basis for direct investigation into the cellular mechanisms by which the brain extracts and processes the information carried in neuronal responses.

Chapter 1

Introduction

For the supreme cunning of the structure of the gray matter is so intricate that it defies and will continue to defy for many centuries the obstinate curiosity of investigators. That apparent disorder of the cerebral jungle, so different from the regularity and symmetry of the spinal cord and of the cerebellum, conceals a profound organization of the utmost subtility which is at present inaccessible.

Santiago Ramón y Cajal (1852-1934)
(Ramón y Cajal, 1989)

At the close of a century that has witnessed remarkable and almost unimaginable advances in biology—the elucidation of the genetic code and the development of experimental tools that will make it possible to unravel the most fundamental aspects of cellular function—it is odd that Cajal could practically have made the same statement today. More than one hundred years after the first anatomical and electrophysiological studies of visual cortex (Gross, 1998), its basic mechanisms of operation are still very poorly understood.

Not all sensory systems are so recalcitrant. Certain invertebrate visual systems, such as the lateral eye of the horseshoe crab, *Limulus polyphemus* (Hartline and Ratliff, 1958; Knight, 1972b), and the motion-sensitive H1 neuron of the blowfly *Calliphora vicina* (Rieke et al., 1997; Ruyter van Steveninck et al., 2000), are relatively well characterized. By contrast, models that successfully predict the responses of populations of neurons in mammalian visual cortex are only beginning to be built (Somers et al., 1995; Knight, 2000; McLaughlin et al., 2000). Constructing such models may be prohibitively difficult: as Cajal recognized, cortical architecture is overwhelmingly complex. We now know that each hemisphere of the primary visual cortex (V1) of the macaque monkey contains approximately 10^8 neurons, and each of those neurons makes approximately 10^3 synaptic contacts with other neurons (Peters, 1987). The anatomic (Tarczy-Hornoch et

al., 1999; Gupta et al., 2000) and functional (Reyes et al., 1998; Thomson, 2000) diversity of these synapses defies simple classification. Nonetheless, though the "century" is rapidly turning into Cajal's "centuries," the central project of cortical physiologists—which would prove him wrong—is still alive and well.

This thesis describes a systematic attempt to unravel some of the rules that dictate the ways in which V1 neurons of the macaque monkey, *Macaca fascicularis*, encode information about visual stimuli. The overriding question at the root of all the work described here is: which aspects of the responses of V1 neurons convey information about which aspects of the visual signal? This question is addressed with techniques of classical neurophysiology: stimuli are presented to the animal's eyes while the activity of individual neurons in the cortex is monitored and recorded. The results, though, depend on three refinements of and advances upon this approach: (1) the use of a wide variety of stimuli to probe a range of responses; (2) the development of novel analysis methods to characterize the qualitative and quantitative aspects of information transmission; and (3) the introduction, for the first time in our laboratory, of techniques to record the simultaneous activity of several well-isolated, nearby visual neurons. Taken together, these advances lead to discoveries that are exciting and remarkable.

The thesis consists of this general introduction, five chapters devoted to the presentation and analysis of responses recorded from V1 neurons, and a brief conclusion. Each of the central chapters has its own self-contained introduction and discussion sections and can therefore be read as a stand-alone research report. Indeed, much of the text for Chapters 3 and 5 has already been published in the form of research papers (Reich et al., 1998; Reich et al., 2000). Methods sections sometimes reference previous chapters. The brief concluding chapter of the thesis is a general discussion of problems solved and raised, and of possible directions for future work.

Chapter 2 provides a general description of the experimental methods used throughout the thesis, with a particular emphasis on the algorithm that is used to assign extracellular waveforms to individual neurons (a process known colloquially as *spike sorting*). The use of tetrodes—four insulated platinum/tungsten wires in close proximity

to one another—to record the simultaneous activity of up to eight nearby neurons is described. The introduction of the tetrode recording system into our laboratory is one of the major technical accomplishments of this doctoral work.

Chapter 3 begins the analysis of real neuronal data, focusing on the responses of retinal ganglion cells and lateral geniculate nucleus neurons of cats, but also including responses of neurons in monkey V1. The goal of the work described in Chapter 3 is to use the time-varying statistics of neuronal responses to characterize, or at least to constrain, the underlying spike-generating processes that are at work in the mammalian brain. In particular, the possibility is addressed, and ultimately rejected, that neuronal responses can be adequately described as being generated by simply modulated, time-varying renewal processes. (The terms are defined in the text of the chapter.) The implication is that the spike-generating processes are sophisticated enough to support the encoding of stimulus-related information into the precise times of occurrence of action potentials ("spikes"), rather than simply into the rates at which those spikes are fired. The chapter closes with the description of a method to compute and characterize "phase locking" in neuronal responses to periodic stimuli, such as drifting sinusoidal gratings. This method, based on *circle maps*, makes it possible—sometimes—to recover underlying preferred spike times from noisy responses. The results in Chapter 3 have been published (Reich et al., 1998).

The analysis in Chapter 4 builds on this foundation by examining the ways in which a particular stimulus attribute—contrast—is encoded into precise spike times in the responses of V1 neurons. The stimuli are flashed sinusoidal gratings, and the responses are analyzed with an information-theoretic method that is based on the embedding of neuronal responses into a metric space (Victor and Purpura, 1996). The results indicate that contrast is encoded into both the overall firing rate and the response latency. More broadly, the results point to a specific role for temporal coding in visual cortex: to facilitate fine discrimination among stimuli that evoke similar firing rates. This hypothesis is ultimately testable, though not with the experimental techniques used in this thesis.

Chapter 5 examines temporal coding in responses to a different set of stimuli—rapidly changing checkerboard patterns. A major advantage of these stimuli is that they make it possible to characterize a neuron's spatiotemporal receptive-field dynamics (Reid et al., 1997). In this chapter, the analysis of these dynamics is driven by the discovery that interspike intervals fired by V1 neurons fall into one of three distinct ranges, the demarcations of which are stereotyped across neurons, and which are likely to reflect intrinsic aspects of neuronal spike generation. It is found that spikes preceded by interspike intervals in different ranges encode different spatiotemporal features of the stimulus. Taken together, the results described in Chapters 4 and 5 reveal that temporal coding can play an important role in responses to both rapidly and slowly changing visual stimuli. The results in Chapter 5 have been published (Reich et al., 2000).

Chapters 6 and 7 present the results of applying a general information-theoretic method to the responses of V1 neurons to a variety of visual stimuli, both rapidly and slowly changing. When sufficient data are available, this so-called *direct method* (Strong et al., 1998) allows the full characterization of the stimulus-related information in spike-train responses. Since obtaining enough data is virtually impossible, methods of estimating full information rates that have less stringent data requirements are introduced. These methods make certain compromises on the accuracy of the calculated information rates, but the compromises are justifiable on physiologic grounds, and their consequences are discussed. An additional novel feature of the analysis described in Chapters 6 and 7 is an extension of the direct method to make possible the parsing of information rates into components that relate to specific stimulus attributes.

In Chapter 6, the analysis is applied to the responses of individual neurons to stimuli that vary in contrast. The results indicate that rapidly varying stimuli do not necessarily evoke higher information rates, as has been suggested (Bura as and Albright, 1999), and that information about contrast is largely independent of stimulus type. The data also reveal that a substantial portion of the overall information cannot be assigned to individual stimulus attributes, but is rather *confounded* between two or more of those attributes. This suggests that the dimensions along which the visual system analyzes stimuli are not the same dimensions along which experimenters present them. Rather, different

stimulus attributes are likely combined into unified representations that make it possible for the visual system to answer questions that require knowledge of their joint values.

Finally, Chapter 7 analyzes the responses of multiple nearby neurons. In many ways, the results in Chapter 7 are the culmination of all the work in earlier chapters. The primary question addressed is whether nearby neurons convey independent or redundant stimulus-related information, or whether they act in concert to encode information not available in the responses of any one neuron alone. The answers to this question have important implications for understanding the functional architecture of sensory cortex. It has long been thought that because nearby neurons, organized anatomically into cortical *columns*, are tuned to stimuli in a similar fashion (Mountcastle, 1957; Hubel and Wiesel, 1962), downstream neurons can improve the signal-to-noise ratio of incoming signals by averaging together the responses of multiple inputs (Shadlen and Newsome, 1998). But the results presented here suggest that not only do the responses of nearby neurons convey for the most part independent information, but that an averaging scheme sacrifices a substantial fraction of the information that would be available if the neuron that fires each action potential were also considered. The demonstration that reliable information can be conveyed without averaging makes it easier to understand how the brain can take advantage of its intrinsic complexity to represent and process a seemingly limitless array of external signals.

Chapter 2

General Methods

This chapter describes the experimental methods used to record activity from neurons in the primary visual cortices of anesthetized macaque monkeys, *Macaca fascicularis*. For details of recordings from neurons in the retina and lateral geniculate nucleus, see Chapter 3. All experiments are performed in collaboration with Jonathan Victor, my co-advisor, and Ferenc Mechler, a postdoctoral fellow in the lab. Collaborative work is necessary because the experiments require round-the-clock attention for a period of typically four or five days. Most of the data analysis, including the spike sorting described later in the chapter, is performed off-line in the weeks following each experiment, although initial characterization of neurons' receptive-field properties (and subsequent determination of stimulus parameters) is done during the course of each recording. All of the experimental procedures comply with the National Institute of Health's guidelines (Van Sluyters and Oberdorfer, 1994) and with institutional requirements at Weill Medical College (where the experiments are conducted).

Vegetative physiology. Most of the data in this thesis come from 14 male adult monkeys weighing roughly 3 kg each. Each animal is premedicated in the Research Animal Resource Center at Weill Medical College with atropine 40 μg IM, and anesthesia is induced 30 minutes later with ketamine 10 mg/kg. Once the animal is sedated, it is transported to the experimental facility. Surgical sites (inner thighs, sternum, scalp, and neck) are shaved with an electric razor and the animal is placed on a temperature-controlled surgical table. Sites are cleaned with betadine and infiltrated with procaine HCl 2% or lidocaine HCl 2%. Access to the two femoral veins and one of the femoral arteries is obtained. Anesthesia is supplemented with methohexital 0.5-1 mg/kg boluses IV, as necessary. A tracheostomy is performed so that the animal can be artificially ventilated after paralysis. Pupils are dilated with atropine 1% drops, ocular inflammation is forestalled with flubriprofen 0.03% drops, and dessication is prevented by eyelid closure. Bacitracin-

neomycin-polymyxin with hydrocortisone 1% ointment is applied to the eyelids to prevent infection. A rectal thermometer is inserted, and a Uro-San pediatric male external catheter (Mentor Urology, Santa Barbara, CA) is applied to facilitate urine collection. Dexamethasone 4 mg is given IV or IM (and is repeated each day) to reduce the potential for cerebral edema. Penicillin G (300,000 units) is given IM as prohyplaxis against bacterial infection. Respiration and heart rate are checked regularly; if any problem is detected, artificial ventilation and appropriate Cardiac Life Support treatments are initiated.

After the preparatory surgery, the animal is transferred to the experimental table and placed on a thermostatically controlled heating pad to keep the core temperature as close to 37°C as possible. Electrocardiogram leads are connected and the arterial catheter is attached to a blood-pressure monitoring device (Patient Monitor 78354A, Hewlett Packard, Palo Alto, CA). Opiate analgesia (sufentanil) is initiated with a 3 μg/kg bolus and maintained at a rate of 1-6 μg/kg/hr; the final rate is determined iteratively and is chosen to keep the heart rate and blood pressure relatively constant (roughly 120 beats per minute and 120/80, respectively). Fluids (5% dextrose in lactated ringers) are delivered through one of the IV lines at the rate of 6-10 mL/hr. Throughout the experiment, blood pressure, heart rate, expired carbon dioxide, and core temperature are continuously monitored. Signs of infection, such as fever or pulmonary congestion, are treated with gentamicin 5 mg/kg/day.

The animal's head is secured in a stereotactic apparatus with ear bars. The scalp is reflected or removed and a craniotomy is performed at a location 12 mm posterior and 14-15 mm lateral to the center of the ear bars. The craniotomy exposes the dura mater and its vessels and often reveals the location of the lunate sulcus. The dura is reflected at least 2 mm posterior to the lunate sulcus to ensure that recordings are made in the striate cortex (Hubel and Wiesel, 1977). Recording sites are confirmed histologically at the end of the experiment (see below).

Paralysis (to prevent eye movements) is induced and maintained with pancuronium 1 mg bolus, 0.2-0.4 mg/kg/hr. The animal is placed on a respirator and ventilated 20-25 times per minute. Stroke volume and exact respiratory rate are chosen to

maintain the pressure of expired CO_2 in the range of 30 mm Hg. Gas-permeable hard contact lenses are used to prevent corneal drying, and artificial pupils (3 mm diameter) are placed in front of the eyes. The optical quality of the animals' eyes is checked repeatedly by direct ophthalmoscopy. A half hour after paralysis is induced, the positions of the foveae and optic discs are mapped onto a tangent screen positioned 114 cm in front from the animal's eyes (at a distance of 114 cm, 1 cm corresponds to 0.5° of visual angle). Optical correction at this viewing distance is determined by slit retinoscopy, and corrective lenses are applied. At a later stage, refraction is corrected to optimize the responses of single neurons to drifting sinusoidal experiments; typically, retinoscopy provides a good estimate of the final refraction.

Electrophysiology. The preparatory stages of the experiments usually last about six hours, after which recording can commence. A relatively vessel-free region of the cortical surface is identified, and the recording electrode is lowered until it reaches the cortical surface. The vertical position of the electrode is controlled with micron precision. A change in the properties of the electrical signal transmitted by the electrode signifies that the surface has been reached. The electrode is then systematically lowered until neural activity that can be modulated by a visual stimulus is encountered. The electrode is then advanced more slowly until one or several distinct waveform types are discerned in the recording. This indicates that the activity of individual neurons can be isolated.

In the first few animals, tungsten-in-glass electrodes (Merrill and Ainsworth, 1972) were used. Electrical signals were amplified by custom-built equipment (including a battery-powered, high input-impedance, unity-gain preamplifier) and band-pass filtered between 300 and 6000 Hz (Model 3988, Krohn-Hite, Avon, MA). The remainder of the recordings were made with tetrodes (Thomas Recording, Giessen, Germany), which consist of four insulated platinum/tungsten wires spaced from 18 to 40 μ apart. The individual leads of the tetrodes are close enough to one another so that potentials originating from a single neuron often appear on more than one lead with different amplitudes (Gray et al., 1995). This greatly simplifies the process of spike sorting, as discussed below. Tetrode signals are fed through a unity-gain preamplifier (NB Labs, Denison, TX or Neura-

lynx, Tucson, AZ) and an eight-channel software-controlled amplifier and filter bank (600 to 6000 Hz, Neuralynx, Tucson, AZ).

For both single electrodes and tetrodes, amplified and filtered signals are fed into the Discovery system (Datawave, Longmont, CO). Thresholds are set independently for each channel of the tetrode, and 32 analog-to-digital samples (sampling frequency: 25 to 30 kHz) are saved to disk around each threshold crossing (usually eight samples before the threshold crossing and 24 afterward). For tetrode recording, 32 samples are saved from each of the four channels when the threshold is crossed on any one channel. The time of each threshold crossing is recorded to the nearest 0.1 msec. In general, thresholds are set relatively low so as to limit the spurious deletion of spikes at this stage. Thus, many of the *spikes* that are fed into the spike sorting algorithm (see below) are actually samples of noise.

Stimuli and experimental protocol. Visual stimuli are created on a green CRT (Model 608, Tektronix, Wilsonville, OR; 270 frames/sec, 150 cd/m^2 mean luminance) by specialized equipment (Milkman et al., 1980). The CRT is positioned 114 cm from the animal's eyes, and a first-surface mirror is used to direct the center of the receptive field(s) onto the center of the CRT. For each neuron or group of neurons encountered, the overall ocular dominance and orientation preference, which are expected to vary little from one neuron to the next at a fixed site, are first determined (Hubel and Wiesel, 1962).

Next, a series of quantitative tests are performed to determine stimulus-response properties and to classify the neurons as either *simple* or *complex* (Hubel and Wiesel, 1962; Skottun et al., 1991). Simple and complex cells are distinguished by their responses to drifting sinusoidal gratings: simple cells respond primarily by modulating their rate of discharge in time with the stimulus, whereas complex cells respond primarily by elevating their overall rate of discharge over the background level, with very little modulation (Skottun et al., 1991).

Quantitative orientation tuning is determined with drifting sinusoidal gratings presented at 16 different orientations, in random order. The spatial and temporal frequencies of the gratings are chosen to produce the greatest audible change in the electrical dis-

charge as judged subjectively by the experimenter. The Michelson contrast of the gratings (difference between maximum and minimum luminance divided by twice the mean luminance) is set to 1. Responses to four-second presentations of each orientation, and also to the presentation of a uniform field at the same mean luminance (*blank*), are collected five to eight times in block-randomized fashion. For the analysis described in this paper, the responses to each stimulus are considered to be one continuous steady-state record. The orientation that evokes the largest response (measured as the elevation in average firing rate for complex cells, or the amplitude of the fundamental Fourier component for simple cells) is used in the remainder of the tuning experiments. When groups of well-isolated single neurons are recorded, the orientation preference sometimes varies within the group by 45° or more. In these cases, subsequent experiments are performed at multiple orientations or else at a compromise orientation. More commonly, neurons within a cluster prefer gratings at the same orientation that move in opposite directions; in these cases, subsequent experiments are performed with both directions of motion.

Spatial and temporal frequency tuning curves are measured in much the same way. Again, if more than one spatial or temporal frequency preference is found within a cluster of neurons, either a compromise value is chosen or else subsequent experiments are performed at multiple values. The final tuning experiment measures the contrast response function. Here, the responses to drifting gratings of the optimal orientation, spatial frequency, and temporal frequency are measured at contrasts geometrically spaced from 0.0625 to 1, in addition to a blank (contrast 0). The data from these experiments are analyzed in Chapters 3, 6, and 7.

Following the tuning experiments, the neurons' receptive fields are mapped using flickering checkerboard patterns in which the temporal pattern in each check is varied according to a binary m-sequence (Sutter, 1992; Victor, 1992; Reid et al., 1997). Details of the stimuli and mapping procedures are given in Chapter 5. M-sequence responses are used for the following purposes: (1) to obtain a detailed picture of the spatiotemporal dynamics of neurons' receptive fields (as in Chapter 5); (2) to determine functional connectivity between simultaneously recorded neurons by cross-correlation analysis (not reported in this thesis); (3) to characterize the statistical properties of neuronal discharges

(for example, interspike interval histograms, as in Chapter 5); and (4) to provide a substrate for calculating the rates of information transmitted by cortical neurons (as in Chapters 5, 6, and 7).

After the m-sequence experiments, additional experiments are performed, the identities of which depend on the neurons' response properties and the needs of the particular experimenter. Among these experiments are stationary sinusoidal grating presentations, analyzed in Chapters 4, 6, and 7. Other types of stimulus, including drifting compound gratings and m-sequences with modulated "surround" stimulation, are also used, but the results of those studies are not included in the thesis.

Histology. Electrode tracks are marked either by electrolytic lesions (5 μA for 5 sec, spaced 1000 μ apart) or fluorescent dyes (DiCarlo et al., 1996). At the end of several days of recording, animals are given an overdose of methohexital and perfused transcardially with saline 0.9% and phosphate-buffered paraformaldehyde 4%. The brains are removed and prepared for standard histological examination. Whenever possible, the locations of the electrode tracks are recovered and a laminar position in the cortex is assigned to each recording site.

Spike sorting. Although the membrane-potential changes characteristic of neuronal action potentials are highly stereotyped when recorded either intracellularly or extracellularly, the problem of separating extracellularly recorded waveforms into cell-specific components remains unsolved (Lewicki, 1998). The obvious problem is that the only way of determining whether a classification is accurate is to compare the spike times obtained from the classification to those obtained from a direct intracellular recording, in which there is no ambiguity. Recently, a number of groups have attempted this experiment—one studying the locust olfactory system (Wehr et al., 1999), and the other studying the rat hippocampus (Harris et al., 2000).

In the neocortex, extracellularly recorded spike waveforms are often close to the background noise, which makes accurate classification even more difficult to achieve (Fig. 2.1). Moreover, even when waveforms have relatively large signal-to-noise ratios, it is quite common for spikes from two neurons to occur nearly simultaneously, so that their

waveforms overlap. Such simultaneous or coincident firing is common in the cortex even for neurons that are separated by hundreds of microns or more (Reid and Alonso, 1996; Singer et al., 1997; Lampl et al., 1999), but it may be especially common for nearby neurons that tend to respond to similar features in a stimulus (Gray et al., 1995). The problem of resolving such *overlaps* has been addressed by a number of investigators (Lewicki, 1994; Sahani et al., 1998; Rebrik et al., 1999). In this thesis, simultaneous spikes (that is, spikes that occur within 1 msec of one another) are completely ignored, except in rare cases where classification is especially clear. This obviously means that fundamental aspects of the responses of populations of neurons may be overlooked, since, at least in some cases, synchronous spikes can carry important information that is not available if such spikes are not distinguished (Dan et al., 1998).

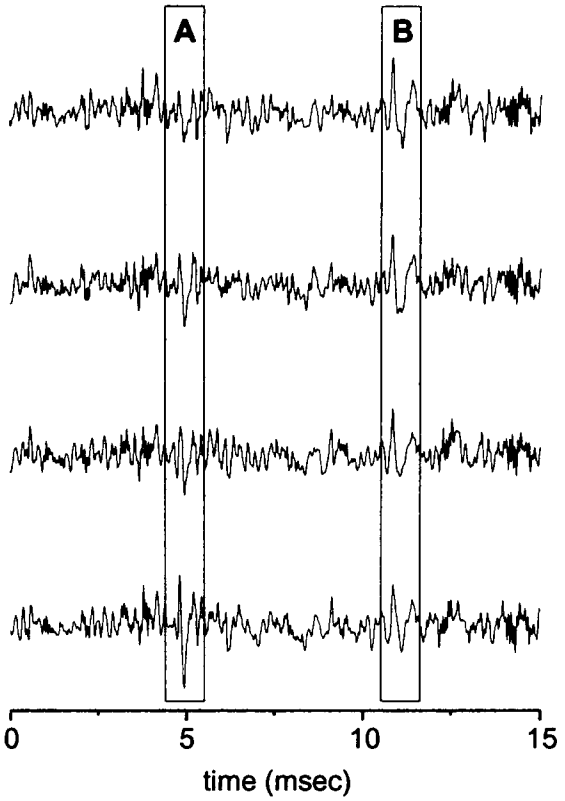

Figure 2.1. *Example of a tetrode recording in V1.* Waveforms are sampled at 30.303 kHz. Two different spikes (marked *A* and *B*) are clearly evident. Spike *A* is most prominent on the bottom channel, whereas spike *B* is most prominent on the top channel. Note that for both spikes, the peak-to-peak amplitude of the waveform is not much larger than the surrounding noise, and several other spikes are clearly buried within the noise.

The present approach to sorting spikes takes its inspiration from earlier work (Fee et al., 1996a; Fee et al., 1996b), albeit with significant modification. The primary differences are: (1) generalization of the analysis from stereotrode (two-channel) to tetrode (four-channel) or electrode (single-channel) recording; and (2) the use of waveform principal components as a basis set for classification (Abeles and Goldstein, 1977). As implemented, the spike-sorting method is only semi-automatic. The algorithm requires specification of a number of parameters (a stable set of which was found in some of the early experiments and used for virtually all subsequent spike-sorting), and final classification requires substantial user intervention. Although the requirement for user intervention may be seen as a drawback of the method, it also reflects the current state of

tion may be seen as a drawback of the method, it also reflects the current state of the art in spike sorting.

It is important to point out that the findings described in this thesis, especially the information calculations across simultaneously recorded neurons, are in general conservative with respect to errors in spike sorting. In particular, the amount of information estimated to be encoded in patterns of spikes across neurons (Chapter 7) is expected to decrease if spikes are randomly shuffled among neurons. In other situations, it is possible that including spurious spikes in a spike train can increase the rate of information transfer, but this is only true if those spurious spikes themselves carry information; spikes at random times cannot increase the information. Likewise, random deletion of spikes that are really part of a recorded neuron's response can only lower the information.

The spike sorting algorithm. All of the following procedures are implemented in Matlab (The Mathworks, Natick, MA). Experimental data are saved in a proprietary format by Datawave's Discovery program; these files are then decoded and read into Matlab data structures. Although many experiments are performed at each recording site, data from the orientation-tuning experiment are usually used as a substrate for initial spike sorting. Occasionally, m-sequence responses are used for this purpose as well.

For initial clustering, 10,000 spikes are randomly chosen from the full data set. As mentioned above, each "spike" consists of 32 consecutive samples per recording channel (one for single electrodes, four for tetrodes) around the time of a threshold crossing. For data sampled at 30.3 kHz (a typical sampling rate in the experiments), this corresponds to 1.06 msec of data at each threshold crossing. Since the most reliable parts of the spike waveform are near the peak, analysis is generally restricted to 21 of the 32 samples, beginning four samples before the peak and ending 16 samples afterward. Spike waveforms from all of the tetrode channels are concatenated, so that the resulting waveform vector has 84 samples rather than 21. Unlike in the method of (Fee et al., 1996a), the waveforms are not resampled or interpolated, because such a procedure greatly slows the calculation, and because initial studies revealed that the final results are not significantly affected.

From the collection of 10,000 vectors, principal components are calculated by singular value decomposition (Press et al., 1992). The number of principal components is equal to the number of samples in each vector (21 or 84). Each principal component is associated with an eigenvalue, which represents the fraction of the variance in the full data set explained by that principal component. Of these principal components, only the first several (usually fewer than ten) are retained, enough to explain 90% of the variance. Taking the dot product of each spike waveform with the surviving principal components transforms the data into the space of the principal components. The result is a dramatic reduction of the dimensionality of the data, and hence a dramatic increase in the speed of the algorithm, without significant loss of features that are relevant to differentiating between spike waveforms from different neurons. Indeed, the procedure may even eliminate variability in the waveforms that would otherwise complicate the spike-sorting process.

Next, the mean of the transformed data is computed, separately for each dimension. To obtain the initial clusters, the mean is sequentially perturbed in two directions by only 0.1% of the standard deviation along each of the vector dimensions. The sign of the perturbation (positive or negative) along each dimension is chosen randomly. Each transformed spike waveform is then assigned to the perturbed vector that is closest in the Euclidean sense. This results in two clusters of transformed waveforms, each of which is bisected as described above. The entire process is iterated five times until 32 initial clusters are obtained; for some data sets, particularly those in which there are many isolable waveforms, six iterations (64 initial clusters) are required.

At this point, every spike in the data set has been assigned to one of the initial clusters. (Fee et al., 1996a) describe a procedure of updating cluster means "on the fly" so that systematic drift in the spike shape is tracked. Since the data sets in the monkey experiments are relatively short, spike drift is instead monitored by subjectively assessing the reliability of each final cluster from one stimulus condition to the next. When clusters become unreliable, or when the shapes of waveforms in a cluster change noticeably, a new classification is performed. This occurs only rarely, however, mostly because the parameters of the algorithm are such that spikes from a single neuron are typically as-

signed to more than a single cluster, and the differences among these clusters are sufficient to account for the spike drift.

The number of spikes in each of the initial clusters can vary substantially. Because small clusters tend to be unreliable, those that contain fewer than 0.5% of all spikes are eliminated, and the affected spikes are placed into a general-purpose cluster (designated *Cluster 0*). Each of these spikes will later be reassigned to one of the surviving clusters, but they do not play a role in determining the properties of the clusters. Each pair of surviving clusters is assigned a *connection strength* based on the interface energy between the set of points in the two clusters; the calculation proceeds as in (Fee et al., 1996a) and is not described here in detail. In general, this approach assigns higher connection strengths to clusters with many spikes along the border between them rather than to clusters that have similar means. The connection-strength calculation is the most computationally demanding in the algorithm, so the results are saved, together with the principal components, to facilitate reclustering of the data, if that is required.

A simple way of aggregating clusters is to combine pairs of clusters that have connection strengths greater than some threshold value. However, there is an additional criterion, following (Fee et al., 1996a), which takes into account a known property of spike train statistics—the refractory period. This criterion is based on whether the *interspike interval index*—the fraction of interspike intervals shorter than 10 msec that are also shorter than 3 msec—will change significantly if the two clusters are combined. Significance is determined by a chi-squared test, and the significance level is set to 5% (two-tailed). If merging two clusters would change this quantity, the clusters are not combined during the automatic portion of the algorithm even if they have a high connection strength. In situations where spikes fired in bursts (that is, spikes with very short interspike intervals) have different waveforms from isolated spikes, clusters that should be combined are kept separate at this stage. In these cases, which are relatively rare, the clusters will be combined in the user-intervention stage of the algorithm, as described below.

After each pair of clusters is combined, connection strengths are recalculated, and cluster pairs are again examined to determine whether they should be aggregated. Once this is completed, clusters are sorted into ascending order based on the interspike interval index. Cluster means (in both the space of principal components and the original waveform space) are recalculated and saved. In the final step of the automatic portion of the algorithm, each spike is independently reassigned to one of the cluster means, again according to the Euclidean distance. These same means are also used as the basis for classification of spikes from subsequent experiments. The choice to base the final reassignment on the Euclidean distance rather than on some interface potential can potentially result in incorrect associations, because waveforms from a particular spike may actually be closer to the mean waveform of another spike and because every spike must be assigned to some cluster. In practice, however, the result is that fewer "noise" spikes are assigned to actual clusters.

After the final reassignment, clusters are displayed graphically. An example is shown in Fig. 2.2. Of the 32 initial clusters, 15 survived the aggregation procedure. For each cluster, the mean spike waveform (thick white line) and 50 individual examples are plotted for each channel of the tetrode. Above each cluster, three numbers are given: the cluster index, the number of spikes in the cluster, and the interspike interval index. Clusters with high values of the interspike interval index are more likely to contain burst spikes.

The final assignment of clusters to single neurons is done by hand. In Fig. 2.2, seven of the 15 clusters (1, 3, 4, 5, 8, 9, 11) are considered noise. The waveforms in some of these clusters (e.g., 5 and 9) resemble extracellularly recorded spikes. However, the amplitudes of these spike-like waveforms are relatively small, and the fact that there is a peak near the beginning of each of these waveforms may simply be a consequence of the fact that waveforms are preselected on the basis of a threshold crossing at this particular sample. The "noise" clusters are more likely to contain spikes from more than one neuron, so to limit data contamination they are removed from further consideration.

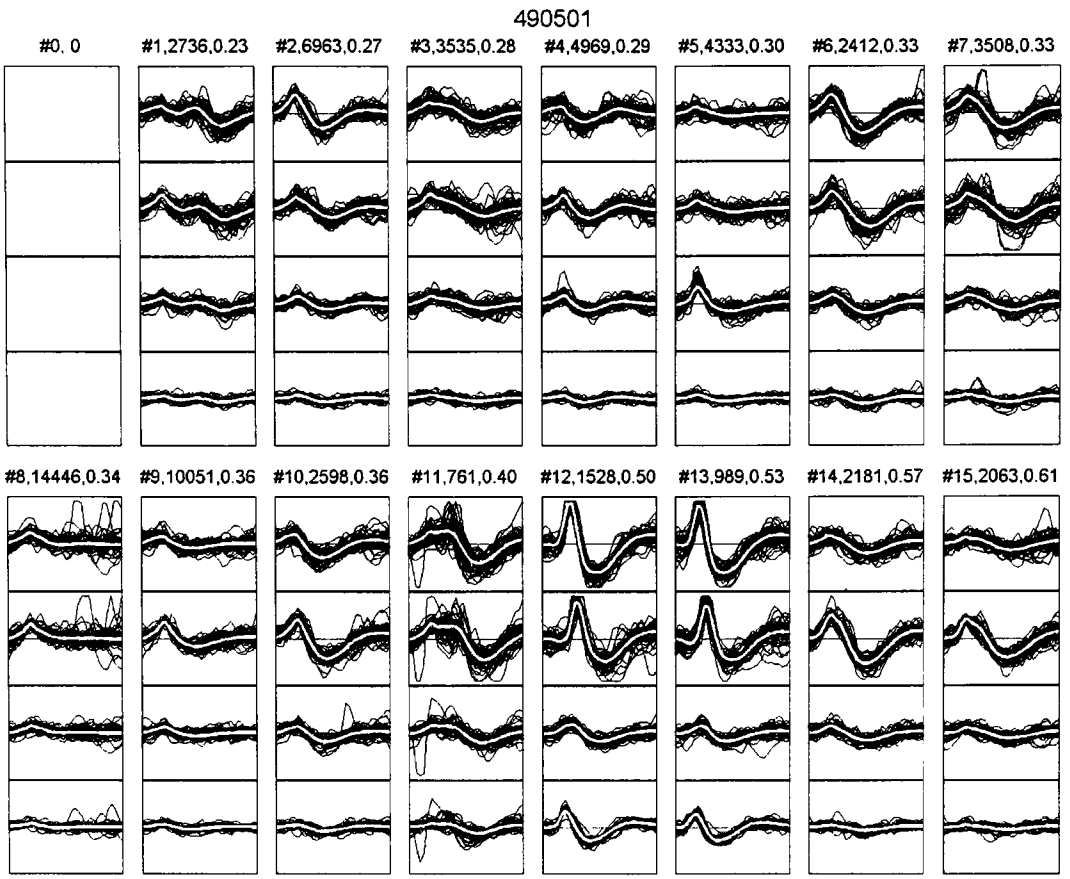

Figure 2.2. *Initial output of the spike sorting algorithm.* See text for details.

The best-isolated spikes in this example are in clusters 12 and 13—indeed, the spike amplitude saturates the sampling on two of the four channels—so that the spikes in these two clusters would be assigned to a single neuron. The spike waveforms in clusters 14 and 15 are also quite similar to one another, both in shape and in interspike interval index. There is a hint that these clusters may contain overlapping spikes from two single neurons: the presence of two cusps near the peak of the average waveform in channel 2 indicates that two underlying neurons might be firing at similar but slightly offset times. However, this may also be a consequence of waveform filtering, and in the absence of further information, the spikes in clusters 14 and 15 are assigned to a single neuron. Three additional single neurons can be identified in this example: cluster 10, the combination of clusters 6 and 7, and cluster 2. The waveforms corresponding to these three sin-

gle neurons are sufficiently dissimilar on the basis of shape, interspike interval index, peak-to-peak amplitude ratios (not shown), and orientation-tuning properties (also not shown). The requirement that orientation tuning be similar for spikes assigned to a single neuron makes it possible to distinguish spikes that seem similar to the experimenter, such as those in cluster 10 and clusters 14 and 15.

Clearly, the spike sorting procedure described here is not complete or perfect by any means. As discussed above, it is also very difficult to objectively evaluate the quality of the algorithm, since there is no additional information, beyond what has been used in the sorting process, about the neuron of origin of each waveform. The procedure, as it stands, is relatively fast (though not fast enough to implement in real time during the experiments), simple to carry out, and seemingly superior to the automatic (Autocut 3.0) and manual (Cluster Cutting 6.0) spike-sorting functionality of the Datawave system. It is likely that a 5% improvement in the quality of the algorithm would require a significant investment of time and energy, with uncertain results. However, it is also clear that neurophysiological recording techniques and the quality of the data obtained by these techniques would stand to benefit greatly from improved spike-sorting algorithms.

Chapter 3

The Power Ratio and the Circle Map

Summary

In this chapter, a method to analyze the dynamics of neuronal spike trains driven by external stimuli is described. The goal is to test the predictions of simple spike-generating models against extracellularly recorded neuronal responses. Through a new statistic called the *power ratio*, two broad classes of responses are distinguished: (1) responses that can be completely characterized by a variable firing rate (for example, modulated Poisson and gamma spike trains); and (2) responses for which firing-rate variations alone are not sufficient to characterize the dynamics (for example, leaky integrate-and-fire spike trains and Poisson spike trains with long absolute refractory periods). The responses of many visual neurons in the cat retinal ganglion, cat lateral geniculate nucleus, and macaque primary visual cortex are shown to belong to the second class, which implies that the pattern of spike times can carry significant information about visual stimuli. A related method, based on circle maps, is also developed. This method can uncover some of the underlying deterministic features of a noisy spike train. Spike trains of X-type retinal ganglion cells, in particular, are shown to be very similar to spike trains generated by a leaky integrate-and-fire model with additive, stimulus-independent noise that could represent background synaptic activity.

Introduction

A central issue in neuroscience is the question of whether neuronal spike trains *in vivo* are essentially random (Shadlen and Newsome, 1994; Shadlen and Newsome, 1998) or have temporal structure that might convey information in some form other than the

mean firing rate (MacKay and McCulloch, 1952; Rieke et al., 1997). Evidence is accumulating in favor of the second hypothesis (Hopfield, 1995; Singer and Gray, 1995): temporal codes have been found in the discharges of individual neurons both *in vitro* (Mainen and Sejnowski, 1995; Nowak et al., 1997) and *in vivo* (Cattaneo et al., 1981; Richmond and Optican, 1987; Mandl, 1993; Victor and Purpura, 1996; Mechler et al., 1998b). Moreover, spike timing in the visual cortex of monkeys bears a well-structured relationship to elementary features of visual stimuli, such as orientation, contrast, and spatial frequency (Victor and Purpura, 1996).

In some retinal ganglion cells and lateral geniculate nucleus (LGN) relay neurons, spike timing is sufficiently precise to be manifest as discrete peaks in peri-stimulus time histograms (PSTHs) (Berry et al., 1997; Reich et al., 1997; Tzonev et al., 1997). This result is consistent with a model that treats retinal ganglion cells as noisy, leaky integrate-and-fire (NLIF) devices, but it is also consistent with simpler models that treat retinal ganglion cell spike trains as modulated Poisson processes for which the measured PSTH is an estimate of the time-varying probability density for spike firing.

The purpose of the analysis described in this chapter is to determine whether the statistical structure of the responses of visual neurons can support the encoding of stimulus-related information in spike times and interspike intervals in addition to modulation of the overall firing rate. Unlike the work presented in later chapters, the analysis does not seek to evaluate the quantity or quality of information conveyed by those responses. Instead, it develops a statistic—the *power ratio*—whereby the responses of two large classes of spike-generating models can be distinguished. In the first class of models, the external stimulus changes the neuron's firing rate but does not affect the dynamics of spike generation; models in this class are termed *simply modulated renewal processes* (SMRPs). In the second class of models, both firing rate and spike-generation dynamics—in particular, the time-dependent interspike interval distribution—are affected by the external stimulus; this class includes the NLIF neuron.

The results indicate that most neurons in the early stages of the mammalian visual system fire spike trains that cannot be considered SMRPs. For X-type retinal ganglion

cells in particular, the NLIF model provides an excellent fit to the statistics of evoked spike trains. Neurons in the primary visual cortex are not as well modeled by NLIF neurons, but they, too, fire spike trains that are not SMRPs. Thus, the results constrain the range of models that can be considered appropriate but simple descriptions of the spiking dynamics of real neurons. At the same time, the results provide independent evidence that real neurons can support temporal coding schemes, which can dramatically increase the rate of information transfer in a neural system.

By an extension of the power-ratio method, a *circle map* can be fit to the responses of retinal ganglion cells to periodic stimuli. The advantage of the circle map formalism is that, in some cases, it allows the calculation of fixed points that reflect the underlying (noise-free) deterministic process that could have given rise to the individual responses. For the X-type retinal ganglion cell responses, dynamical features identified by the circle maps are also consistent with the NLIF model.

Methods

In this section, the procedures for obtaining experimental data are described. In addition, the fundamental analysis methods used in this chapter, particularly the generation of model data for comparison with real data, are explained. Both data-driven models, which involve resampling real data or drawing samples from parametric descriptions of real data, and first-principle-driven models, which involve simplified versions of detailed, biophysics-inspired models, are described. Calculation of the *interval maps*, the *power ratio* statistic, and *circle maps* is described as necessary in Results.

Electrophysiology. The results presented in this chapter are based on extracellular recordings of LGN neurons and their retinal inputs in anesthetized cats, made in collaboration with colleagues at The Rockefeller University (Ehud Kaplan, Pratik Mukherjee, Mavi Sanchez-Vives, and Tsuyoshi Ozaki), and of primary visual cortex (V1) neurons in anesthetized macaque monkeys, made in collaboration with colleagues at Weill Medical College (Jonathan Victor and Ferenc Mechler). Experiments were performed on nine

male and three female adult cats, and on three male adult monkeys, which weighed roughly 3 kg each. The general outline of the monkey experiments is described in Chapter 2 and is reiterated here in brief; cat experiments are similar and are described here, as well.

For the cats, anesthesia is initiated by intramuscular injections of xylazine 1 mg/kg (Rompun) and ketamine 10 mg/kg (Ketaset) and is maintained throughout surgery and recording with intravenous injection of thiopental 2.5%, 2-6 mg/kg/hr (Pentothal). Paralysis is induced and maintained with vecuronium 0.25 mg/kg/hr (Norcuron). For the monkeys, anesthesia is induced with ketamine 10 mg/kg, supplemented as needed by methohexital 0.5-1 mg/kg (Brevital) boluses during the preparatory surgery, and maintained with sufentanil 3 μg/kg bolus, 1-6 μg/kg/hr (Sufenta). Paralysis is induced and maintained with pancuronium 1 mg bolus, 0.2-0.4 mg/kg/hr (Pavulon).

Gas-permeable hard contact lenses are used to prevent corneal drying, and artificial pupils (3 mm diameter) are placed in front of the eyes. The optical quality of the animals' eyes is checked regularly by direct ophthalmoscopy. Optical correction with trial lenses is added to optimize grating responses at a viewing distance of 114 cm. Blood pressure, heart rate, expired carbon dioxide, and core temperature are continuously monitored and maintained within the physiological range.

Tungsten-in-glass electrodes (Merrill and Ainsworth, 1972) record extracellular potentials from individual cat LGN neurons and from their primary retinal-ganglion-cell inputs in the form of synaptic (S) potentials (Kaplan et al., 1987), or from monkey V1 neurons. The electrode signals are amplified, filtered, and monitored conventionally. Action potentials of single neurons are selected by a window discriminator (Winston Electronics) for the cats. For the monkeys, analog waveforms are identified and differentiated by criteria such as peak amplitude, valley amplitude, and principal components (Datawave). Visual stimuli are created on a white CRT (Conrac 7351, 135 frames/sec, 80 cd/m^2 mean luminance) for the cats, or on a green CRT (Tektronix 608, 270 frames/sec, 150 cd/m^2 mean luminance) for the monkeys, by specialized equipment (Milkman et al., 1980). Action potentials are timed to the nearest 0.1 msec.

Cat retinal ganglion and LGN neurons are classified as X-type or Y-type and *on*-center or *off*-center (Enroth-Cugell and Robson, 1966). Spatial frequency tuning and contrast response functions are measured with drifting sinusoidal grating stimuli; between six and ten separate contrasts or spatial frequencies are sampled for each neuron. Responses to each contrast or spatial frequency are usually recorded for 16 sec, but occasionally for longer periods of time (up to 256 sec), before the next stimulus is presented.

Monkey V1 neurons are classified as simple or complex on the basis of whether their response to a drifting grating of high spatial frequency is predominantly a modulated response at the driving frequency, for simple cells, or else an elevation of the mean firing rate, for complex cells (Skottun et al., 1991). Contrast response functions are also measured with sinusoidal grating stimuli presented at the optimal orientation, spatial frequency, and temporal frequency. The stimuli are presented for four to ten seconds at each contrast, in random order, and the entire set of contrasts is presented, in different random orders, four to eight times. For the analysis described in this chapter, the responses to each stimulus are considered to be one continuous steady-state record.

Simple models. ***Poisson.*** A resampling technique (Victor and Purpura, 1996) is used to create artificial spike trains with the same PSTH as a measured spike train. Each spike in the original spike train is associated with a randomly chosen response cycle, an operation that preserves the set of spike times (and hence the PSTH) but destroys the distribution of those times among the individual cycles (and hence the interspike interval histogram, or ISIH). The resulting spike train has the statistics of a modulated Poisson process.

Modified Poisson. To test the hypothesis that firing rate is in part determined by slow variations in responsiveness, and that such slow variations can account for any difference between recorded data and Poisson-resampled data, a procedure equivalent to the *exchange resampling* of (Victor and Purpura, 1996) is performed. Each response cycle is assigned the same number of spikes as in the corresponding cycle of the real response, but the spike times themselves are drawn at random, without replacement, from the entire

collection of spikes. Since all spikes are used exactly once, the PSTH of the resampled spike train is identical to the PSTH of the original spike train.

Gamma. Artificial spike trains with similar (though not identical) PSTHs to those of measured spike trains, but with the interspike interval statistics of n^{th}-order modulated gamma processes, are also generated. A gamma process may be considered to have a relative refractory period, the duration of which changes with the stimulus strength. For very high-order gamma processes, the firing is clock-like and approaches the behavior of a non-leaky integrate-and-fire model. Gamma processes have been used in the past to describe the statistics of retinal ganglion cell spike trains (FitzHugh, 1958; Troy and Robson, 1992). Here, modulated gamma spike trains are generated by drawing a random number to determine whether a spike is fired in each 0.1 msec time bin, where the probability for spike firing in each bin is determined from the linearly interpolated PSTH. For an n^{th}-order gamma process, the model is given n chances to fire in each bin, and every n^{th} spike is kept in the final spike train.

Absolute refractory periods. To generate spike trains with absolute refractory periods, the n^{th}-order gamma model is modified to hold the firing probability at zero for a fixed time, equal to the desired refractory period, after each spike. This procedure effectively shifts the overall ISIH toward longer times, leaving a gap equal in duration to the refractory period.

NLIF model. The model discussed here is a noisy variation of the leaky integrate-and-fire model (Knight, 1972a). It is a highly reduced version of the Hodgkin-Huxley equations for neuronal firing, in which the state variable $V(t)$ plays the role of the membrane potential. The model neuron "fires" when $V(t)$ reaches a threshold V_{th}, after which $V(t)$ is reset to zero. In the simulations in this chapter, the input to the model is a sinusoidally modulated current. Poisson-distributed noise shots of steady rate, uniform size, and random polarity are added to the state variable at each 0.1 msec time step. In the absence of noise, this model phase locks: if the leak rate is sufficiently fast compared to the stimulus cycle, and the stimulus sufficiently strongly modulated, the spike times in all stimulus cycles are identical.

Formally, the model is described by:

$$\frac{dV}{dt} = -\frac{V(t)}{\tau} + [S_0 + S_1 \cos(2\pi f t + \varphi)] + N(t), \qquad (3.1)$$

where $V(t)$ is the state variable of the model; t is the time within the stimulus cycle (sec); τ is the time constant of the leak (sec); S_0 is the mean input level (sec^{-1}); S_1 is the contrast (sec^{-1}); f is the input temporal frequency (Hz); φ is the input phase (radians); and $N(t)$ is the input Poisson shot-noise (sec^{-1}). The overall firing rate of the neuron depends on the threshold V_{th}, the noise, and the deterministic input—both the driving force (when $S_1>0$) and the shot noise process.

The results in this paper are based on the model's responses to stimuli of ten different contrasts about a mean of S_0=1 sec^{-1}, ranging from 0% (S_1=0 sec^{-1}) to 100% (S_1=1 sec^{-1}). The threshold is chosen to be 75% of the steady-state value of the state variable in the absence of input modulation, or 0.015 in the dimensionless units of the model. The time constant τ is set to 20 msec, the temporal frequency f to 4.2 Hz, and the phase φ to π radians, which aligns the period of strongest firing with the middle of the response cycle. Several different noise shot sizes are tested, ranging from 0 to ±0.0016, but the shot rate is kept constant at 1000 shots/sec.[1] Integration is performed numerically by Euler's method with a time step of 0.1 msec (Press et al., 1992).

The state variable $V(t)$ is measured in dimensionless units, following (Knight, 1972a). These units can be considered voltages, since the state variable loosely corresponds to the membrane potential of real neurons. Dimensionless units are preferred over actual voltages because the NLIF model does not describe the detailed biophysical processes that occur in real neurons. Nonetheless, despite the difference in units, the model is similar, in many ways, to the one described by (Shadlen and Newsome, 1998), the primary difference being the absence of the deterministic input in the latter model. The presence of the deterministic input makes it possible to use fewer, smaller-amplitude

[1] In the absence of leakage, when τ is very large, the rate of change of the variance in the accumulated charge due to the Poisson shots is simply q^2r, or 2.6 msec^{-1} for the largest q used here. For general τ, the induced voltage noise reaches its steady state standard deviation of $q\sqrt{r\tau/2}$, or 0.0051 for the largest q, with a time constant $\tau/2$, or 10 msec.

noise shots. Even so, the noise is sufficient to cause substantial jitter in spike timing without dominating the response statistics, in contrast to the model of (Shadlen and Newsome, 1998), where the noise itself is dominant.

Results

This section is divided into three major parts. First, a brief discussion of renewal processes is presented. Second, a multi-step procedure for classifying neuronal responses into one of the two classes mentioned in the Introduction is described. Third, the use of circle maps and their fixed points to characterize neuronal responses is explained.

The multi-step response classification procedure itself has three parts. The first step is to apply a data-driven time transformation that flattens the PSTH and converts SMRPs into unmodulated renewal processes. The second step is to plot the distribution of the interspike intervals on the transformed timescale. The third step is to calculate an index that is sensitive to variations in the interspike interval distribution, and to compare that index to the one obtained from Poisson processes with the same PSTH.

Renewal Processes. Spike trains of renewal processes are characterized by the fact that all interspike intervals are independent and identically distributed (Papoulis, 1991). This implies that the firing rate is necessarily constant, on average. The probability that a spike is fired within a brief time window dt at a particular time t since the previous spike can be written as:

$$p(t)dt = rg(rt)dt\,, \tag{3.2}$$

where r is the firing rate and g is some dimensionless function that integrates to 1. This function g describes the shape of the interspike interval distribution from which successive spikes are drawn at random. For a Poisson process, the simplest renewal process, g is exponential, so:

$$p(t)dt = re^{-rt}dt\,. \tag{3.3}$$

Real neuronal spike trains have a time-varying firing rate, so the concept of a renewal process must be expanded to encompass this. This is done by eliminating the requirement that all interspike intervals be identically distributed, while maintaining the requirement that the intervals be independent. Thus, intervals may depend on the stimulus, but they do not reflect the firing history before the previous spike. The effect of this modification is to create a *modulated renewal process* for which the firing probability now depends on a variable firing rate $r(t)$ and on an interval distribution that changes in time. Thus, the probability that a spike at time t_0 is followed by a spike in time window dt at time t_0+t can be written as $p(t|t_0)dt$.

Interval Maps and the Power Ratio. ***Time transformation.*** In order to compare responses to different stimuli, a *demodulation transformation* is applied. This time transformation replaces the original time axis by the integral of the PSTH (FitzHugh, 1957; Gestri, 1978; Cattaneo et al., 1981). For each real time t, a transformed time $u(t)$ is given as:

$$u(t) = \int_0^t \frac{r(t')}{\bar{r}} dt' \tag{3.4}$$

where $r(t)$ is the firing rate at time t, estimated by the PSTH, and $\bar{r}$ is the mean firing rate over the entire cycle. This invertible transformation effectively expands time during portions of the response when the firing rate is high and compresses time when the firing rate is low, so that the PSTH in transformed time is flat. The transformation changes the internal clock of the neuron from one that ticks in units of real time into one that ticks in units of instantaneous firing probability. Distributed across all response cycles, the same number of spikes are fired in each unit of transformed time.

Because spike trains are inherently discontinuous, this transformation is simple to implement. The transformed time of a given spike is equal to the fraction of spikes (across all cycles) that occur earlier—that is, its fractional rank within the sorted list of spike times—multiplied by the cycle duration. Ties are broken randomly. Fig. 3.1 shows

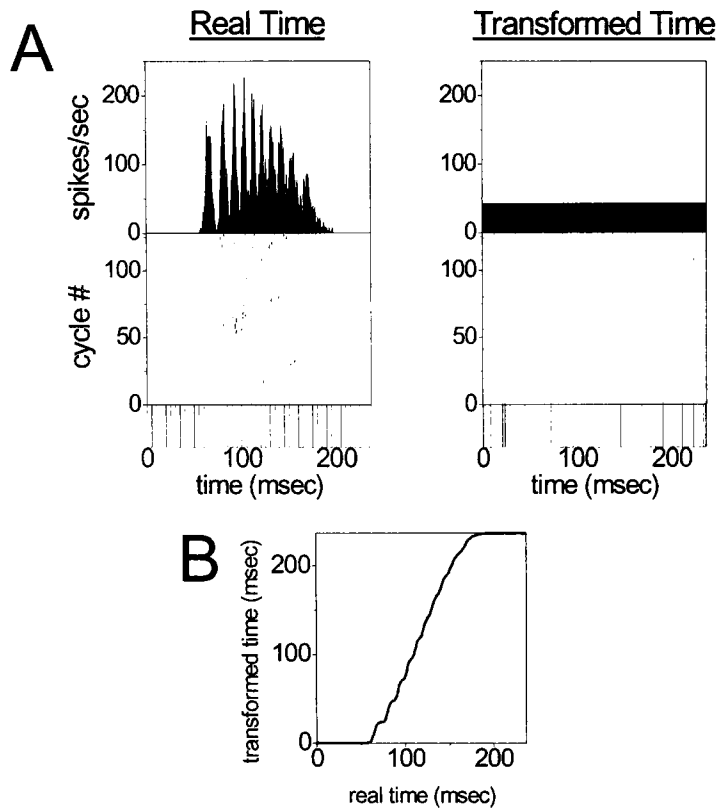

Figure 3.1. *Time transformation.* Results from 128 cycles of the response of a noisy, leaky integrate-and-fire (NLIF) model (shot size 0.0004) to a 4.2 Hz sinusoidal input current at unit contrast. **A**: response in real time (left panel) and transformed time (right panel). In the middle of each panel is a raster plot that shows the spike times in each cycle, which are collected in 1 msec bins to form the peri-stimulus time histograms (PSTHs) shown above the raster plots. Scaling the spike times on the left by the integral of their PSTH yields the demodulated, *transformed-time* version of the spike train (right panel), for which the PSTH is flat. Evenly spaced tick marks in real time (bottom of left panel), are separated non-uniformly by the time transformation (bottom of right panel)—that is, the distance between adjacent ticks is expanded when the response is strong, in the middle of the cycle, and contracted when the response is weak, early and late in the cycle. The apparent discrepancy between the number of tick marks in the two panels is due to the fact that many of the transformed-time tick marks fall on top of one another. **B**: another view of the time transformation for this spike train. The thick solid line, equivalent to the integral of the real-time PSTH, shows the value of transformed time to which each value of real time is mapped. The thin line along the diagonal represents the null transformation, in which transformed and real times are identical. When the slope of the thick solid line is greater than 1, the transformation expands time, and when the slope is less than 1, the transformation contracts time.

the effects of the time transformation for data derived from the NLIF model at unit contrast and shot size 0.0004.

Time transformation and renewal processes. The term *simply modulated renewal process* (SMRP) is introduced here to denote a subset of modulated renewal processes, the spike trains of which are uniquely converted by the time transformation into spike trains of unmodulated renewal processes with the same mean firing rate (Gestri, 1978). Examples of SMRPs are modulated Poisson and gamma processes.

Many modulated renewal processes are not SMRPs, however. Examples include spike trains generated by the NLIF model and by models with fixed absolute refractory periods (Table 3.1). Models

that do not produce SMRPs contain parameters, such as leak time and refractory period, that are not affected by external stimuli or firing-rate variations. Although these parameters are fixed in real time, they are not fixed in transformed time because they get scaled by the local firing rate. As explained below (*Specificity of the power ratio*), this results in interspike intervals that are not identically distributed in transformed time, as they would be for a true renewal process.

Interval maps. To distinguish between SMRPs and other models that could have produced a measured spike train, it is useful to plot the *interval map*. The interval map relates each spike time (plotted on the horizontal axis) to the subsequent interspike interval (plotted on the vertical axis). Examples of interval maps are shown in Figs. 3.2A (real time) and 3.2B (transformed time). The left column uses a spike train generated by the NLIF model, and the right column uses a Poisson-resampled spike train with the same PSTH (see Methods). The interval map is reminiscent of the *intervalogram* (Funke and Worgötter, 1997), but, unlike the intervalogram, the interval map uses no binning or averaging. It includes all the information necessary to reconstruct both the PSTH and ISIH of a given data set. To obtain the PSTH, simply add up the number of points in each time bin along the horizontal axis, and to obtain the ISIH, add up the number of points in each time bin along the vertical axis. In all the interval maps in this paper, the PSTH is plotted above the interval map, and the ISIH, rotated 90°, on the right-hand side.

Model	SMRP?	Power ratio in SMRP range?
Noisy, leaky integrate-and-fire (NLIF)	No	No (except for low contrasts)
Poisson-resampled spike train	Yes	Yes
Modified Poisson spike train	Yes	Yes
Modulated Poisson process	Yes	Yes
Modulated gamma process	Yes	Yes
Non-leaky integrate-and-fire	Yes	Yes
SMRP, fixed absolute refractory period	No	Yes (except for excessively long refractory periods)

Table 3.1. *Simple spike-generating models.* For each model, two questions are asked: (1) Is the model a simply modulated renewal process (SMRP)? and (2) Does the power ratio fall in the range expected for an SMRP?

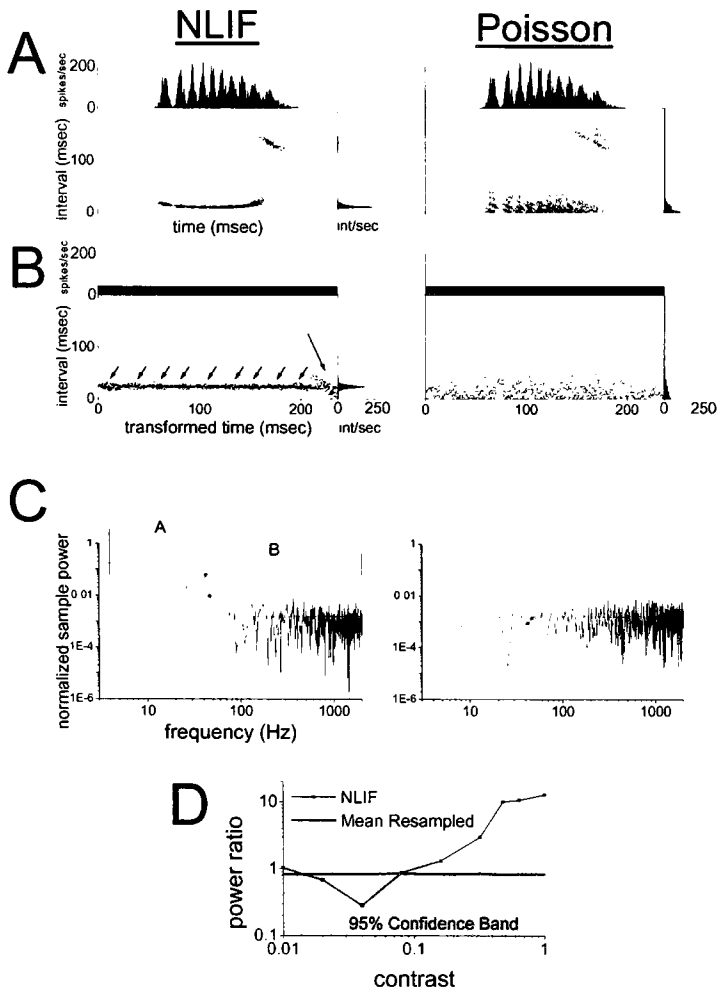

Figure 3.2. *The power ratio.* Same spike train as in Fig. 3.1. In panels **A-C**, the left column shows data taken from the NLIF model, and the right column shows the same data resampled into a modulated Poisson process (see Methods). **A**: interval maps in original time; **B**: interval maps in transformed time. The large arrow in the left panel of **B** represents the resetting that occurs during the silent period of the response to each cycle, and the small arrows represent small-scale resettings that occur within the response to each cycle. The marginal distributions are also shown—the PSTH along the horizontal axis and the ISIH along the vertical axis; **C**: sample power at each harmonic normalized by the total non-DC sample power. The solid line represents the normalized sample power of the test data, whereas the dashed line represents the mean normalized sample power at each harmonic for 1000 Poisson-resamplings of the data. The first *n* harmonics, where *n* is the smallest integer larger than the mean number of spikes in each response cycle, are indicated by circles. Filled circles bracket the cell's firing rate. Because the preponderance of the sample power for the NLIF spike train occurs in the first *n* non-DC harmonics, the ratio of the mean sample power in the first *n* harmonics ("A") to the mean sample power in all non-DC harmonics ("B") is the calculated *power ratio*; **D**: power ratio of the NLIF spike train as a function of stimulus contrast, and mean and 95% confidence band for the power ratio of 100 Poisson-resampled spike trains at each contrast.

In Fig. 3.2A, the distinct cluster of points at the end of the response cycle in both interval maps corresponds to the final interval in each cycle. In real time, these final intervals, which span the portion of the stimulus cycle during which no spikes are fired (that is, when the PSTH is zero), are far longer than the other intervals. This is true for both NLIF and Poisson spike trains, because the two spike trains have identical PSTHs. In transformed time, however, spikes are equally likely to be fired at all points in the stimulus cycle, so the PSTH is never zero, and there is therefore no well separated long-

interval cluster. For the Poisson process in particular, the distribution of interspike intervals is largely independent of transformed time, so the transformed-time interval map of the Poisson-resampled spike train is nearly uniform (Fig. 3.2B, right panel). In other words, there are no privileged spike times for a Poisson process. For the NLIF spike train, however, the long-interval cluster is clearly retained (Fig. 3.2B, left panel, large arrow), indicating that the distribution of interspike intervals is not independent of transformed time, and that some spike times are, in fact, privileged over others. The explanation for this lies in the leakiness of the NLIF model, which ensures that if the external stimulus is sufficiently weak (for example, the negative phase of a high-contrast sinusoid), the state variable falls to near its minimum before beginning to recharge as the stimulus strengthens again. This resynchronizes the state variable and causes the first spike in each cycle to occur at a highly reliable, privileged time, regardless of the time of the previous spike.

Thus, the first spike in each cycle is independent of the last spike in the previous cycle for the NLIF spike train, whereas for the Poisson process, the two spike times depend strongly on one another. This is opposite to the relationship between the final interspike interval and the last spike time in each cycle, which are correlated for the NLIF spike train and independent for the Poisson process. Thus, for the NLIF spike train, because of the resynchronization, the final interval is longer when the last spike occurs relatively early, and shorter when the last spike occurs relatively late. The time transformation does not eliminate this dependence, which is reflected in the distinct final-interval cluster (large arrow). To a lesser extent, the leakiness and resetting properties of the NLIF model are also reflected in the smaller interval clusters that occur throughout the cycle (small arrows).

A power-spectral approach is used to quantify the extent to which the interval map of a particular spike train deviates from that of a Poisson process with the same PSTH. The absence (for Poisson processes) or presence (for spike trains that could not have been generated by a Poisson process) of slow changes in the interval map across transformed time can be detected in the power spectrum. Practically, the sample power in the transformed-time interval map at each harmonic of the stimulus cycle is calculated and normalized by the total sample power of the modulated (non-DC) harmonics (Fig.

3.2C). The prominent clusters visible in the transformed-time interval map of the NLIF spike train selectively increase the sample power in the low-frequency harmonics (Fig. 3.2C, left panel). For this reason, it is useful to focus on the sample power in the first n harmonics, where n is the mean number of spikes in each response cycle, rounded up to the next integer. These harmonics are signified by circles in Fig. 3.2C, and the firing rate itself is bracketed by the pair of solid circles.

The *power ratio* is defined as the mean sample power in the first n frequency components of the interval map divided by the mean sample power of all modulated (non-DC) components:

$$PR = \frac{\frac{1}{n}\sum_{k=1}^{n}|H_k|^2}{\frac{2}{N}\sum_{k=1}^{N/2}|H_k|^2}. \tag{3.5}$$

Here, N is the total number of interspike intervals and H_k is the discrete Fourier component at the k^{th} harmonic of the transformed-time interval map. The discrete Fourier components are given as:

$$H_k = \sum_{j=1}^{N} h_j e^{\frac{2\pi i k t_j}{T}} \tag{3.6}$$

(Bendat and Piersol, 1971), where t_j is the transformed time of the j^{th} spike, h_j is the j^{th} interspike interval (in transformed time), and T is the duration of the stimulus cycle. The dashed lines in both panels of Fig. 3.2C represent the mean normalized sample power of 1000 Poisson resamplings of the original spike train (NLIF on the left, Poisson on the right).[2] The first n harmonics of the interval map reflect features that occur no more than n times during the stimulus cycle—once per spike, on average. When more than n com-

[2] Although one might have expected the mean power spectrum of a Poisson interval map to be flat, it actually has a low-frequency cutoff. This is because the interspike intervals (on the vertical axis) determine the values of successive spike times (on the horizontal axis), so the interval map is weakly correlated, even for a Poisson spike train. The unnormalized power at the k^{th} harmonic of a Poisson interval map depends explicitly on the firing rate r, and it can be shown to have an expected value of

$$|H_k|^2 = \frac{2}{(rT)^2}\left(\frac{(2\pi k / rT)^2}{1+(2\pi k / rT)^2}\right) \tag{3.7}$$

For small k, power is proportional to k^2; as k increases, the power grows toward an asymptotic value of $2/(rT)^2$. This accounts for the frequency dependence of the power spectrum at low harmonics (Fig. 3.2C).

ponents are considered, the ability to distinguish between NLIF and Poisson spike trains is diminished, since the power spectra of the interval maps are similar at high frequencies.

The power ratio of the NLIF spike train in Fig. 3.2 is 12.92, and of the Poisson-resampled spike train, 0.80. In order to determine whether each spike train could have been generated by a Poisson process, a large number of Poisson resamplings, each of which has exactly the same PSTH as the original spike train, are performed. A spike train is considered to deviate significantly from the Poisson expectation if its power ratio is larger than the power ratios of 95% of the Poisson-resampled spike trains. For the NLIF spike train of Fig. 3.2, the deviation is highly significant ($p<0.0001$), whereas for the Poisson-resampled spike train, not surprisingly, it is not.

In Fig. 3.2D, the power ratio of NLIF spike trains, together with the mean and 95% confidence region of the power ratios from 1000 Poisson resamplings at each contrast, are plotted against stimulus contrast. It is clear that at the four highest depths of modulation, 0.32 and above, the NLIF spike trains are readily distinguished from spike trains generated by Poisson processes with the same PSTH, and the spike trains become less and less Poisson-like as the contrast increases.

Of course, it is not always necessary to calculate the power ratios in order to distinguish between the transformed-time interval maps of Fig. 3.2B. The Poisson interval map plainly differs from the NLIF interval map not only in the lack of clusters—the main feature captured by the power ratio—but in many other ways as well, including the shape of the summed ISIH on the vertical axis (exponentially decaying for the Poisson spike train, peaked for the NLIF spike train). In fact, the ISIH of the NLIF data, measured in transformed time, resembles much more closely the interval distribution of a high-order gamma process. However, the power ratio is a useful tool to distinguish SMRPs as a class from other modulated renewal processes. Thus, the power ratio distinguishes the NLIF spike train in Fig. 3.2 even from spike trains generated by a high-order gamma process with a similar PSTH and ISIH (see below, *Specificity of the power ratio*, and Fig. 3.5).

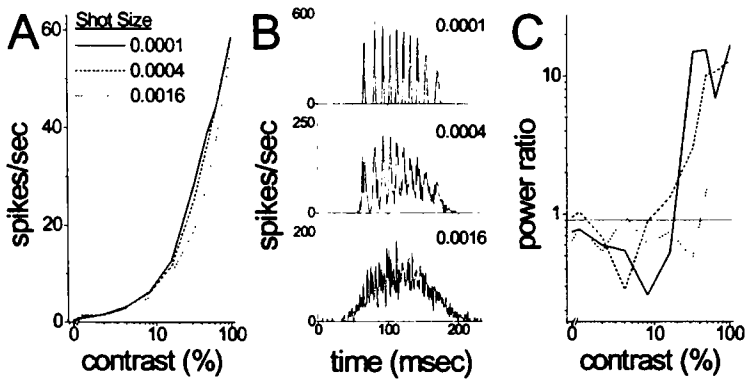

Figure 3.3. *NLIF model with different amounts of input noise.* **A**: response strength, as a function of contrast, measured as the magnitude of the Fourier component at the driving frequency (4.2 Hz), for three noise levels; **B**: PSTHs for three noise levels at unit (100%) contrast. Peaks in the PSTH, very sharp when the input noise is low, disappear when the input noise is made large; **C**: power ratios for the same noise levels as in **A**, calculated as a function of stimulus contrast. The power ratio can distinguish the responses at all three noise levels from Poisson spike trains with the same PSTH, as long as the contrast is sufficiently high. The solid horizontal line represents the mean Poisson expectation for the power ratio across all data sets. Notice that the power ratio is smaller when the input noise is larger, indicating that NLIF spike trains with large input noise are more Poisson-like.

Sensitivity of the power ratio. The behavior of the NLIF model with different amounts of input noise is investigated. The stimuli are 4.2 Hz sinusoidal currents at ten contrasts, ranging from 0 to 1. The magnitude of the response at the driving frequency is largely insensitive to the input noise (Fig. 3.3A). The shape of the PSTH at high contrast, however, depends significantly on the amount of input noise. Spike times are highly precise and reproducible when the input noise is low, which is reflected in a peaked PSTH (Fig. 3.3B, top panel). When the input noise is high (Fig. 3.3B, bottom panel), the PSTH peaks disappear, indicating that in this situation the noise dominates the deterministic input and that the spike times are no longer precise and reproducible. However, even with high input noise, the power ratio is capable of distinguishing high-contrast responses from Poisson spike trains (Fig. 3.3C).

The contrast-dependence of the deviation of NLIF spike trains from Poisson spike trains is also seen in Fig. 3.4A, which shows the fraction of 25 independently generated NLIF spike trains that have power ratios outside the Poisson range, as a function of contrast, for the same three noise levels. When the noise is low, the spike trains are either always Poisson-like or always inconsistent with Poisson processes—hence the jump from 0 to 1 between contrasts 0.16 and 0.32 contrast for the shot size of 0.0001. As the noise increases, spike trains of intermediate contrast are sometimes consistent with Poisson processes and sometimes not. But even at the highest noise level, the responses to unit-contrast stimuli are nearly always inconsistent with Poisson processes. In other words, the

power ratio distinguishes Poisson spike trains from non-Poisson spike trains even when the PSTH shows no evidence of precise spike times.

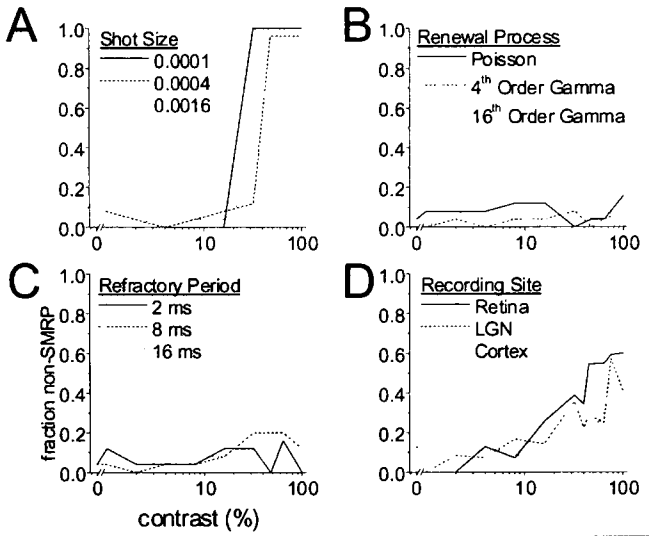

Figure 3.4. *Summary of analysis of model and real spike trains.* Each panel shows the fraction of spike trains with non-SMRP power ratios for a variety of model and real neurons, as a function of stimulus contrast. **A**: NLIF models at different noise levels (shot sizes); **B**: gamma processes of different orders; **C**: Poisson processes with refractory periods of different durations; **D**: retinal ganglion, LGN, and V1 neurons. Note that stimuli of higher contrast tend to evoke non-SMRP spike trains in real neurons and in NLIF and long-refractory-period models. However, when the refractory period is in the physiologic range (on the order of a few msec), higher contrasts do not cause the spike trains to deviate reliably from the SMRP expectation, suggesting that the addition of a fixed absolute refractory period to a Poisson process does not adequately account for the firing patterns of real neurons.

Specificity of the power ratio. Typical SMRPs cannot be empirically distinguished from Poisson-resampled spike trains by means of the power ratio. Fig. 3.5 shows the interval maps in real and transformed time for spike trains generated by several models. Again, the marginal distribution along the horizontal axis (plotted above the interval map) is the PSTH, and the marginal distribution along the vertical axis (plotted to the right of the interval map) is the ISIH. Recall that in transformed time, the PSTH is flat by construction. The power ratio is listed in the right column, along with the p value from the multiple-resamplings significance test. Power ratios with p values less than 0.05 are considered to indicate deviation from the Poisson expectation. Fig. 3.5A again shows the NLIF response to a unit-contrast stimulus; the power ratio is highly significant. Fig. 3.5B shows the same NLIF response transformed into a modified Poisson process by the exchange-resampling procedure (see Methods). Fig. 3.5C shows the spike train of a fourth-order gamma process, and Fig. 3.5D shows the spike train of a 16^{th}-order gamma process, where the firing probabilities in each case are derived from the PSTH of the original NLIF spike train. In all three cases (Figs. 3.5B-D), the power ratio is well within the Poisson range ($p>>0.05$), indicating that the power ratio cannot distinguish between different SMRPs, even when the ISIHs are different.

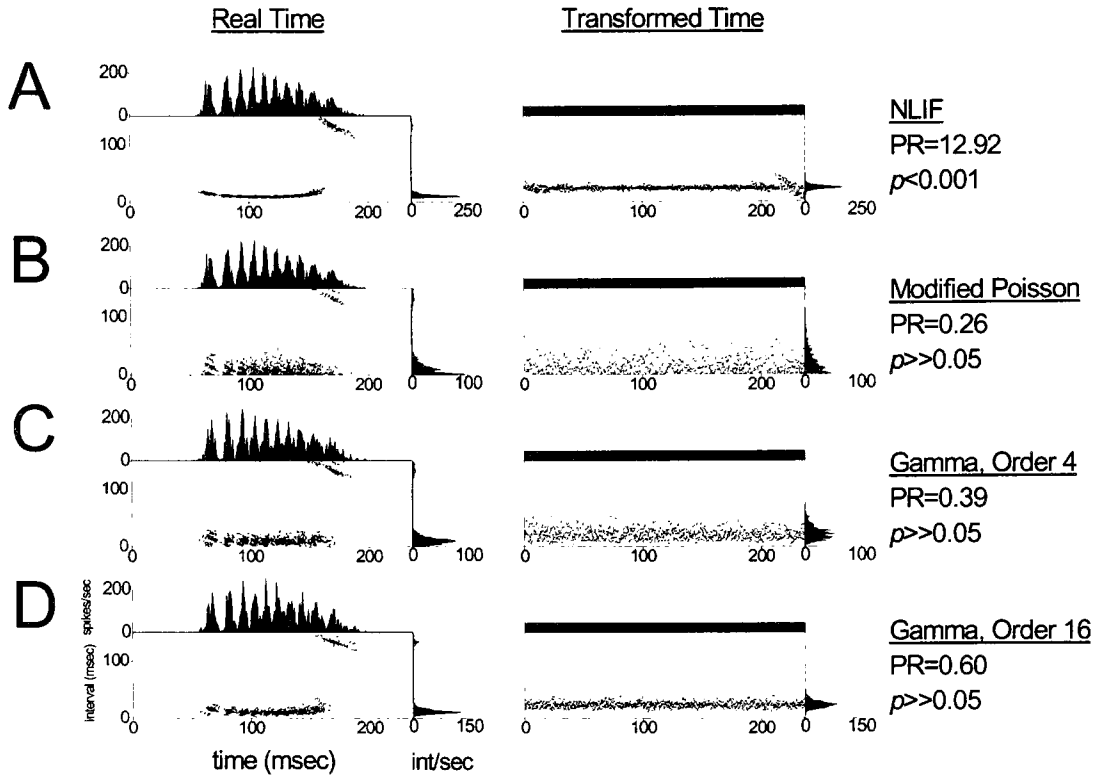

Figure 3.5. *Interval maps, histograms, and power ratios for four simple spike-generating models.* The interval maps are presented in both real and transformed time. Modified Poisson responses are generated by the exchange-resampling procedure, whereas gamma responses are generated from estimates of the PSTH (see Methods). The power ratio for each response and its significance level are shown in the right column. **A**: NLIF model (noise level 0.0004); **B**: modified Poisson process; **C**: fourth-order gamma process; **D**: 16th-order gamma process. Note that all of the responses derived from SMRPs have power ratios in the Poisson range and are therefore indistinguishable by this index.

These results are summarized in Fig. 3.4B. The fraction of spike trains inconsistent with SMRPs is plotted against stimulus contrast for each type of SMRP. Again, 25 responses are simulated at each of ten contrasts, and the firing probability for each condition is set to match the PSTH for an example of the NLIF model with a shot size of 0.0004. At all contrasts, SMRP spike trains are empirically indistinguishable from Poisson spike trains, and hence from one another, by the power ratio. These results also indicate that the presence of a refractory period, by itself, does not produce a power ratio distinguishably different from Poisson, since gamma-distributed spike trains have a relative refractory period in which the duration of the refractory period is deterministically related to the strength of the input.

Models that contain fixed refractory periods measured in units of real time, however, are not SMRPs. Such models are reasonable on biophysical grounds, because absolute refractory periods are thought to result from fundamental properties of neurons' membranes and ion channels, independent of external stimuli. Because these refractory periods are measured in real time, as explained above, they are distorted non-uniformly by the time transformation, which is determined by the overall modulation of a neuron's response and thus varies during the response. When the firing probability is low, the transformation compresses time so that the transformed-time refractory period is very short. Conversely, when the firing probability is high, the transformation expands time and thus stretches the transformed-time refractory period. Therefore, the minimum interspike interval in transformed time varies throughout the response, producing a modulated transformed-time interval map. These modulations may be picked up by the power ratio, which may fall outside the Poisson range.

Fig. 3.6 shows examples of interval maps from non-SMRP spike trains that are derived from the same NLIF spike train. The firing probability in each case is derived from the observed PSTH of the original NLIF spike train. The gray-shaded areas at the bottom of the interval maps in Figs. 3.6B-D represent the durations of the refractory period in real (left column) and transformed (right column) time; interspike intervals that fall in these gray-shaded areas are disallowed. In transformed time, as expected, the duration of the refractory period varies during the course of the response in all three cases. Fig. 3.6B shows the spike train of a Poisson process with a 2 msec refractory period, an appropriate duration for retinal ganglion cells (Berry and Meister, 1998). The power ratio for this spike train is well within the SMRP range ($p>>0.05$) despite the presence of the physiological refractory period. Fig. 3.6C shows the spike train of a Poisson process with a 16 msec refractory period, which is excessively long for a real neuron. In this case, the overall firing rate falls significantly, and the PSTH barely resembles the PSTH of the original spike train (Fig. 3.6A) because of the limitations imposed on the maximum firing rate by the refractory period. Indeed, because the refractory period is so long, it is impossible to obtain a PSTH identical to that of the original data. It is not surprising that this power ratio is well outside the SMRP range, but it is perhaps surprising that it is still so much lower than the power ratio of the NLIF spike train. Fig. 3.6D shows the response of

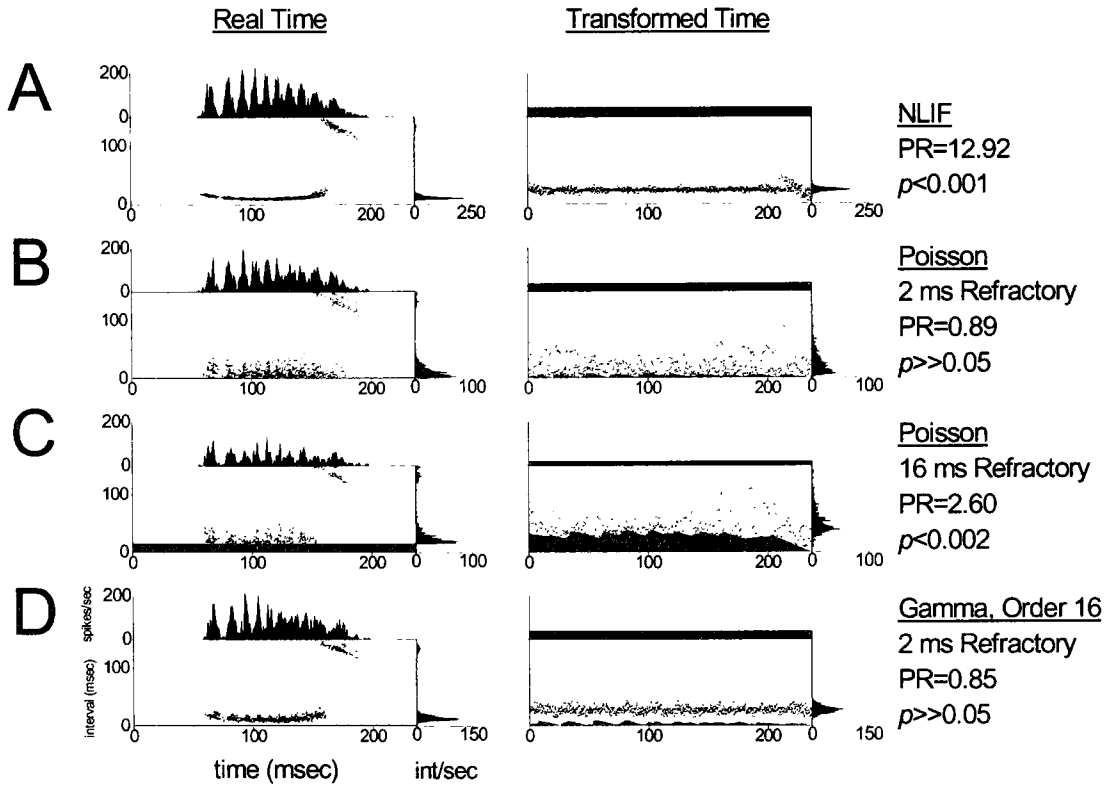

Figure 3.6. *Interval maps, histograms, and power ratios for models with fixed absolute refractory periods.* **A**: NLIF model (noise level 0.0004); **B**: Poisson process with 2 msec refractory period; **C**: Poisson process with 16 msec refractory period; **D**: 16th-order gamma process with 2 msec refractory period. The gray-shaded area at the bottom of each interval map for the refractory-period models covers the range of disallowed interspike intervals. The minimum interspike interval is fixed in real time but variable in transformed time. This means that refractory-period models are not SMRPs, even though an unphysiologically long refractory period (here, 16 msec) is required to push the power ratio outside the SMRP range.

a 16th-order gamma process with a 2 msec refractory period, shorter than the typical relative refractory period of the gamma process itself. Despite the presence of both absolute and relative refractory periods, the power ratio is within the SMRP range.

The results for the refractory period models are summarized in Fig. 3.4C. Again, for each model, 250 spike trains are created, 25 at each of ten contrasts. The firing probability for each condition, derived from the PSTH of the NLIF model with a shot size of 0.0004, dictates the target modulation of the firing rate. However, the degree to which the target modulation is achieved varies with the duration of the refractory period, as discussed above. Only when the refractory period is very long—above 16 msec—do the power ratios deviate significantly and consistently from the SMRP range. As mentioned

above, such refractory periods are unreasonably long for visual neurons of the types studied here.

Neuronal responses. The neuronal data presented here come from 12 cats and three monkeys. For the cats, data from retinal ganglion cells (recorded as synaptic potentials in the LGN) and LGN relay neurons are presented. In many cases, the responses of the LGN neurons are recorded simultaneously with their predominant retinal inputs. For the monkeys, the data come from neurons in the primary visual cortex (V1). In both cats and monkeys, stimuli are drifting sinusoidal gratings of several contrasts at fixed spatial and temporal frequencies, although for a few of the retinal ganglion and LGN cells, the contrast is fixed at 0.4 and the spatial frequency is varied. Contrasts are geometrically spaced, so that in most experiments the contrast is less than 0.4. In general, the neuron's entire classical receptive field is stimulated by the drifting gratings, but in a few cat neurons the center of the receptive field is stimulated in isolation while the surround illumination is kept fixed at the same mean luminance.

Retinal ganglion cells. Altogether, 342 spike trains from 39 retinal ganglion cells were recorded. Sample interval maps from four neurons, in both real and transformed time, are shown in Fig. 3.7. Fig. 3.7A presents data from an X-type, *on*-center retinal ganglion cell. The power ratio, listed in the right column, is well outside the SMRP range, as it is for the X-type, *off*-center retinal ganglion cell in Fig. 3.7B. Note the presence of clusters of spikes in the transformed-time interval maps in Figs. 3.7A-B, in particular the large final cluster. These are similar to the clusters of spikes seen in the transformed-time interval maps of NLIF spike trains (Figs. 3.2, 3.5, and 3.6), which suggests that phenomena similar to the reset and leak of the NLIF model may underlie spike generation in X-type retinal ganglion cells.

For the Y-type retinal ganglion cell responses shown in Figs. 3.7C-D, the power ratio is also outside the SMRP range, but NLIF-like clusters are not obvious. The differences between X-cell and Y-cell interval maps may be due to differences in the underlying spike generating mechanisms. However, it is more likely that the explanation lies in the fact that Y-cells' responses to drifting gratings typically involve a prominent eleva-

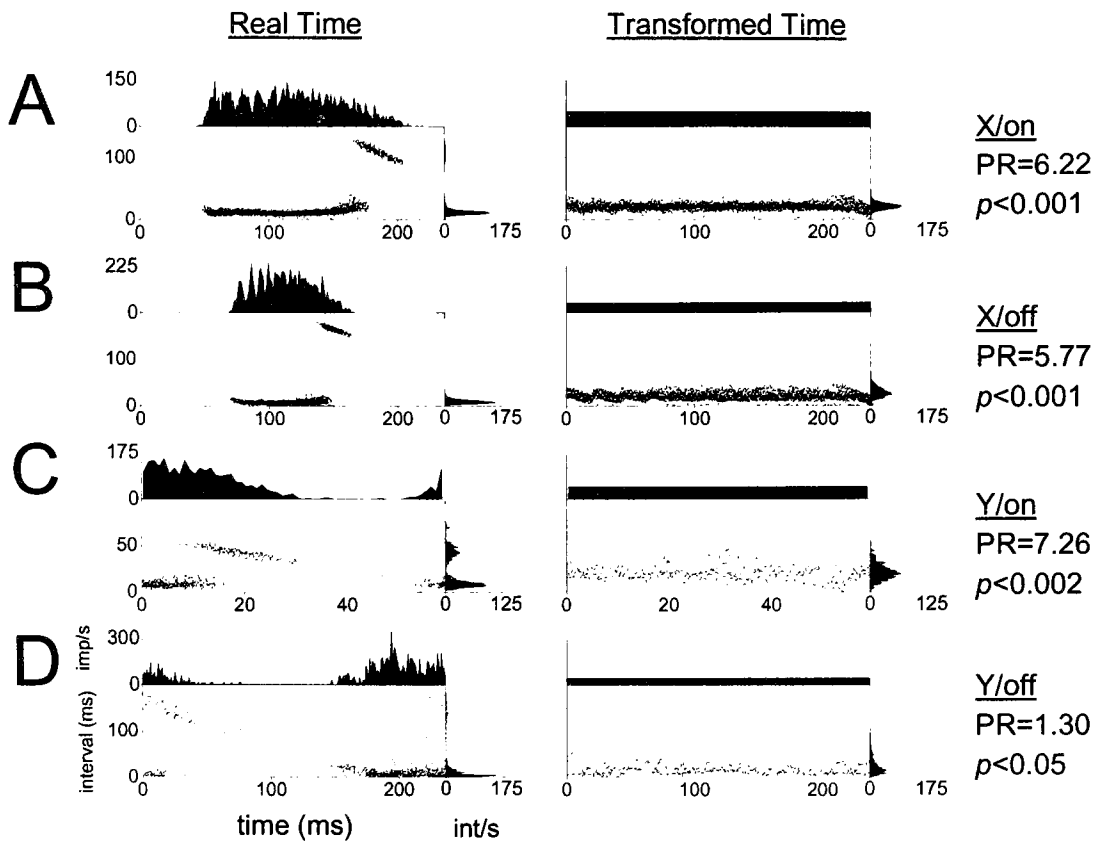

Figure 3.7. *Interval maps, histograms, and power ratios for four cat retinal ganglion cells.* The stimuli are all drifting sinusoidal gratings of optimal spatial frequency, and the recordings are made from synaptic potentials in the LGN. **A**: X-type, *on*-center cell; unit contrast; 4.2 Hz; **B**: X-type, *off*-center cell; unit contrast; 4.2 Hz; **C**: Y-type, *on*-center cell; unit contrast; 16.9 Hz; **D**: Y-type, *off*-center cell; 0.4 contrast; 4.2 Hz. The spike train of panel D is the dominant retinal input to the LGN cell of Fig. 3.8D.

tion of the mean firing rate and a smaller modulated component than do X-cells' responses, at least at high spatial frequencies (Enroth-Cugell and Robson, 1966). Since the transformed-time interval map, and by extension the power ratio, are highly sensitive to interactions between the modulated component of the response and spike train dynamics, and since Y-cell responses to drifting gratings are less modulated than X-cell responses, the interaction of the response modulation and the dynamics of spike generation may be less obvious in Y-cells.

LGN relay neurons. 322 spike trains from 36 LGN neurons were recorded. Four examples are shown in Fig. 3.8. The power ratios for the spike trains of Figs. 3.8A-C are well outside the SMRP range, and the interval maps are, accordingly, highly non-

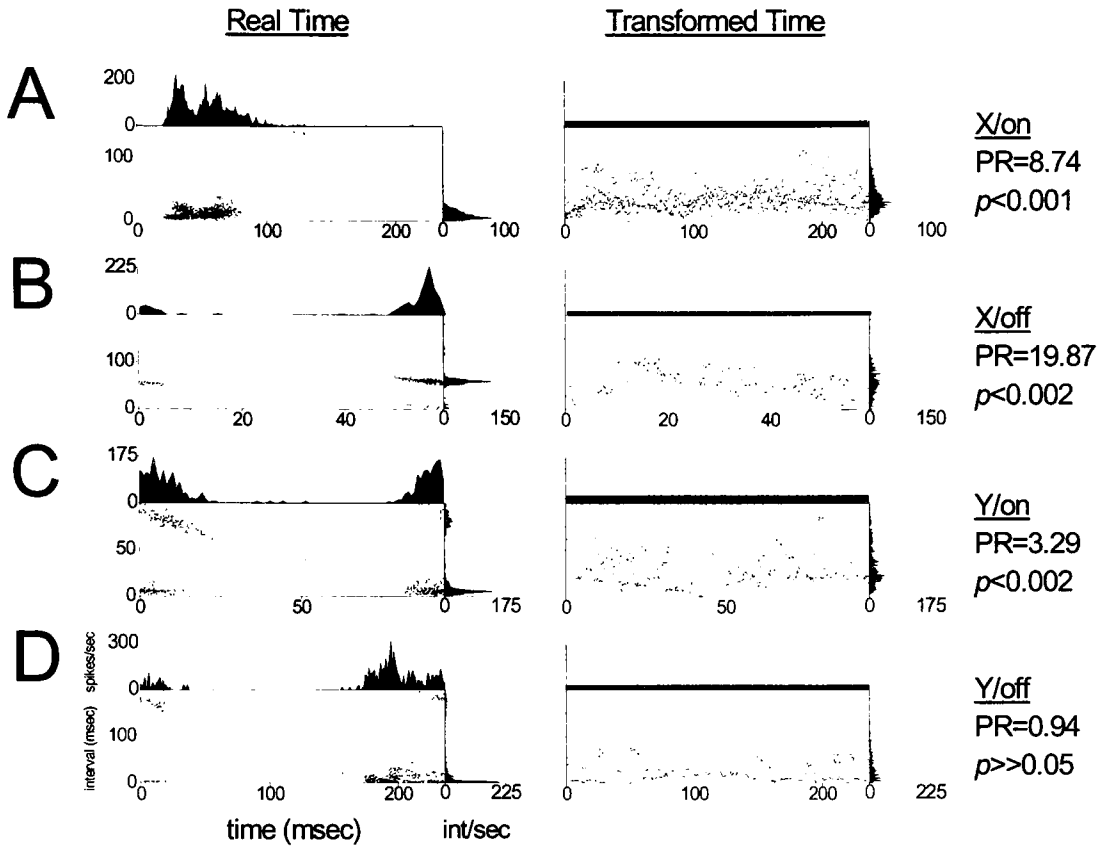

Figure 3.8. *Interval maps, histograms, and power ratios for four cat LGN neurons.* The stimuli are again drifting sinusoidal gratings of optimal spatial frequency. **A**: X-type, *on*-center cell; unit contrast; 4.2 Hz; **B**: X-type, *off*-center cell; unit contrast; 16.9 Hz; **C**: Y-type, *on*-center cell; 0.75 contrast; 10.6 Hz; **D**: Y-type, *off*-center cell; 0.4 contrast; 4.2 Hz. The response in panel D is driven primarily by the spike train shown in Fig. 3.7D.

uniform. For the data in Fig. 3.8A, recorded from an X-type, *on*-center LGN neuron, the transformed-time interval map seems to have several horizontal bands at the beginning (Funke and Worgötter, 1997), a single band in the middle, and the hint of a broad NLIF-like cluster at the end. Furthermore, at a transformed time of 100 msec, the mean interval becomes abruptly longer. A similar interval map is seen for the Y-cell of Fig. 3.8C, which also has a non-SMRP power ratio. These transformed-time interval maps differ strikingly from the transformed-time interval maps of NLIF and X-type retinal ganglion cell spike trains.

However, the response of the X-type, *off*-center LGN cell of Fig. 3.8B conforms more closely to the NLIF expectation. In this response, evoked by a 16.9 Hz, unit-contrast drifting grating, there is typically only one spike per cycle. This can be inferred

from the real-time interval map, which contains a prominent band at an interval of roughly 60 msec (corresponding to one spike per cycle), as well as much fainter bands below and above it (corresponding, respectively, to two spikes per cycle and one spike every other cycle). In transformed time, the interval map consists of a single, downward-sloping NLIF-like cluster. This distinct non-uniformity, which is not a simple consequence of the fact that there is typically only one spike per cycle (Poisson-resampled spike trains with the same PSTH do not have this non-uniformity), induces a remarkably high power ratio of 19.87, well outside the SMRP range.

Of the 36 LGN neurons, 31 were recorded simultaneously with their retinal input, accounting for 276 spike trains at each recording site (the number of stimulus conditions for each neuron was not identical). Interval map and power ratio analysis shows concordant power ratios in 81% of the cases. However, 43 (16%) of the paired spike trains were inconsistent with SMRPs in the retina but not the LGN, whereas ten (4%) were inconsistent with SMRPs in the LGN but not the retina. An example is shown in Figs. 3.7D and 3.8D, where the interval maps are derived from the responses of a Y-type, *off*-center LGN neuron and its retinal input. Although the real-time PSTHs for the two cells are similar, the power ratios and transformed-time interval maps are not. In particular, the response of the retinal ganglion cell is not consistent with an SMRP, whereas the response of the LGN neuron is. This is not simply due to the fact that the retinal ganglion cell response has more spikes than the LGN response: reducing the number of trials in the retinal ganglion cell response to make the number of spikes equal to the number of spikes in the LGN response does not bring the power ratio of the retinal ganglion cell response into the SMRP range. This result may signify a fundamental change in the underlying spike-generating mechanism from the retina to the LGN, which may selectively delete certain spikes, or, alternatively, may enter a mode in which "bursts" of spikes are fired (Mukherjee and Kaplan, 1995).

V1 neurons. In monkey V1, 113 spike trains from 19 neurons were recorded, of which four examples are shown in Fig. 3.9. In all four spike trains—two from simple cells and two from complex cells—the power ratio is significantly outside the SMRP range. The transformed-time interval maps of both simple-cell responses (Figs. 3.9A-B)

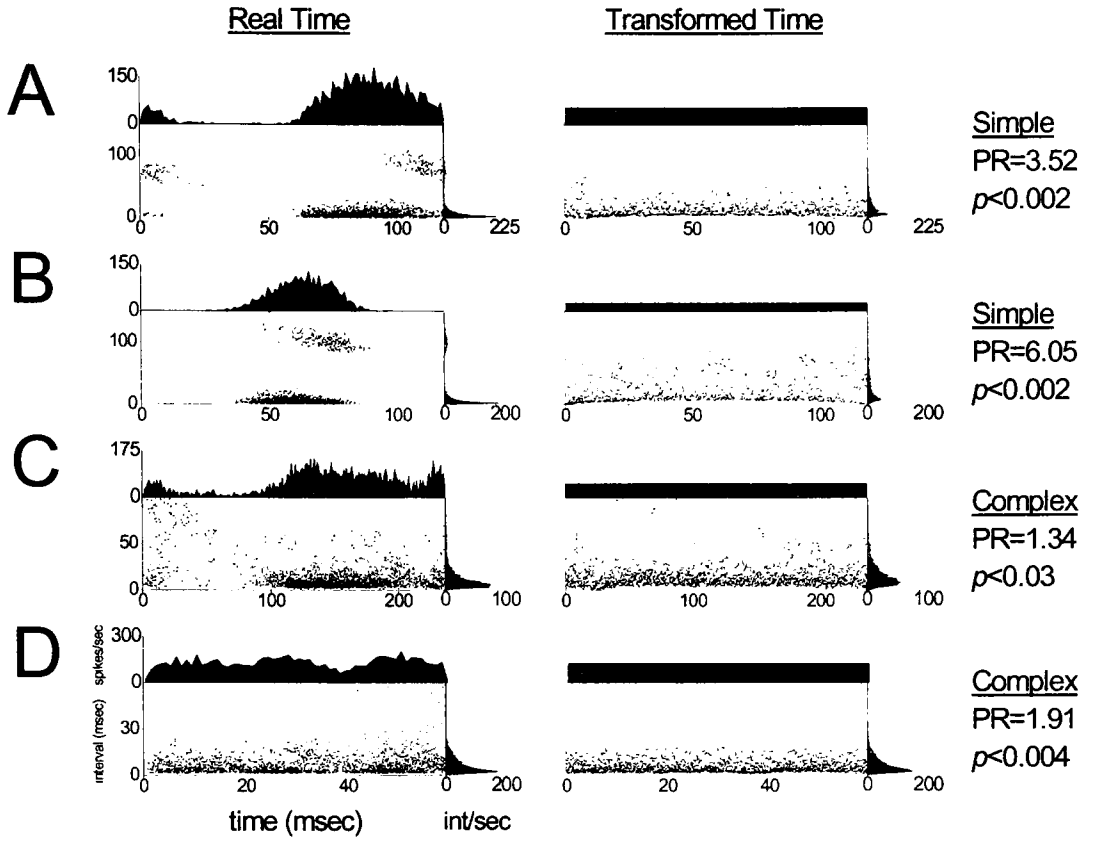

Figure 3.9. *Interval maps, histograms, and power ratios for four monkey V1 cells.* The stimuli are again drifting sinusoidal gratings of optimal spatial frequency. **A**: simple cell; unit contrast; 8.4 Hz; **B**: simple cell; unit contrast; 8.4 Hz; **C**: complex cell; unit contrast; 4.2 Hz; **D**: complex cell; unit contrast; 16.9 Hz.

show no clear evidence of NLIF-like clusters, but rather reveal a dense and non-uniform band of points along the bottom margin. These points are likely to correspond to bursts of spikes fired within a few milliseconds of one another. Since the intervals between burst spikes are stereotyped in real time, they become variable in transformed time. The transformed-time interval map is therefore non-uniform as well, causing the power ratio to fall outside the SMRP range. Thus, the presence of bursts, which are prominent in cortical cells and are sometimes thought to convey stimulus-related information (DeBusk et al., 1997) (see Chapter 5), is indicative of an underlying spike-generating mechanism that is not an SMRP.

The complex cell in Fig. 3.9C does not fire spikes in clear, stereotyped bursts. For this cell, the primary non-uniformity in the transformed-time interval map occurs near 25

msec in transformed time and is sufficient to lift the power ratio out of the SMRP range. The power ratio of the complex cell of Fig. 3.9D, by contrast, is outside the SMRP range, but there is no obvious explanation for the modulation. This last cell has an atypically high firing rate (around 125 spikes/sec), and the interval map is constructed from over 2400 spikes. The large amount of data gives rise to an extremely reliable estimate of the local interspike interval distributions, so that even small deviations from uniformity are likely to be picked up by the power ratio. Thus, even though the power ratio for this response is only 1.91, it is still well outside the SMRP range.

Results across all recordings are summarized in Table 3.2. The fraction of spike trains inconsistent with SMRPs decreases twofold from the retina to the cortex. The fraction of neurons at each recording site that fire at least one spike train inconsistent with an SMRP also falls nearly twofold. Because multiple spike trains are collected from each neuron, Bonferroni's correction (Bland, 1995) is used for the comparison across neurons to control for the possibility that a response is significant by chance alone. Thus, if m spike trains are collected from a given cell, at least one of those spike trains is required to have a power ratio outside the SMRP range with a p value of $0.05/m$. The decrease in number of non-SMRP spike trains from retina to LGN is significant by a χ^2 test (1 d.o.f., $p<0.001$), whereas the decrease from LGN to V1 is not significant. However, comparing retinal ganglion and LGN responses, which are recorded in cats, with cortical responses, which are recorded in monkeys, is tenuous at best.

At all three recording sites, the fraction of spike trains that fall significantly out-

Recording Site	Spike Trains		Cells			
			All Stimuli		High-Contrast Stimuli	
	Total	*Non-SMRP*	*Total*	*Non-SMRP*	*Total*	*Non-SMRP*
Retina	342	130 (38%)	39	26 (67%)	31	21 (68%)
LGN	322	76 (24%)	36	17 (47%)	28	14 (50%)
V1	92	21 (19%)	19	7 (37%)	19	9 (47%)

Table 3.2. *Summary of results from three recording sites (cat retinal ganglion, cat LGN, and monkey V1).* For each recording site, the fraction of spike trains with power ratios outside the SMRP range is provided. Also provided is the fraction of cells that fired non-SMRP spike trains in response to drifting-grating stimuli of a wide range of contrasts (most of which were below 0.4) and in response to stimuli of high contrast (greater than 0.4) alone.

side the SMRP range depends strongly on the stimulus contrast (Fig. 3.4D). Responses of real neurons to high-contrast stimuli are more often inconsistent with SMRPs than responses to low-contrast stimuli, just as they are for the NLIF and long-refractory-period models. Since the stimulus set is heavily weighted toward low contrasts—67% of the stimuli have contrasts of 0.4 or less—the Bonferroni correction likely results in an overly conservative calculation of the number of cells that fire non-SMRP spike trains. Restricting the analysis to stimuli with contrast greater than 0.4 typically reduces m, or the number of responses per cell, by a factor of three. On the basis of these high-contrast responses, a somewhat higher proportion of neurons at each site—nearly 50% in the cortex—fire non-SMRP spike trains (Table 3.2).

Statistics and utility of the power ratio. The power ratio method does not rely on the use of sinusoidal stimuli or steady-state responses—it could equally well be applied to spike trains evoked by repeated, transiently presented stimuli. Moreover, it is perhaps surprising that a large number of spikes are not required to obtain a useful estimate of the power ratio: 200 to 300 spikes, distributed over at least 16 trials or cycles of the drifting grating, are often sufficient to indicate the presence of a non-SMRP response. Thus, the number of spikes required to apply this method is comparable to the number of spikes required to estimate the PSTH. Nonetheless, the power ratios of non-SMRP responses increase as more and more cycles are added. This is because these power ratios detect non-uniformities in the transformed-time interval maps that are reinforced by the spikes in the additional trials. In this way, the power ratio can be considered to be a measure of the signal-to-noise ratio of a spike train's deviation from SMRP dynamics. Although a data set of only 200 to 300 spikes may be sufficient to indicate that a spike train is inconsistent with an SMRP, longer data sets provide greater sensitivity.

Circle maps. So far, the transformed-time interval map has served as a basis for determining whether a given spike train could have been generated by an SMRP. The results indicate that the interval maps of retinal ganglion cell responses are quite similar to those of spike trains generated by the NLIF model, but that LGN and V1 interval maps

are qualitatively different, despite also having power ratios that fall outside the Poisson (and SMRP) range. Here, a stronger test of the NLIF model is developed and discussed. This test takes advantage of the transformed-time interval maps to analyze the detailed dynamics of spike trains fired in response to periodic stimuli. This test uses interval maps to construct a model of NLIF and retinal ganglion cell dynamics on the basis of circle maps (Keener et al., 1981).

As above, the formalism is developed around the responses of a deterministic, leaky integrate-and-fire process with superimposed stimulus-independent noise. For sufficiently large input, the spike times in every response cycle are nearly identical as a consequence of the leak-related phase-locking (Knight, 1972a). The stimulus-independent input noise serves to "blur" the spike times, which induces trial-to-trial variability in the exact time of each spike and is reflected in the fact that the PSTH peaks are far from the delta functions that would be anticipated in the absence of noise (Fig. 3.3B). In other words, the underlying deterministic process produces precise spike times, whereas the input noise produces a dispersion around those times.

Circle maps can be used to retrieve the underlying precise spike times from such noisy responses. A circle map is a tool from dynamical systems theory that provides a powerful framework for the analysis of periodically-forced nonlinear systems (Procaccia et al., 1987). It is an example of a *first-return map*, in which the current value of a parameter is related to the subsequent value of the same parameter. In this analysis, the relevant parameter is the time of a spike within the response cycle.

A *circle map* is a one-dimensional function that maps a circle onto itself by carrying each point of the circle clockwise to a new image point. Here, the circle is the stimulus cycle. It is represented by the interval $[0,T)$, where T is the period of the stimulus, and the sequence $\{\theta_1,\theta_2,\ldots\}$ indicates the phases (occurrence times) of individual spikes. Formally, the circle map is written:

$$\theta_{n+1}=g(\theta_n) \tag{3.8}$$

where θ_{n+1} is computed mod T. It maps the time θ_n of one spike to the time θ_{n+1} of its successor spike. The m^{th} iterate of the circle map, $g^m(\theta_n)$ $[\equiv g(g^{m-1}(\theta_n))]$, maps the time of

one spike, θ_n, to the time of its m^{th} successor, θ_{n+m}. A circle map may have *fixed points*, which are values θ^* such that $g(\theta^*)=\theta^*$. Additionally, the circle map may have m^{th}-*iterate fixed points*, that is, fixed points of $g^m(\theta_n)$. For neural data, in which there are m spikes in each stimulus cycle, it is these m^{th}-iterate fixed points that are sought. A *stable* fixed point has the property that all nearby points converge to it upon repeated iteration of the map. Conversely, an *unstable* fixed point has the property that all nearby points diverge from it upon repeated iteration of the map. The presence and stability of individual fixed points are unaffected by topological transformations of the stimulus cycle, such as the time transformation. Stable m^{th}-iterate fixed points correspond to a stable pattern of firing of m spikes per stimulus cycle.

The overall procedure for building circle maps from spike data is as follows. First, a subset of the responses is selected for further analysis. Second, Fourier analysis is used to smooth the first return map of the selected responses, measured in transformed time, which results in an analytic circle map. Third, the derived circle map is iterated to find the iterate fixed points, both stable and unstable. Finally, the fixed points are converted into real time and related to the original PSTH.

In greater detail, the first step of the procedure is to select for analysis only stimulus periods that elicit identical numbers of spikes. This ensures that the *winding number* of the circle map (the number of iterates of the map required to traverse the full 360°) will be an integer. Note that the number-of-spikes criterion does not require that spikes fall at particular times during each trial, or even at reliable times across trials. If two selected trials are separated by several unselected ones (because of differences in the number of spikes), the final actual interspike interval of the first selected response is used, rather than the interval to the first spike of the next selected response.

Next, the transformed-time interval map of the spikes in the selected trials is constructed. From the interval map, a function $f(t)$ is calculated by Fourier analysis. The transformed spike time t is used interchangeably with the phase θ as the independent variable. The function $f(t)$ includes only the significant frequency components through the n^{th} harmonic, including the DC, where n is the number of spikes per response cycle

(Fig. 3.10A). This cutoff is chosen because the first n harmonics are the ones that are elevated for NLIF-like non-SMRPs. A Fourier component is considered significant if the normalized sample power at that harmonic is larger than the corresponding value in a criterion fraction (0.98) of many (typically 100) Poisson-resampled data sets. Thus, for data generated by an SMRP (unlike the NLIF data of Fig. 3.10), 2% of the Fourier components are expected to measure as significant.

To find the trajectory of a given point in the circle map, the *cobwebbing* procedure is used (Strogatz, 1994). Specifically, the circle map is iterated m times and the values of θ_n at which the m^{th} iterate crosses the diagonal $\theta_{n+m}=\theta_n$ are computed. If the iterated circle map

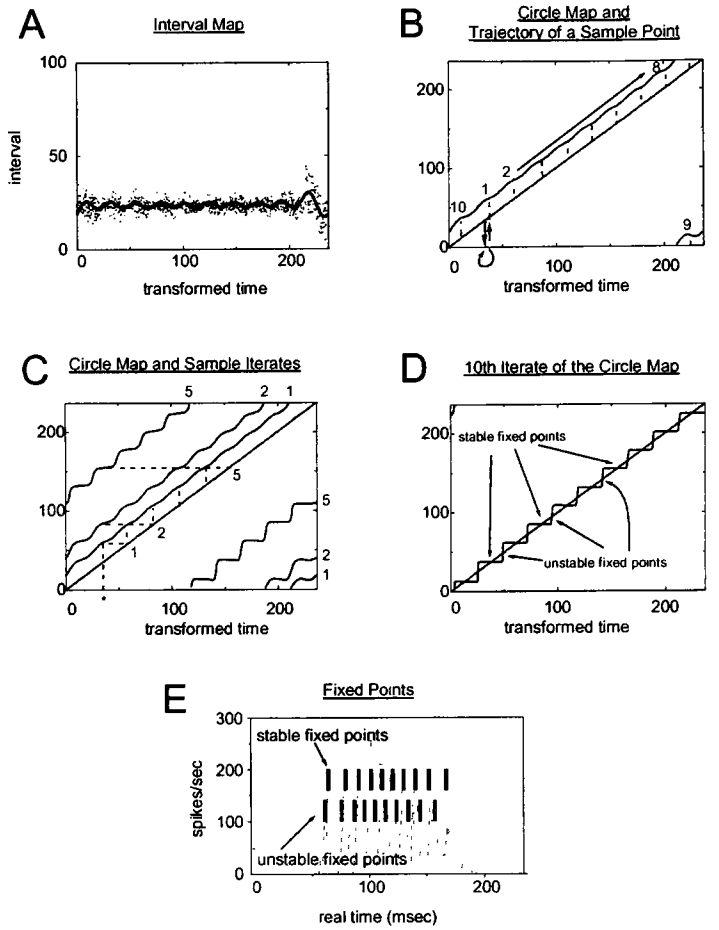

Figure 3.10. *Construction of circle maps from spike data.* The data are taken from the NLIF model (input noise 0.0004). **A**: Fourier series fit to the transformed-time interval map, including only the low-frequency components that meet a significance criterion; **B**: the circle map itself, which represents, essentially, a fit to the *successor map* in which each spike time is related to the subsequent spike time. Also shown, represented by the dotted lines, is the *cobwebbing* procedure used to find the trajectory of any point and its successors on the circle; **C**: circle map and two sample iterates (the second and fifth) of the circle map. The dotted lines again depict the cobwebbing procedure that is used to find the successors to the value of any given starting point, represented here by the *, and any given iterate of the circle map; **D**: fixed point iterate of the circle map. In this case, the responses have ten spikes per cycle, so the fixed point iterate is the tenth. Stable fixed points correspond to times at which the iterated circle map crosses the diagonal with slope less than one, and unstable fixed points correspond to crossings with slope greater than one; **E**: positions of the stable and unstable fixed points in real time. The stable fixed points overlap the peaks in the PSTH, whereas the unstable fixed points occur near the troughs.

crosses the diagonal with slope greater than one, the fixed point is unstable. If the slope is less than one, the fixed point is stable. Intuitively, a circle map derived from spike data should have stable fixed points that correspond to the underlying noise-free firing times for the NLIF model. Simulated SMRP spike trains do not have stable fixed points, since in transformed time only the DC (unmodulated) component of interval maps are significant. This makes the derived circle map topologically equivalent to a rigid rotation (Knight, 1972a).

The dotted line in Fig. 3.10B is a graphical representation of the cobwebbing procedure. This line follows the positions of the successor times of a randomly-chosen starting point. At each step, the successor time is given by the circle map; graphically, the value on the vertical axis becomes the next value on the horizontal axis. The step (in this case, the tenth) that takes the original point as close as possible to its starting value—that is, the step that takes the point approximately 360° around the circle—is of particular interest. Following the trajectories of all starting points, not only the entire circle map, but also its iterates $g^m(\theta)$, are obtained. Two of these iterates are shown together with the original circle map in Fig. 3.10C. Notice that the iteration procedure accentuates small ripples in the circle map, so that successive iterates are more staircase-like.

Since, for the NLIF data shown in Fig. 3.10, all of the selected responses have ten spikes per cycle, the tenth iterate of the circle map is examined for fixed points (Fig. 3.10D). The number of stable (and hence unstable) iterate fixed points is exactly equal to the number of iterates required to find them, if they exist. The fixed points are extracted, and the inverse of the time-stretching transformation is applied (Fig. 3.10E). Thus, the fixed points are converted from transformed time to real time, which facilitates the comparison with features of the PSTH.

A test of the reliability and robustness of the circle map method is shown in Fig. 3.11. The stable fixed points are plotted for different amounts of input shot noise in the NLIF model. Fig. 3.11A shows the fixed points at an intermediate contrast (0.32) that is nonetheless high enough to cause single-cycle phase locking in the absence of noise. When the noise is sufficiently small, fixed points are easily uncovered, and their times

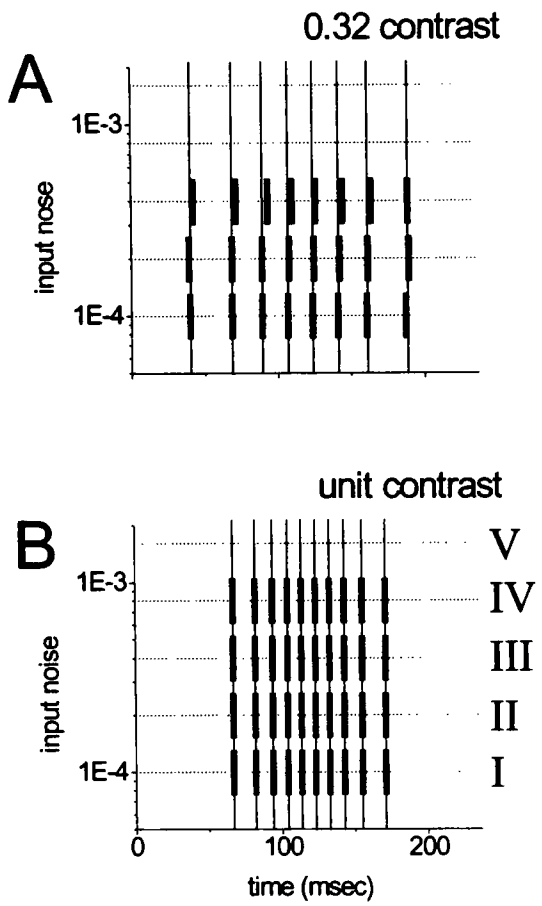

Figure 3.11. *Stable fixed points for the NLIF model.* The addition of significant amounts of noise does not have a large effect on the calculated times of the stable fixed points. Dotted horizontal lines are plotted at each of the noise levels tested. Thin vertical lines represent spike times in the absence of noise. **A**: 0.32 depth of modulation; **B**: unit depth of modulation. Roman numerals refer to the amount of shot noise (I: 0.0001; II: 0.0002; III: 0.0004; IV: 0.0008; V: 0.0016). The PSTHs for noise levels I, III, and V are shown in Fig. 3.3B. Even at high levels of input noise (row III for 0.32 contrast, and row IV for unit contrast), when the PSTH does not have clear peaks, the circle map still has fixed points that are in excellent agreement with the noise-free spike times. However, when the noise is too high (V), the circle maps have no fixed points.

and number are similar to the times of the spikes in the noise-free case (represented by thin vertical lines). This is true even at a noise level (0.0004) for which the PSTH peaks are no longer evident. However, the fixed points are eventually lost when the noise becomes large enough to overwhelm the effects of the underlying deterministic process. At unit contrast (Fig. 3.11B), the fixed points, like the PSTH peaks, are maintained more reliably than at low contrasts, and they only disappear at the highest noise level tested.

Circle maps derived from real neuronal data also have fixed points under comparable conditions. Fig. 3.12 shows the example of an X-type, *on*-center retinal ganglion cell. At contrast 0.16 (Fig. 3.12A), the circle map has no fixed points, even though the power ratio falls outside the Poisson range. This is consistent with the fact that the leaky integrate-and-fire model does not phase-lock to a single cycle of the stimulus when the modulation depth is low. At very high contrasts (Fig. 3.12C), however, where there are evident peaks in the PSTH, the circle map typically has iterate fixed points. The number of stable (or unstable) fixed points is equal to the number of spikes in each response. The stable fixed points fall near the centers of the PSTH peaks, and the unstable fixed points fall near the troughs. At an intermediate contrast (Fig. 3.12B), where the PSTH is clearly

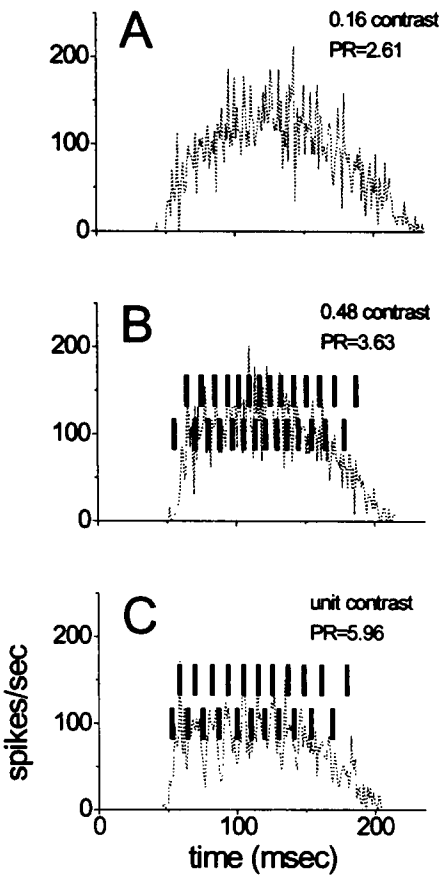

Figure 3.12. *Circle map fixed points at three noise levels for an X-type, on-center retinal ganglion cell.* The stable fixed points are plotted above the unstable fixed points, and they are superimposed on the PSTH of the response. **A**: 0.16 contrast, no fixed points even though the spike train could not have been generated by a simply modulated SMRP; **B**: 0.48 contrast, fixed points even without PSTH peaks; **C**: unit contrast, stable fixed points that correspond to the centers of the PSTH peaks, and unstable fixed points that correspond to the troughs.

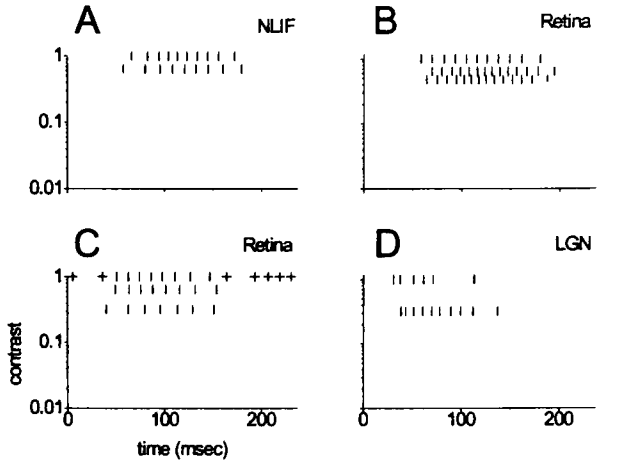

Figure 3.13. *Stable fixed points for different stimulus contrasts.* **A**: NLIF model (input noise 0.0008); **B**: X-type, *on*-center retinal ganglion cell; **C**: X-type, *on*-center retinal ganglion cell. Two different experimental conditions are superimposed: (1) only the central portion of the receptive field is stimulated (|); (2) only the receptive field surround is stimulated (+); **D**: X-type, *on*-center LGN cell.

modulated but peaks are not evident, the corresponding circle map nevertheless has fixed points.

Fig. 3.13 shows the stable fixed points for the NLIF model and three cat cells (two X-type, *on*-center retinal ganglion cells and one X-type, *on*-center LGN cell) at multiple stimulus contrasts. In Fig. 3.13C, two experiments on the same retinal ganglion cell are superimposed. In the first experiment (|), the stimulus covers only the central region of the receptive field; in the second experiment (+), the stimulus covers only the surround. Fixed points appear at lower contrast for the center stimulation, which indicates that stimulation of the center of the receptive field is likely to elicit more precise spike firing.

Discussion

The interval map/power ratio method effectively distinguishes spike trains generated by two broad classes of models. The first class, simply modulated renewal processes (SMRPs), includes modulated Poisson and gamma processes. The spike-generating mechanisms in this class of models are characterized by the simple manner in which they are affected by an external stimulus: the stimulus changes the firing rate, or, equivalently, modulates the running speed of an internal clock. The time transformation is the unique map that regularizes the clock and thus transforms spike trains generated by these models into renewal processes.

The second class of models is characterized by underlying spike-generating mechanisms for which the effect of an external stimulus is not equivalent to a modulation of the internal clock. Such models are not SMRPs because they contain parameters that are measured in units of real time and that do not covary with the stimulus. The effects of these real-time parameters survive the time transformation. The NLIF model falls into this class because it resets after each spike is fired and because it is *leaky*. These features are reflected in the clusters of points at specific locations in the transformed-time interval maps of its spike trains (Fig. 3.2B). Models with refractory periods that are fixed in real time (Berry and Meister, 1998) also fall into this class because the time transformation distorts the refractory period, changing the duration of the refractory period non-uniformly through the response. This distortion induces a modulation in the transformed-time interval maps. Modulations that are large enough, such as those that occur when the refractory period is long, are reflected in the power ratio. The simulations in Fig. 3.6 reveal that the refractory period needs to be quite long—unphysiologically long—in order to evoke a power ratio outside the SMRP range.

The power-ratio statistic is designed to be sensitive to a particular kind of structure in the transformed-time interval maps—namely, deviations from the uniformity expected of an SMRP. Many features of the interval map that could distinguish among different SMRPs do not affect the power ratio at all. For example, modulated gamma processes of different orders have different ISIHs, which are characterized by different means

and variances. Their transformed-time interval maps are all uniform throughout the stimulus cycle, but the shape of the local interval distributions depends on the order of the gamma process.

Certain spike trains generated by non-SMRP models have power ratios in the SMRP range. Examples include NLIF spike trains evoked by low-contrast stimuli as well as spike trains generated by modulated renewal processes with fixed refractory periods in the physiological range. Furthermore, the power ratio is insensitive to serial correlations provided that those correlations are stimulus-independent. It is likely that indices other than the power ratio could distinguish these spike trains from SMRP spike trains. Thus, the power ratio test is a conservative one.

Despite the lack of sensitivity of the power ratio, a surprisingly large fraction of spike trains at all three recording sites are inconsistent with SMRPs. The fraction of neurons at each recording site that have non-SMRP underlying spike generating mechanisms is also surprisingly large, and it is even larger when only high-contrast responses are considered. These results suggest that SMRPs are, in general, poor models for neurons in all three brain areas, especially if the underlying spike-generating mechanisms are relatively similar across neurons.

For cat X-type retinal ganglion cells in particular, several findings suggest that the NLIF model provides a useful reduced description of the spike-generating mechanism. First, such a model *phase locks* in response to sinusoidal input that is sufficiently strongly modulated (Knight, 1972a), just as real retinal ganglion cells do (Reich et al., 1997). Second, when the stimulus modulation depth is sufficiently high, even in the presence of significant noise, evidence of the phase locking can still be seen in the PSTH; this, too, is true for real retinal ganglion cells (Reich et al., 1997). Third, as the contrast increases, both model and real responses undergo a gradual transition from firing spike trains that are consistent with SMRPs to firing spike trains that are not. Fourth, the transformed-time interval maps of both real and model responses contain prominent resetting clusters. Finally, circle maps, which can uncover fixed points from noisy model data that correspond

to phase-locked spikes in noise-free model data, can also uncover similarly behaved fixed points from real retinal ganglion cell responses.

Thus, the NLIF model provides a single explanation for many of the salient features of retinal ganglion cell spike trains, including details of their temporal behavior. If multiple responses of a single NLIF neuron are considered to be interchangeable with individual responses of multiple, parallel NLIF neurons (Knight, 1972a), then the results may provide an explanation for the response synchronization that has been seen across multiple retinal ganglion cells (Meister et al., 1995). It should be noted, however, that the NLIF model provides a highly simplified description of retinal ganglion cell spike generation, and no explicit attempt has been made to fit the model to any particular retinal ganglion cell. Furthermore, the model does not contain many of the features of full-fledged models, such as realistic ion channels, and it does not account for the serial correlations between consecutive interspike intervals that are a well-described feature of unmodulated retinal ganglion cell spike trains (FitzHugh, 1958; Levine, 1991; Troy and Robson, 1992).

Because the data reported here come from three successive stages of visual processing, the results also provide some insight into changes in the temporal properties of spike trains as information is transmitted through the visual system. This may be of some value in addressing the yet-unknown mechanisms of cortical information processing. The data presented in this chapter suggest that retinal ganglion, LGN relay, and V1 neurons contain intrinsic temporal structure that is a consequence of their distinctive spike-generating dynamics.

Chapter 4

Temporal Coding of Contrast: When, What, and Why

Summary

The results of Chapter 3 reveal that the dynamical structure of the responses of many neurons in V1 can support the encoding of stimulus-related information in spike times, in addition to the overall firing rate. In this chapter, the information that V1 responses convey about the contrast of static visual stimuli is explicitly calculated. These responses often contain several easily distinguished temporal components, which will be called *latency*, *transient*, *tonic*, and *off*. Calculating the information about contrast conveyed in each component and in groups of components makes it possible to delineate aspects of the temporal structure that may be relevant for contrast encoding. The results indicate that as much or more contrast-related information is encoded into the temporal structure of spike train responses as into the firing rate, and that the temporally coded information is manifested most strongly in the latency to response onset. Transient, tonic, and off responses contribute relatively little. The results also reveal that temporal coding is important for distinguishing subtle contrast differences, whereas firing rates are useful for gross discrimination. This suggests that the temporal structure of neurons' responses may extend the dynamic range for contrast encoding in the primate visual system.

Introduction

Stimulus contrast offers several advantages as a paradigm for studying the ways in which information is encoded into the responses of visual neurons. Contrast encoding

is highly nonlinear: the firing rate of V1 neurons tends to vary with contrast in a sigmoidal fashion (Albrecht and Hamilton, 1982). As with retinal ganglion cells (Shapley and Victor, 1978) and lateral geniculate nucleus neurons (Sclar, 1987), the responses of V1 neurons exhibit prominent contrast gain control (Ohzawa et al., 1982; Bonds, 1991) that may be modeled as a divisive inhibitory process (Heeger, 1992). Moreover, in the case of stationary stimuli, which serve as useful substrates for the analysis of temporal coding (Victor and Purpura, 1996), variation of stimulus contrast does not necessarily entail variation of spatial phase, whereas variation of other stimulus parameters, such as orientation and spatial frequency, does.

In the past, both moving and stationary stimuli have been used to study contrast encoding. Favorite stimuli have included sinusoidal gratings, which may be either drifted uniformly or flashed briefly for a specified period of time. Typically, responses are characterized by average measures, such as the mean firing rate (especially for complex cells) and the fundamental Fourier component (especially for simple cells, when the stimulus is periodic) (Skottun et al., 1991). However, recent studies (Gawne et al., 1996b; Victor and Purpura, 1996; Mechler et al., 1998b) have indicated that such measures may ignore an important part of the information about contrast that is encoded in the temporal structure of neurons' responses—that is, in the detailed timing of action potentials relative to the stimulus time course. Such *temporal coding* of contrast is more prominent in responses that have transient components, such as those elicited by drifting edges, than in responses to narrowband stimuli, such as drifting sinusoidal gratings (Mechler et al., 1998b). Moreover, much of the information about contrast is encoded into a single response variable—the latency from stimulus onset to neuronal firing—that can in some cases vary independently of the overall firing rate as the spatial structure of the stimulus changes (Gawne et al., 1996b).

This chapter presents the results of a systematic study of the responses of V1 neurons to transiently presented sinusoidal gratings that vary in contrast. The goal is to characterize aspects of the responses that are relevant for contrast representation and to determine whether temporal coding plays some specific, identifiable role. A metric-space approach is used to estimate information about stimulus contrast (Victor and Purpura,

1996). The full response is analyzed, as are its various temporal components, including the latency, the initial transient period of high firing rate, and the longer period of tonic firing that lasts until the stimulus is turned off. The results indicate that different response components can convey both independent and redundant information about contrast. The fraction of information encoded into the temporal structure, as opposed to the firing rate, can vary from component to component within the same response. Taken together, the results lead to a hypothesis for the role played by the encoding of information into the temporal structure of neuronal responses—namely, that temporal coding allows the visual system to distinguish among stimuli that evoke similar firing rates.

Methods

Stimuli. As in the previous chapter, the data presented here represent the activity of single neurons with parafoveal receptive fields in the primary visual cortices of sufentanil-anesthetized macaque monkeys. Stimuli consist of transiently presented, stationary sinusoidal gratings that have fixed orientation, spatial frequency, and spatial phase but that vary in contrast. For each neuron encountered in the experiments, the orientation and spatial frequency of the drifting sinusoidal grating that maximizes either the firing rate, for complex cells, or the response modulation at the driving frequency, for simple cells, are determined. For groups of neurons containing more than one well-isolated individual neuron, the orientation and spatial frequency that are optimal for the best-isolated or most robustly responding neuron are used; experience and quantitative studies (DeAngelis et al., 1999) suggest that the optimal values of orientation and spatial frequency vary relatively little among nearby neurons.

The third parameter of the stimuli—spatial phase—is more difficult to choose. Spatial phase preference can vary dramatically from one neuron to its neighbor, especially among simple cells (DeAngelis et al., 1999). Although a spatial-phase tuning experiment, which uses stationary sinusoidal gratings, is performed for each neuron or group of neurons encountered, and although the spatial phase that evokes the largest firing rate in one or more neurons is selected, it is impossible to be sure that the chosen spa-

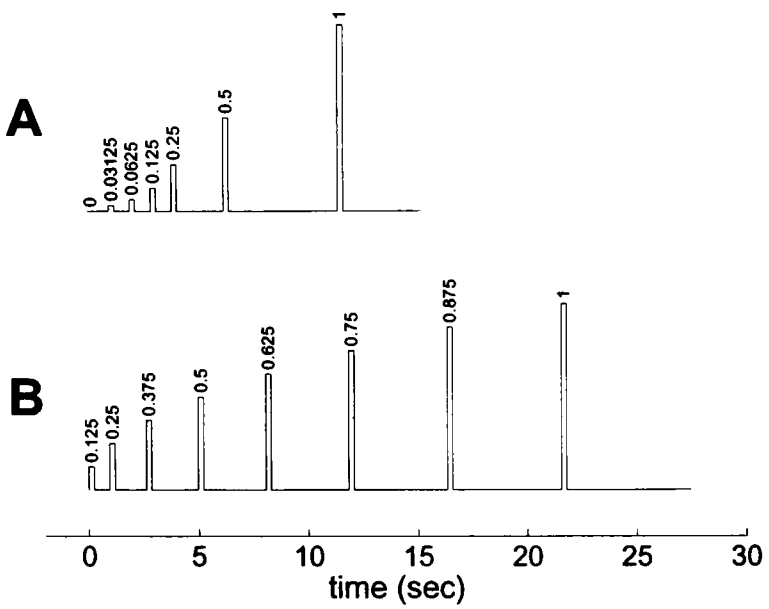

Figure 4.1. *Timeline of contrast experiments.* **A**: geometric series of contrasts. **B**: arithmetic series of contrasts. Grating stimuli are presented for 0.237 sec each (pulses) and are then replaced with a uniform field at the same mean luminance (effectively, contrast 0). The duration of the uniform-field presentation following each grating presentation increases with contrast, as shown in the figure, to approximate a uniform state of contrast adaptation. In the figure, pulse heights are proportional to contrast, and the contrast value is listed above each pulse.

tial phase is actually "optimal" in any sense. This is true even for simple cells, which can be exquisitely sensitive to spatial phase (Movshon et al., 1978b; Victor and Purpura, 1998): the spatial phase that evokes the largest response may not, for example, evoke the most reliable responses.

After fixing the orientation, spatial frequency, and spatial phase, one of two possible sets of stationary-grating stimuli is presented. The first set consists of a geometric series of six contrasts, and the second set of an arithmetic series of eight contrasts (Fig. 4.1). In both sets, gratings replace a uniform field (4.5° x 4.5°) of the same mean luminance (150 cd/m^2) for a period of 237 msec, after which the uniform field reappears for a minimum of 710 msec. The amount of time between grating presentations increases as a function of the contrast of the preceding grating. For example, the amount of time following the 0.5 contrast presentation is 2.84 sec, and following the 0.875 contrast presentation, 4.26 sec. This strategy is used to approximate a uniform state of contrast adaptation (Sclar et al., 1989). The entire series of contrasts is typically presented 100 times. For each trial, the spikes that occur in the first 350 msec after stimulus onset are analyzed. Also analyzed are multiple 947 msec periods of uniform-field stimulation.

Information estimation. Information theory provides a method of measuring the fidelity with which responses to similar stimuli form distinct clusters in some response

space. The information-theoretic measures calculated here are sensitive to both the number of spikes in a response (the firing rate) and the timing of those spikes. The method of estimating information involves embedding neuronal responses into metric spaces rather than Euclidean vector spaces, which tend to be sparsely populated (Victor and Purpura, 1996; Victor and Purpura, 1997). Pairwise distances between individual spike trains are calculated under the spike-time metric (Victor and Purpura, 1996), which computes the shortest path by which one spike train can be converted into another through elementary steps that include adding and deleting spikes as well as shifting spikes in time. The analysis depends on the value of a parameter, called q, which represents the cost per unit time of moving the occurrence time of a spike during the conversion of one spike train into another. When q=0 sec^{-1}, the distance is the difference in number of spikes between the two trains. At very large values of q, the distance approaches the sum of the number of spikes that do not fall at identical times in the two trains. At intermediate values of q, the distance lies between those two extremes.

The mutual information H, calculated in bits, is a measure of the degree to which responses to the same stimulus are more similar to each other than to other responses. Clustering of responses into stimulus classes—the prerequisite for the information calculation—is described in (Victor and Purpura, 1997). Each response is considered in turn, and the median distances to the responses in each stimulus class are calculated. The response under consideration is assigned to the cluster associated with the shortest median distance. Here, the median distance, rather than the generalized mean distance as in (Victor and Purpura, 1997), is used as the basis for the clustering because simulations show that the median behaves more robustly for responses with small numbers of spikes (see Appendix). When the number of spikes is large, the details of the clustering matter less.

For N equally probable stimuli, the maximum possible information is $\log_2 N$. As a first step in comparing multiple data sets, information values are normalized by the appropriate maximum value. Within each data set, a bias-corrected mutual information H is calculated as a function of q, the cost parameter. From the plots of H vs. q, several parameters are extracted. H_0, the mutual information at q=0 sec^{-1}, represents the amount of

information contained in the spike count or firing rate and is calculated directly. The remaining parameters are extracted from a fit to the information curve:

$$H_{\text{fit}}(q) = k\left(\frac{1+Aq^{c}}{1+Bq^{b}}\right). \tag{4.1}$$

This function is chosen empirically because it gives good fits, not because its parameters are likely to have any physiological relevance. Fitting and parameter extraction are done with the interior-reflective Newton method as implemented by the routine LSQCURVEFIT in the Optimization Toolbox of Matlab 5.3.1 (The Mathworks, Natick, MA).

The following parameters are extracted from the fits. $\boldsymbol{H_{\text{peak}}}$ is the mutual information at the peak of the curve; if it occurs at $q_{\text{peak}}>0$ sec^{-1}, there is more information in the temporal structure of the response than in the spike count alone, and the extra information is given by H_{peak}-H_0. The informative temporal precision limit (in msec) is $2000/q_{\text{cut}}$, where $\boldsymbol{q_{\text{cut}}}$ is the value at which $H_{\text{fit}}(q)=H_{\text{peak}}/2$.[3] Finally, an index of temporal coding $\boldsymbol{\Theta}$, the percentage of stimulus-related information that is carried in the temporal structure of the neuron's response at q_{peak}, is given by $\Theta=100(H_{\text{peak}}-H_0)/H_{\text{peak}}$.

Bias in the information calculation. Due to small sample sizes, mutual information is likely to be overestimated (Miller, 1955; Treves and Panzeri, 1995). To correct this bias, a resampling technique is applied to estimate the information that would be obtained from an equivalently sized set of responses with no stimulus dependence (Victor and Purpura, 1996). The resampling is implemented by estimating the information in ten random associations of responses with stimuli. Simulations (Victor and Purpura, 1997) reveal that this resampling procedure tends to overcorrect the mutual information estimates, which are therefore likely to be conservative. The total amount of overcorrection is expected to be small and independent of the cost q (see Appendix).

---

[3] Generally, q_{cut} is a more reliable index of temporal precision than q_{peak}, because the information curves often have no sharp peak value; see Appendix. In the expression $2000/q_{\text{cut}}$, a factor of 1000 comes from the conversion of sec to msec, and a factor of two from the fact that the spike-time distance has a natural time scale of $2/q$, which is the maximum separation of a pair of spikes that are considered to have similar times (Victor and Purpura, 1996).

In addition to the bias, the information estimates themselves are random variables and therefore have some uncertainty. This uncertainty is estimated for the data in Figs. 4.3 and 4.5 by the bootstrap method (Efron and Tibshirani, 1998). Specifically, for each data set, 100 resamplings are made, in which the spike trains are drawn from the original data set, with replacement and separately for each stimulus condition. The bootstrap estimate of standard error is:

$$SE_{\text{boot}} = \sqrt{\frac{1}{B-1}\sum_{b=1}^{B}\left[s(\mathbf{x}^{*b}) - s(\cdot)\right]^2} \tag{4.2}$$

where B is the number of resamplings, $s(\mathbf{x}^{*b})$ is the b^{th} resampling, and $s(\cdot)$ is the mean of the resamplings, or $\sum_{b=1}^{B} s(\mathbf{x}^{*b})/B$. Error bars in Figs 4.3 and 4.5 are *bias-corrected root-mean-square errors*, obtained by combining the bias and the bootstrap standard error as follows (Efron and Tibshirani, 1998):

$$SE_{\text{bias-corrected}} = \sqrt{SE_{\text{boot}}{}^2 + \text{bias}^2}\,. \tag{4.3}$$

Note that the bootstrap procedure itself can provide an estimate for the bias (Efron and Tibshirani, 1998). The bias correction based on random association of stimuli and responses is preferred here because its properties in the context of metric-space information calculations have been extensively investigated (Victor and Purpura, 1997). For the data presented here, the two bias estimates are of the same order of magnitude.

Latency. For responses to stationary gratings, the *onset latency* (Sestokas and Lehmkuhle, 1986) is determined by a method similar to that of (Maunsell and Gibson, 1992). This method identifies the earliest time, for each stimulus, that visual information reaches the neuron under study. Other methods of finding the latency (Levick, 1973; Lennie, 1981; Bolz et al., 1982) are designed for different purposes, such as determining the peak of neuronal activation following stimulus onset. In the present method, the background spike-count distribution is estimated by dividing the response to the uniform field (zero contrast) into 1 msec bins and tabulating the observed spike counts in those bins across multiple repeats of the stimulus. For the response to each non-zero contrast, the latency is taken to be the first bin in which the number of spikes is significantly higher, in that bin and the three subsequent ones, than the background spike count. Sig-

nificance is determined in a non-parametric fashion by directly comparing the observed spike counts to the distribution of background spike counts, and by requiring that the observed spike count be in the top 20% of the background spike counts in each bin. This gives a significance level of $0.0016=(0.2)^4$ over four consecutive bins, assuming independence. In a few cases, robust latency values could not be obtained with this significance criterion for low-contrast responses, and the cutoffs had to be relaxed to 30% ($p<0.0081$) or 40% ($p<0.0256$). In other cases, a 10% cutoff ($p<0.0001$) could be used.

A similar method is adopted to find the boundary between the transient (phasic) and tonic (sustained) portions of the responses to stationary gratings (see Fig. 4.2). For this purpose, the estimate of baseline activity is taken to be a section of 100 msec of each response that is identified by eye to be part of the tonic response. From the beginning of the identified section, a backward search proceeds, bin by bin, until four consecutive bins are found in which the spike counts are significantly greater than the baseline spike counts. The last of these bins (the first one encountered in the backward search) is chosen as the boundary point. Since off responses are often quite small and difficult to delineate, they are uniformly considered to begin 237 msec (the duration of each grating stimulus) after the response onset, and to have the same duration as the transient response. This choice corresponds to the assumption that the latency to the on (transient) response is exactly as long as the latency to the off response.

Results

Fig. 4.2A shows the responses of a simple cell in macaque V1 to a series of stationary sinusoidal gratings presented at an arithmetic series of eight contrasts. The grating stimulus appears for 237 msec and is then replaced by a uniform field at the same mean luminance. Responses are presented as post-stimulus time histograms (PSTHs), binned at 1 msec resolution. PSTHs represent the average firing rate at all times after the onset of the visual stimulus, which occurs at time zero. Stimuli are each presented 100 times.

Despite the simple appearance-disappearance time course of the stimulus, the PSTH has a complicated temporal waveform (Ikeda and Wright, 1975; Movshon et al., 1978b). At least four distinct components of the response can be discerned. The division of the unit-contrast PSTH into these four temporal components—denoted *latency*, *transient*, *tonic*, and *off*—is shown in Fig. 4.2B. Boundaries between the components are chosen as described in Methods. Without seeking to determine the biophysical and physiological mechanisms underlying the distinctions among response components, or even whether they are generated by discrete mechanisms in the first place, the analysis that follows examines the degree to which each temporal component encodes the contrast of the visual stimulus.

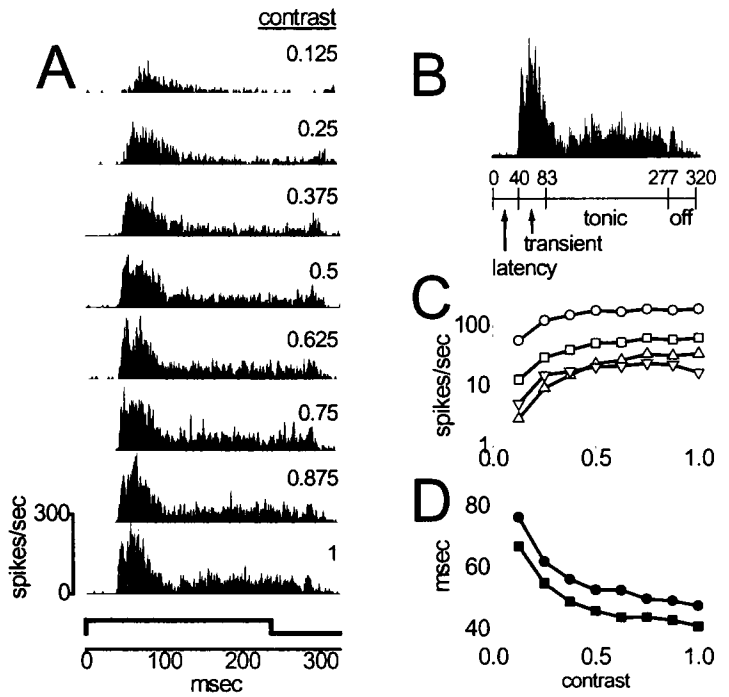

Figure 4.2. *Response of a simple cell to stationary sinusoidal gratings.* **A**: Post-stimulus time histograms (1 msec bins) of the response of a V1 simple cell (35/1s) to sinusoidal gratings presented 100 times at eight contrasts. The stimuli appear at time zero and are turned off at 237 msec. Response onset is abrupt, with a latency that decreases as a function of contrast. **B**: Typical responses can be divided into four components: the onset latency, the transient period of high firing rate, the elevated tonic firing that follows the transient and persists until the stimulus is turned off, and the off response itself. The procedure for assigning boundaries between response components is described in Methods. **C**: Firing rate vs. contrast for each response component: full response (□), transient (○), tonic (△), and off (▽). **D**: onset latency (■) and median first post-latency spike time (●) as a function of contrast.

The latency (Maunsell and Gibson, 1992) is defined as the amount of time between stimulus onset and the beginning of the neural response. In V1 neurons, its duration reflects, at the very least, the time required for a response to be evoked in the photoreceptors and for the neural signal generated in the photoreceptors to pass through the various retinal cell layers and the lateral geniculate nucleus. Latency in V1 neurons de-

creases as the contrast of the visual stimulus increases (Sestokas and Lehmkuhle, 1986; Gawne et al., 1996b), a phenomenon that is related to the temporal phase advance of responses to drifting gratings with increasing contrast, in both the retina (Shapley and Victor, 1978) and cortex (Dean and Tolhurst, 1986; Albrecht, 1995). The decrease in latency can be appreciated by scanning down the column of PSTHs in Fig. 4.2A and observing that the onset of the response becomes progressively earlier as contrast increases.

The transient portion of the response is the relatively brief period of intense firing that begins when the visual signal first reaches the neuron. The bulk of the stimulus-related information in neuronal responses has already been transmitted by the end of the transient component (Purpura et al., 1993; Heller et al., 1995; Bura as et al., 1998; Müller et al., 1999a). In the responses of the neuron presented in Fig. 4.2, the firing rate increases quickly after the response onset and remains high for 40 to 50 msec before declining to a tonic level that is relatively sustained until after the stimulus has been turned off. The decline in firing rate has been thought to reflect a process of short-term adaptation, perhaps involving synaptic depression (Chance et al., 1998; Müller et al., 1999b). After the tonic response ends, at around 277 msec, a brief off response appears. The off response is much smaller than the transient response, even though the change in contrast is identical (0 to 1 or vice versa). In this simple cell, the relative size of the transient and off responses is largely a function of the spatial phase of the stimulus (not shown). This neuron's response is similar to that of the "nonlinear simple cell" recorded from cat V1 and depicted in Fig. 5 of (Movshon et al., 1978b).

Of the 50 neurons analyzed here, only 20 had distinct transient and tonic response components that were easily separable by eye and by the boundary-search method (see Methods). The other neurons had responses that decayed slowly over time or else remained constant until after the grating was removed. Thus, conclusions about the contrast-encoding properties of the response latency are based on data from 50 neurons, whereas conclusions about the contrast-encoding properties of the transient, tonic, and off responses are based on data from 20 neurons.

Fig. 4.2C and D show different scalar measures of the contrast responses, plotted against contrast. In Fig. 4.2C, the firing rate is plotted on a logarithmic scale separately for each temporal response component. The firing rate for the full response (□) has a dynamic range of about 40 spikes/sec, but most of that range is evoked by contrasts of 0.5 and lower. Above this contrast, the firing rate saturates, making it very difficult to distinguish high-contrast stimuli on the basis of firing rate alone. This type of saturation is a common feature of many V1 neurons' contrast response functions (Maffei and Fiorentini, 1973; Tolhurst et al., 1981; Albrecht and Hamilton, 1982; Ahmed et al., 1997), and it is prominent in all temporal components of the response.

Fig. 4.2D shows the dependence on stimulus contrast of latency (■) and median first post-latency spike time (●). The latency is a PSTH-based measure of the earliest time that the response rises above the baseline firing rate. Like the firing rate, both of these response measures change rapidly at low contrast and less rapidly at high contrast, though the degree of saturation is arguably less for latency and first spike time than for firing rate. The decrease in response latency as a function of contrast has been proposed to be a primary way in which visual neurons represent contrast (Cleland and Enroth-Cugell, 1970; Bolz et al., 1982; Gawne et al., 1996b; Wiener et al., 1999).

The relative contribution of each temporal response component to the encoding of contrast is assessed by comparing the information conveyed by subset spike trains that consist only of spikes within a particular response component, with average latency information either left intact or removed. In the following sections, *full response* refers to spikes that occur between the onset latency and a cutoff time 237 msec later. For responses that include a clear transient component (20 of 50 neurons), the full response is extended by the duration of the transient so as to include the off response, regardless of its actual size or duration. Latency is determined independently for each contrast but assumes a fixed value for all spike trains recorded at that contrast, since it is a measure derived from the average response at each contrast. Only contrasts that evoke a clear response onset are considered, so that the full set of contrasts is not necessarily analyzed for each neuron. For each neuron, approximately 100 responses are recorded at each contrast.

Latency. Latency and firing rate, *a priori*, are independent response measures. As shown in Fig. 4.2, both measures covary with contrast. Across all 50 neurons in the sample, stimulus contrast is correlated with both firing rate (median Pearson's $R = 0.94$) and latency ($R = -0.78$), and, consequently, the correlation between firing rate and latency is high ($R = -0.86$). The degree of correlation is nearly twice as high as what is found in cat retinal ganglion cells for stimuli that vary in spatial position within the receptive field (Levick, 1973).

To evaluate the separate contributions of firing rate and response latency to the coding of contrast information in V1, it is useful to compare the information transmitted by the full response with the information transmitted by two derived responses. The two derived responses are complementary: one contains only the latency information, and the other removes latency information entirely. This means that if the sum of the contrast-related information in the derived responses exceeds the contrast-related information in the full response, the two derived responses can be said to convey redundant information. Alternatively, if the derived responses convey independent information, their information curves are expected to sum to the information curve of the full response. If one of the derived responses conveys more information than the full response, then the additional features in the full response can be said to provide confusing information about contrast (although in that case they may provide information about other stimulus features, such as spatial phase).

The first derived response is obtained by subtracting the contrast-specific (but not trial-specific) latency from all spike times recorded at each contrast. This preserves the relative spike times within and across trials at a single contrast but removes the overall latency shift across contrasts. The resulting derived response, which contains the same number of spikes as the original response, as well as the same interspike intervals, is used to evaluate the amount of contrast-related information that is not contained in the latency. This information could be carried by spike counts and by aspects of temporal pattern other than the time of the first spike (for example, the time of the second spike or the occurrence of "bursts"). The second derived response is obtained by selecting only the first post-latency spike in each trial of the full response. Trials in which no spikes are fired are

ignored, so that each trial in the derived response has exactly one spike. This removes the confounding effect of differences in spike count, since spike-free trials are more likely to occur at lower contrasts. The result is a derived response that is used to evaluate the amount of contrast-related information encoded specifically into the response latency.

Fig. 4.3 shows the results of applying the spike metric method to the full and derived responses for two separate neurons. For each cost q, information is expressed as a percentage of the maximum information that would have been obtained from the set of responses if the responses to different contrasts were perfectly distinguishable. This maximum value, in bits, is $\log_2 N$, where N is the number of stimulus conditions. All information estimates (both actual and normalized) are corrected for the small-sample bias by subtracting the information expected from chance clustering. Individual points are fit to an empirical five-parameter curve as described in Methods. The parameters H_{peak} (peak information, possibly equal to the spike-count information H_0), Θ (temporal coding index), and q_{cut} (temporal precision limit) are extracted from these fits. H_0 itself is estimated directly.

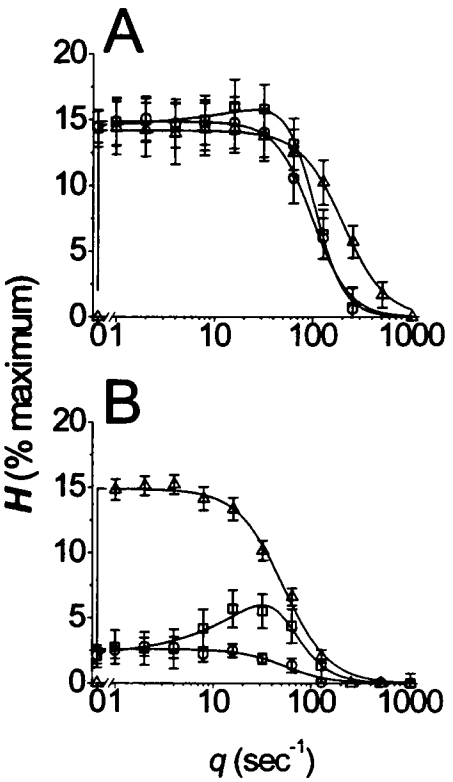

Figure 4.3. *Information transmitted in the onset latency.* Contrast-related information in the full response (□), in the response with latency removed (○), and in the first spike alone (△). Transmitted information is evaluated by the spike-metric method. **A**: simple cell from Fig. 4.2 (35/1s). **B**: complex cell (38/2s). Solid lines are empirical fits to the data points (see Methods). Information values are represented as a percentage of the maximum possible information ($\log_2 N$, where N is the number of stimuli), after an estimated bias due to limited sample size is subtracted (see Methods). Error bars represent one standard error of the mean and are derived from a bootstrap resampling procedure (see Methods).

Fig. 4.3A shows the information curves for the simple cell of Fig. 4.2. Essentially all of the information in the full response (□) is encoded in the spike count: H_0 is 14% ± 1.4% (standard error of the mean derived from 100 bootstrap resamplings), whereas

H_{peak} is 16% (derived from a fit to the information curve; see Methods). Although approximately Θ = 8.5% of the contrast-related information is transmitted by a temporal code, this value is not significantly different from zero. The same is true in the derived response with latency information removed (O; Θ = 2.5%, not significantly different from zero). For this derived response, H_0 is by construction equal to H_0 of the full response (up to discrepancies in the estimate of information bias; see Methods), since the number of spikes in each trial of the full and latency-free responses is the same. Remarkably, the information in the first spike alone ($\triangle$) is as high as the information in the full response (H_{peak} = 15%), which means that the information contained in the time of the first spike is redundant with the information contained in the spike count. For the first spike alone, H_0 is 0% by construction: all trials in the derived response have exactly one spike, so that no information at all is transmitted by the spike count.

In addition to showing the amount of contrast-related information transmitted by the first spike compared to the full response, the information curves in Fig. 4.3A provide insight into the *informative temporal precision* of spikes in the full and derived responses. As discussed in the Appendix, this is a measure of the precision with which spike times can be used to distinguish one stimulus from others, but it is not explicitly related to the reliability of spike times across trials for particular stimuli. One measure of the temporal precision limit is q_{cut}, the value of q at which the fitted information curve $H_{fit}(q)$ reaches $H_{peak}/2$. For the full response, q_{cut} is 110 sec^{-1}, giving an informative temporal precision limit of $2000/q_{cut}$ = 18 msec. The temporal precision limit of the response with latency removed is similar (20 msec). These values represent a kind of average over the entire response. However, when the analysis is limited to the first spike alone, the temporal precision limit (assuming there is at least one spike in the response) is about twice as fine (10 msec).

Fig. 4.3B shows the corresponding information curves for a complex cell. When the latency information is removed from the response of this neuron, less than half of the contrast-related information remains: H_{peak} of the derived response with latency removed is only 44% of H_{peak} of the full response. Whereas 62% of the information in the full response is temporally coded, the percentage declines to 15% when the latency information

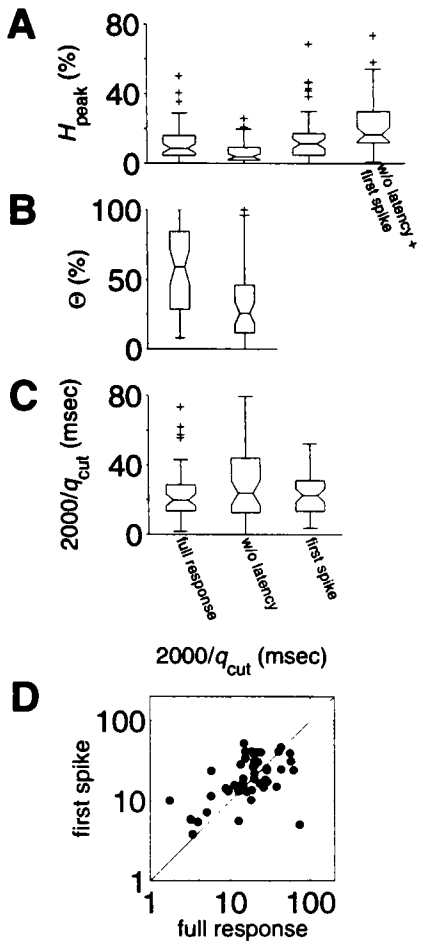

Figure 4.4. *Summary of contrast-related information in the response latency.* Distributions are derived from the responses of 50 V1 neurons to stationary sinusoidal grating stimuli. Information parameters are taken from fits to information curves such as the ones shown in Fig. 4.3. Box plots depict medians (box centers), interquartile ranges (box boundaries), data within 1.5 interquartile range of the medians on both sides (whiskers), and 95% confidence intervals of the medians (notches). Symbols (+): outliers within the displayed vertical-axis range. Since none of the information parameters can take on values less than zero, and since one of them (Θ) is bounded above by 100%, boxes and whiskers may merge at these extreme values. **A**: peak information values (H_{peak}), expressed as a percent of the maximum possible information. **B**: estimated percentage of information that is represented by a temporal code (Θ). Because all *first-spike* responses have exactly one spike, Θ=100% by definition. **C**: informative temporal precision limit, in msec ($2000/q_{cut}$). **D**: informative temporal precision limit of first spike alone vs. informative temporal precision limit of full response for each neuron.

is removed. In this neuron's response, the first spike alone (△) transmits far more information than the full response (□), meaning that later spikes actually impair contrast discrimination. However, the temporal precision limit of the first spike (40 msec) is coarser than the temporal precision limit of the full response (24 msec). Indeed, the responses of this neuron highlight the difficulty of interpreting the temporal precision limits derived from the information curves: even though the temporal precision limit is lower for the first spike than for the full response, the first spike transmits more information about contrast than does the full response at any given temporal precision (q).

The information curves across all 50 neurons are summarized in Fig. 4.4, which uses box plots to represent the distributions of each information parameter for the full response and the two derived responses. As expected (Fig. 4.4A), more contrast-related information is conveyed in the full response (median H_{peak}: 8.6%) than in the response with latency removed (3.9%, $p<0.001$, direct comparison with 1000 bootstrap resamplings).

However, the first spike alone typically conveys more information than the full response (12%, $p<0.001$). Not surprisingly, then, the sum of the information conveyed by the two derived responses (right-most distribution, 16%) is also larger than the information conveyed by the full response ($p<0.001$), indicating that the contrast-related information in the two derived responses is redundant (Gawne et al., 1996b). There are no significant differences between simple ($N=22$) and complex ($N=28$) cells in the median value of H_{peak} for either the full response or the two derived responses ($p>0.05$, direct comparison with 1000 unpaired bootstrap resamplings).

Fig. 4.4B shows that contrast is encoded in the full response primarily through a temporal code: the median temporal coding fraction (Θ) is 61%. Most of this is due to latency variations, and the temporal coding fraction declines to 19% when those variations are removed ($p<0.001$). This confirms that most of the temporally coded information is in the response latency, as indexed by the time of the first spike. Again, there are no significant differences between simple and complex cells in terms of the median value of Θ.

Fig. 4.4C depicts the distribution of informative temporal precision limits ($2000/q_{cut}$) across all neurons. The full responses have a median temporal precision limit of 20 msec, significantly finer than the temporal precision limits of the responses with latency removed (median: 24 msec, $p=0.004$), but not significantly different from the temporal precision limits of the first spikes alone (23 msec, $p=0.06$). Fig. 4.4D shows that the temporal precision limits for the full response and the first spike alone are correlated (Spearman's rank correlation coefficient: 0.51, $p=0.002$, direct comparison with 1000 paired bootstrap resamplings). There are no significant differences between simple and complex cells with respect to q_{cut}.

It is important to reiterate that in the context of the results described here, the informative temporal precision limit gives an estimate of the time differences that are relevant for distinguishing between different contrasts, and not of the reliability of a particular spike time within a response across repeated trials (although the two numbers may be correlated). As discussed above, the temporal precision limit cannot be considered in iso-

lation from the overall information, which is typically significantly higher for the first spike alone than for the full response.

Transient, tonic, and off responses. In the responses of 20 of the 50 neurons, transient, tonic, and off components are clearly delineated. To analyze the contrast-encoding properties of each temporal component separately, derived responses are constructed that consist only of spikes that occur during one of the response components. From each spike time, the starting time of its associated response component is subtracted. For example, the contrast-specific latency is subtracted from all spike times in the transient response at each contrast. This subtraction means that the comparison of response components is not confounded by differences in the onset times of those components; without the subtraction, responses would be extremely easy to distinguish and information values would be spuriously high.

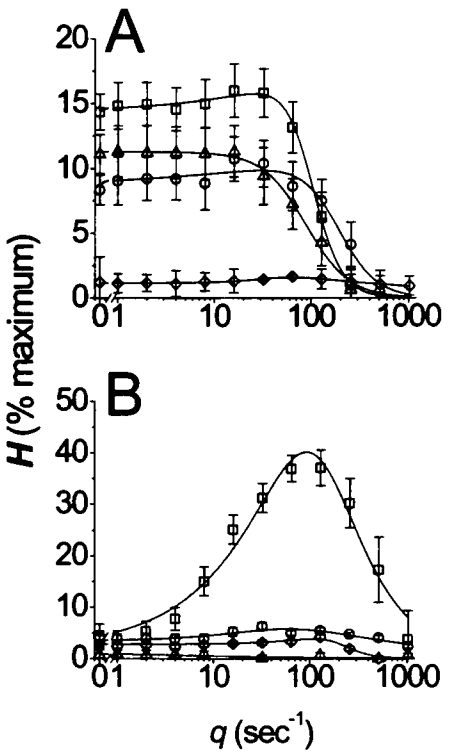

Figure 4.5. *Information curves for full response and response components.* Full response (□), transient (○), tonic (△), and off (◇). Details are as in Fig. 4.3. **A**: simple cell (35/1s), same as Figs. 4.2 and 4.3A. **B**: complex cell (43/11s).

Fig. 4.5A shows the information curves for the simple cell of Figs. 4.2 and 4.3A. The squares are taken directly from Fig. 4.3A and represent the contrast-related information conveyed by the full response. The circles represent the transient response, which conveys at most 63% of the peak contrast-related information in the full response and does so with an informative temporal precision limit of 10 msec. The triangles represent the tonic response, which conveys 71% as much contrast-related information as the full response with a temporal precision limit of 22 msec. Within the range of the decoding schemes parameterized by q, the information contents of the full, transient, and tonic response are most easily evaluated by counting spikes. Finally, the diamonds represent the

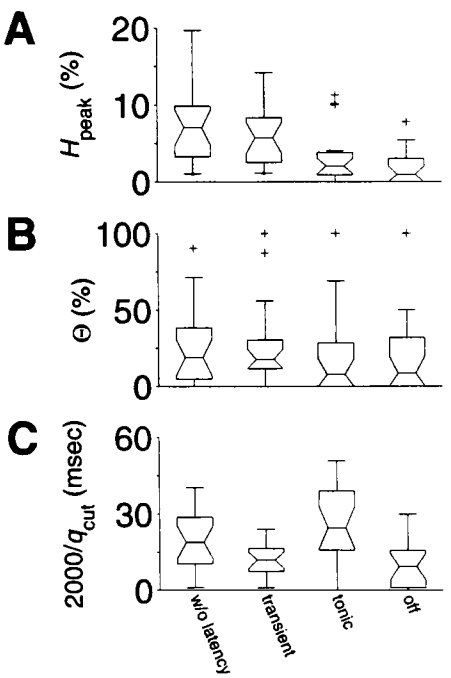

Figure 4.6. *Summary of contrast-related information in the temporal response components.* Distributions are derived from the responses of 20 V1 neurons to stationary sinusoidal gratings. These responses all have distinguishable temporal components (transient, tonic, off). Box-plot details as in Fig. 4.4.

off response, which is relatively weak in this neuron at this spatial phase (see Fig. 4.2A). Not surprisingly, contrast is least well encoded by the off response.

Fig. 4.5B shows the information curves for a complex-cell response with very prominent transient and off response components and a tonic response close to the background firing level. For this neuron, the individual response components encode contrast poorly, whereas the full response encodes, at its peak, 41% of the available information about contrast. This information is almost exclusively (Θ=89%) temporally coded and is conveyed in the latency rather than in the temporal structure of the response components (not shown).

Across all neurons, the transient and tonic responses, and especially the off response, encode substantially less contrast-related information than the latency or full response, at least for the set of contrasts used here. The median peak information (H_{peak}) is significantly lower for each response component than for the full response or the first spike alone (p<0.001; distributions not shown). A more appropriate comparison, though, is the one shown in Fig. 4.6A, where the information-parameter distributions for the full response with latency removed are presented in the first column. This is because the responses derived from the three response components also do not preserve latency variation, as discussed above. H_{peak} for each of the three components is significantly lower than H_{peak} for the full response with latency removed (p<0.004). However, when the values of H_{peak} are summed across the three response components separately for each neuron, the result is significantly greater than H_{peak} of the response with latency removed (medians: 11% vs. 7.1%, respectively, p<0.001). This means that the information about contrast conveyed by the three response components is redundant.

To the extent that the different response components do encode contrast, the transient is significantly more effective (higher H_{peak}) than either the tonic or off component ($p<0.004$). The timing of spikes within the transient and tonic components is not likely to play a primary role in the encoding of contrast (Fig. 4.6B), although the time at which the transient component begins—equivalent to the latency—is clearly important. Fig. 4.6C shows that for the subset of neurons (approximately 50%) that actually encode contrast in the temporal structure of transient and tonic responses, spikes are significantly more precise in the transient than in the tonic ($p=0.02$). The fine informative temporal precision estimate for the off responses is most likely an artifact of poor fits to the information estimates. Finally, among the neurons with responses that could be clearly divided into transient, tonic, and off components, there were no reliable differences between simple ($N=7$) and complex cells ($N=13$) in any of the information measures.

Information estimates depend strongly on the sampling range and density of stimulus contrast. For each neuron tested, the stimulus set consisted of stationary sinusoidal gratings at either an arithmetic series of eight contrasts or a geometric series of six contrasts. The results presented in Figs. 4.3-4.6 are derived from responses to contrasts that evoked a robust change in each neuron's firing rate, in which response-component boundaries could be estimated. Thus, the number of contrasts and the particular contrast values analyzed differ from neuron to neuron. In order to make across-neuron comparisons, information values are normalized by the maximum information available in the stimulus set (i.e., $\log_2$ of the number of stimulus categories).

However, the calculated information values depend strongly on the particular contrasts that are analyzed, and not just on the number of contrasts. Intuitively, this makes good sense: the greater the difference in contrast between two stimuli, the easier to distinguish them and, by extension, to distinguish a neuron's responses to them. Thus, the information transmitted about a pair of stimuli at contrasts 0.125 and 0.25 is expected to be lower than the information transmitted about a pair of stimuli at contrasts 0.125 and 1. As more contrasts are added to the stimulus set, two things happen. First, the maximum information that can be transmitted in response to the entire set increases, because the number of stimuli is larger. Second, there is a greater potential for confusing the various

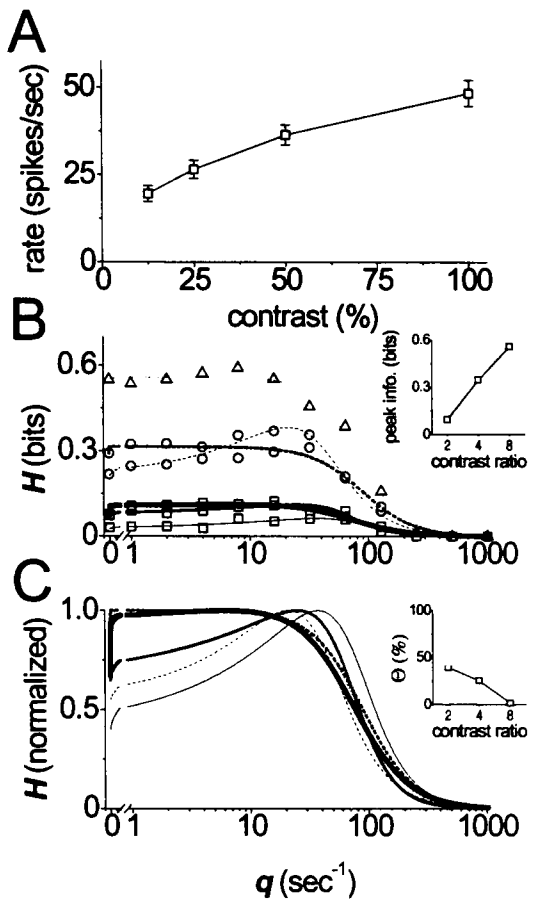

Figure 4.7. *Information about contrast depends on the set of stimuli.* Responses of a V1 simple cell (43/10s) to stationary gratings of four contrasts (0.125, 0.25, 0.5, and 1). **A**: firing rate as a function of contrast. Error bars represent two standard errors of the mean. **B**: information curves (bits) and empirical fits for all pairwise combinations of contrasts. Symbols and line type indicate the contrast offset in the pair. Solid lines and □: factor of two. Dashed lines and ○: factor of four. Dotted lines and △: factor of eight. Thickness indicates the lower contrast in each pair; thin: 0.125; medium: 0.25; thick: 0.5. **Inset**: mean value of H_{peak} for fixed values of the ratio of higher to lower contrast. **C**: fits from **B**, each normalized to its own maximum. In this representation, the information at q=0 sec^{-1} is the fraction of the total information that is contained in the spike count or firing rate (i.e., 1-Θ/100). This plot reveals that spike count plays a relatively more important role when the contrasts are sparsely sampled (contrast ratio 8); conversely, temporal coding is more important when the contrasts are densely sampled (contrast ratio 2). **Inset**: mean value of the temporal coding fraction Θ for fixed values of the contrast ratio.

contrasts, both because there are more of them and because the particular contrasts may evoke similar responses. The balance of these effects determines whether the transmitted information is larger or smaller when responses to more contrasts are included. Also of interest is whether changing the stimulus set determines the aspects of the responses that are most informative in discriminating among contrasts—in particular, whether the information is primarily encoded in the spike count or spike times, and with what precision.

Fig. 4.7A shows the dependence of spike count upon contrast for a V1 simple cell stimulated with stationary sinusoidal gratings. The response increases over the entire range of contrasts but shows signs of saturating at the highest contrasts. The information curves derived from this neuron's responses to all six pairwise combinations of the four contrasts (0.125, 0.25, 0.5, and 1) are shown in Fig. 4.7B. Here, information is given in bits rather than percentages, and the maximum possible transmitted information is 1 bit. Line thickness indicates the lowest contrast in the pair—thin for 0.125, medium for 0.25, and thick for 0.5. Solid lines and squares represent pairs in which the two contrasts differ

by a factor of two, dashed lines and circles by a factor of four, and the dotted line and triangles by a factor of eight.

Intuitively, one expects that closely spaced contrasts are difficult to distinguish, and that information estimates calculated from neuronal responses to pairs of closely spaced contrasts should consequently be low. This intuition accounts for the clustering of information curves by line type (solid, dashed, or dotted), corresponding to different contrast ratios of the stimuli, along the vertical axis in Fig. 4.7B. The mean value of H_{peak} increases with the *contrast ratio* (ratio of the higher to lower contrast; inset of Fig. 4.7B). Unexpectedly, the fraction of temporally encoded information ($\Theta/100$) decreases with the contrast ratio. This can be seen in Fig. 4.7C, which plots the fits from Fig. 4.7B, each normalized to its own maximum. The fraction of information that is encoded in the firing rate ($1-\Theta/100$) is given by the value of each curve at q=0 sec^{-1}; this fraction is greatest when the contrasts of the two stimuli are 0.125 and 1, and least when the contrasts are 0.125 and 0.25. The mean values of Θ are plotted as a function of the contrast ratio in the inset of Fig. 4.7C.

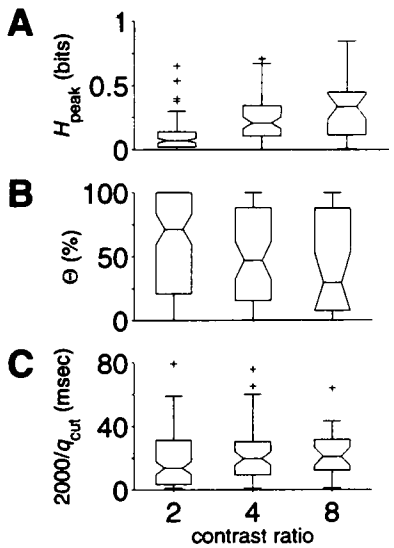

Figure 4.8. *Summary of information parameters for responses to pairs of contrasts.* Parameters derived from fits to the information curves. Distributions across 50 neurons. Box-and-whisker plots as in Fig. 4.4. N=29 for contrast ratio 8, N=63 for contrast ratio 4, and N=111 for contrast ratio 2. **A**: peak information (H_{peak}), expressed in bits, greatest for widely separated contrasts (contrast ratio 8). **B**: temporal coding fraction (Θ), greatest for closely separated contrasts (contrast ratio 2). **C**: temporal precision limit ($2000/q_{cut}$), generally (but not significantly) finest for closely separated contrasts (contrast ratio 2).

The neuron of Fig. 4.7 is typical of the population. This is summarized in Fig. 4.8, which shows the distributions of the three key response statistics as a function of the contrast ratio. Fig. 4.8A shows that the peak contrast-related information H_{peak} increases with contrast ratio (p<0.0001, Kruskal-Wallis non-parametric ANOVA). On the other hand (Fig. 4.8B), the relative amount of temporal coding in the response is largest when the contrasts are closely spaced (low contrast ratio) and smallest when the contrasts are far apart (p=0.03). The informative temporal precision limits of the spikes that contribute to distinguishing contrasts does not change significantly with

contrast ratio (Fig. 4.8C; p=0.15). On the basis of these results, it is proposed that a major role of temporal coding is to enable the visual system to distinguish among stimuli even when there is little change in firing rate. Ultimately, of course, as the difference between the two contrasts is decreased, even the temporally coded information must fall to zero, and the precision of coding is then undefined.

Discussion

The results in this chapter address two important issues regarding contrast encoding by V1 neurons. First, they provide insight into the detailed temporal structure of responses to stationary sinusoidal gratings and the degree to which information about contrast is encoded in each distinct temporal component. Second, they suggest a hypothesis for the role played by temporal coding in contrast discrimination.

From a clinical point of view, the results may be relevant for understanding aspects of visual loss in patients with multiple sclerosis, who tend to have defects in contrast sensitivity that are out of proportion to their loss of visual acuity (Regan et al., 1977). The major pathophysiological effect of chronic demyelination—the primary lesion in multiple sclerosis—is slowing of conduction, which could easily disrupt the finely tuned variations in response latency that are so informative about the contrast of visual stimuli. Thus, it is reasonable to speculate that disturbances of the temporal structure of responses may play a critical role in the visual defects seen in demyelinating diseases.

Role of different response components. The temporal structure of the spike-train response of a visual neuron can be shaped by a number of factors. Most importantly, perhaps, the temporal structure of the response can directly reflect temporal changes in the stimulus. However, information about static features of the stimulus (for the stimuli considered here, contrast, spatial frequency, and orientation) can be multiplexed into the temporal structure of the response (McClurkin et al., 1991; Victor and Purpura, 1996). For stimuli that vary rapidly in time, these two sources of temporal modulation in the response are likely to be confounded. The temporal modulation in the stationary stimuli—

here, sinusoidal gratings that, after an abrupt onset, are present for 237 msec and then replaced by a uniform field at the same mean luminance—is relatively simple. This makes it possible to study the ways in which stimulus contrast, *per se*, affects the responses of V1 neurons. Consistent with previous reports, the results indicate that contrast is encoded in both the firing rate and temporal structure of stationary-grating responses.

The summary distributions plotted in Figs. 4.4 and 4.6 show that, for these stimuli, nearly all the available information about contrast is contained in some combination of firing rate and latency, and that at least some portion of that information is encoded redundantly into both aspects of the response. That neurons encode contrast-related information into these two response parameters has been known for some time (Hartline, 1938). The present results show that latency, the variation of which can depend precisely on contrast, conveys significantly more contrast-related information than does firing rate in the responses of monkey V1 neurons. In this context, it is important to point out that although the method provides an estimate of the informative temporal precision limits in these responses, it does not prove that such temporal precision is used by the brain. To examine this issue directly, experiments would need to be performed in which the perceptual or behavioral consequences of manipulating the fine temporal precision of V1 neurons' responses are examined.

It is not immediately clear that cortical neurons can actually obtain an accurate measure of latency, which is a necessary prerequisite for decoding the contrast-related information encoded therein. Determination of latency requires a comparison of response onset to stimulus onset, but the response onset is itself the neuronal representation of stimulus onset. A number of solutions to this problem can be proposed. One possibility is that there is an overall population activation in V1 that occurs regardless of the particular visual stimulus. Latency information could then be extracted through a comparison of the response times of particular neurons to the time of this general activation. In all likelihood, the characteristics of the general activation change with stimulus contrast, just as they do for individual neurons. In this case, latency information could potentially be extracted by downstream neurons that measure the distribution (in particular, the variance) of the onset times of responses in an ensemble of nearby neurons. Additionally, if latency

is correlated with the degree of synchrony across multiple neurons, postsynaptic "coincidence detectors" would be able to extract the information contained in the latency (Singer, 1999b). Finally, neurons might be able to measure response latency through a comparison of response onset to the time of occurrence of a preceding saccade, which could be taken as á sign that a new stimulus is present. Each of these solutions can in principle be tested explicitly, although to do so would be a challenge to current experimental techniques.

Beyond firing rate and latency, V1 neurons transmit very little information about the contrast of transiently presented stimuli. This is consistent with the results of other investigations (Gawne et al., 1996b; Wiener et al., 1999) and raises the question of why the responses to stationary gratings contain such prominent temporal variation, reflected in the transient, tonic, and off response components. The present results, together with earlier work, suggest that these response components may primarily transmit information about other stimulus parameters such as orientation, spatial frequency, and spatial phase (Victor and Purpura, 1996). For simple cells, in particular, the timing and magnitude of the transient, tonic, and off response components strongly depend on spatial phase (Movshon et al., 1978b) and therefore convey a great deal of information about that stimulus attribute (Victor and Purpura, 1998). This would follow naturally if different response components reflect the contributions of distinct receptive-field subunits, as has been suggested (Movshon et al., 1978b).

Role of temporal coding in the representation of contrast. Previous work has explicitly evaluated the role of temporal coding in the representation of stimulus contrast. (Victor and Purpura, 1996) find that contrast is encoded with higher temporal precision than other stimulus attributes: between 10 and 30 msec. In a related study, (Mechler et al., 1998b) confirm that contrast can be encoded into the temporal structure of spike train responses but demonstrate that the temporal structure conveys more information when stimuli have transient components than when they do not. Thus, temporal coding is prominent in the responses to drifting edges or square-wave gratings, just as it is when stimuli appear and disappear abruptly, but less prominent in the responses to drifting sinusoidal gratings. This is perhaps surprising given that the responses to drifting sinusoi-

dal gratings exhibit a prominent phase advance (Dean, 1981; Albrecht, 1995), which should be reflected in the information measurements. That the phase advance does not give rise to substantial contrast-related information suggests that the variability in the response phase when the stimulus has no transient overwhelms the informative contrast-dependent variation.

The present results extend the analysis of contrast representation by demonstrating (Figs. 4.7 and 4.8) that temporal coding plays a relatively (but not absolutely) more important role as the contrast ratio decreases. This agrees with findings in the locust olfactory system, in which precise spike times make it possible for the animal to discriminate among odorant stimuli that evoke similar firing rates (Stopfer et al., 1997). In V1, temporal coding can better be used to distinguish contrasts that differ by a factor of two than by a factor of eight. This is not to say, however, that contrasts beyond the saturation point, which give rise to responses with similar firing rates, can be distinguished as efficiently as lower contrasts even when spike timing information is taken into account—indeed, the opposite is true (Fig. 4.8). Nonetheless, the important implication is that temporal coding—in particular, variations in response latency—extends the dynamic range of V1 responses with respect to contrast representation beyond what would be available from differences in firing rate alone.

Appendix

In this Appendix, simulations are presented to test the suitability of the metric-space method for calculating information when there are only one or two spikes per trial. This is akin to the situation, described in Results, in which derived spike trains that contain only the first spike in a response are evaluated. More general simulations that test the metric-space method under a variety of conditions are presented in (Victor and Purpura, 1997).

The simulated response sets analyzed in this Appendix each consists of two stimulus conditions. Within each condition, the spike times are drawn from either one or

two Gaussian distributions, one spike per distribution. The means and standard deviations of the distributions, and the overall offset of the distribution means between conditions, are the parameters that are varied in the simulations. For each condition, a fixed number of trials are simulated, each of which contains the same number of spikes (either one or two). The simulated spike trains are subjected to the metric-space analysis, and the results are compared with the results of an analysis in which a simulated response is assigned to the stimulus category that has the highest likelihood of giving rise to spikes at the observed times (Cover and Thomas, 1991). The latter analysis is performed on 10,000 simulated spike trains from the same underlying distributions—100 times more trials than are available from real data. In cases when there is only one spike per trial, it is easy to calculate the information explicitly from the probability distributions; the information values derived from the likelihood-ratio analysis match the true information values very well.

Fig. 4.9A shows the effects of varying the cost parameter q and the number of simulated trials on the estimated information. In this example, there is only one spike per trial. The separation between distribution means across conditions is 16 msec, and the distributions have a standard deviation of 8 msec (so that there is considerable overlap which can lead to ambiguity in assigning responses to the appropriate stimuli). The left panel in Fig. 4.9A shows the bias-corrected information values as a function of q for four different sets of simulated responses, each of which contains a different number of trials: 16 (□), 64 (○), 128 (△), and 1024 (▽). The solid horizontal line shows the actual information (0.37 bits) calculated analytically. For low values of q, the calculated information is close to the actual information, at least when there are a sufficient number of trials. For higher values of q, the information declines to zero, which is expected because spike times measured with extremely high temporal precision do not provide information about the stimulus (Victor and Purpura, 1997).

The right panel of Fig. 4.9A shows estimated information plotted against the number of trials simulated for four different values of q: 1 (□), 8 (○), 64 (△), and 512 (▽) sec^{-1}. When the number of trials is on the order of 100—the number of trials at each contrast in the real data described in this chapter—the calculated information approaches the actual information when q is sufficiently low. The similarity of the curves for differ-

ent values of q indicates that the accuracy of the bias-correction method does not depend strongly on q, even though it can do so in principle (Victor and Purpura, 1997).

Fig. 4.9B-D shows the results of varying the mean separation and standard deviation of the Gaussian distributions. The left panels plot H_{peak} (the maximum calculated information), and the right panels plot $2000/q_{cut}$ (the measure of informative temporal precision used in this chapter). Both parameters are derived from fits of the information curves (see Methods). Different symbols correspond to different ratios of mean separation to standard deviation: 1 (□), 4 (○), and 16 (△). For each condition, 100 simulated spike trains are created. Solid lines in the left panels correspond to likelihood-ratio information measures. Slanted lines in the right panels correspond to expected temporal-precision limits, which are

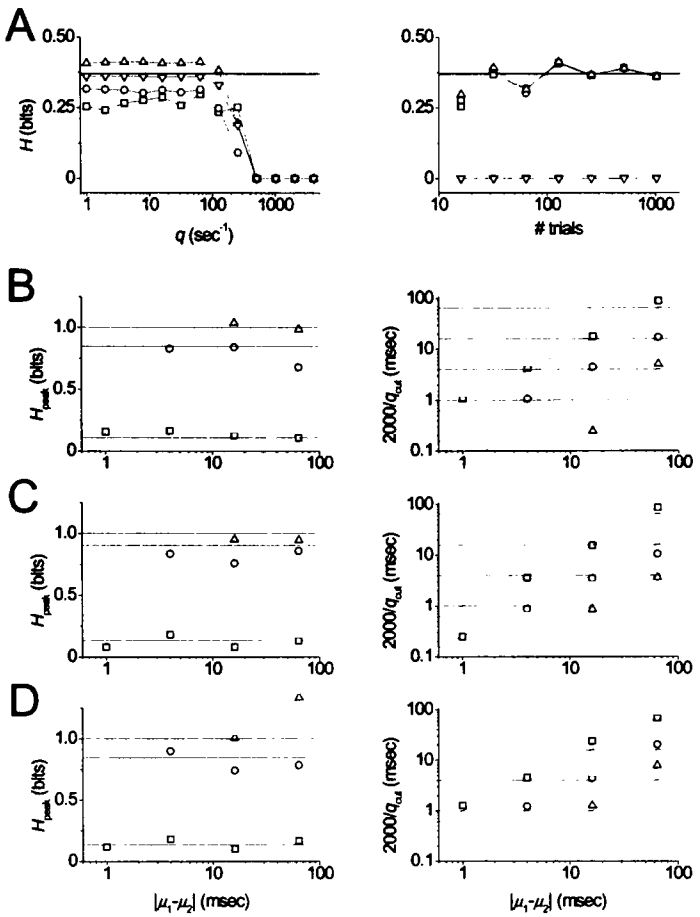

Figure 4.9. *Metric-space results for simulated data.* Test of the suitability of the metric-space method for estimating information and temporal precision when the number of spikes per trial is low. **A**: bias-corrected information as a function of the cost q (left) and the number of simulated trials (right). In the left panel, different curves represent different numbers of simulated trials: 16 (□), 64 (○), 128 (△), and 1024 (▽). In the right panel, different curves represent different values of q: 1 (□), 8 (○), 64 (△), and 512 (▽) sec^{-1}. **B-D**: peak information (H_{peak}) and temporal precision limits ($2000/q_{cut}$) derived from fits to the information curves. Solid lines in the left panels show expected information values from a likelihood-ratio analysis. Solid lines in the right panels show expected temporal precisions for different mean separations (slanted lines) and standard deviations (horizontal lines) of the Gaussian distributions that are used to generate spike times. Symbols correspond to different ratios of mean separation to standard deviation: 1 (□), 4 (○), and 16 (△). **B**: one spike/trial. **C**: two spikes/trial; spike-time distributions within a trial have a mean separation of 32 msec. **D**: two spikes/trial (same data as in **C**), but only the first spike in each trial is analyzed.

related to the mean separation between distributions used in the simulations. They are independent of the standard deviations (horizontal lines) because the informative temporal precision corresponds to reliable variations in spike timing across stimulus conditions that can provide information about the stimulus, rather than jitter of particular spike times across trials within a condition. The estimated temporal precisions for the stimuli used in these simulations are expected to fall at the intersections of the horizontal and slanted lines.

Fig. 4.9B shows the case of one spike per trial; a subset of these data is plotted, in more detail, in Fig. 4.9A. Fig. 4.9C shows the case of two spikes per trial, where the individual spike-time distributions within each trial are separated by 32 msec. Fig. 4.9D shows the results of extracting the first spike from each of the trials in Fig. 4.9C and performing the metric-space analysis; this corresponds to extracting the first post-latency spike at each contrast in real data. The results indicate that the estimated information values for 100 trials are generally very close to the expectation, regardless of the mean-separation-to-standard-deviation ratio, in all three simulations (left panels). Moreover, the calculated temporal precision limits fall near the intersections of the horizontal and slanted lines in the right panels, indicating that the temporal precision is recovered relatively accurately by the metric-space method. Together with the results of (Victor and Purpura, 1997), the simulations reveal that the metric-space method, as used in this chapter, can accurately estimate information and informative temporal precision limits from spike data.

Chapter 5

Interspike Intervals, Receptive Fields, and Responses to Rapidly Varying Stimuli

Summary

Chapter 4 focuses on the information that is temporally encoded into the responses of V1 neurons to transiently presented, stationary stimuli with simple spatial structure. This chapter concerns itself with the responses to stimuli that change rapidly in time and that have far more spatial complexity than sinusoidal gratings. The interspike intervals in these responses can be grouped into classes on the basis of reproducible features in the interspike-interval histograms, when these histograms are constructed in a novel manner. The kind and amount of information conveyed by a spike depends on the duration of the immediately preceding interspike interval. Spikes in different classes bear different relationships to the visual stimulus, both qualitatively (in terms of the average stimulus preceding each spike) and quantitatively (in terms of the amount of information encoded per spike and per second). In particular, spikes preceded by very short intervals (3 msec or less) contribute disproportionately to the overall RF properties of V1 neurons. Although many of the results would be expected for neurons that use a constrained firing-rate code to transmit information, the evidence suggests that visual neurons are well-equipped to decode stimulus-related information on the basis of relative spike timing and interspike-interval duration.

Introduction

Responses to stimuli that change slowly or rarely, such as those presented in Chapter 4, provide an excellent substrate on which to examine the ways in which neurons

encode stimulus-related information into time-varying changes in their average firing rates and into precise spike times. The results indicate that a great deal of information is carried in the spike times, particularly in the time of the first spike in the response, and that this time of occurrence can be very precise.

Analysis of responses to stimuli that change rapidly in time present a different challenge because spike timing can be affected by the mechanisms at work in the responses to stationary stimuli, but also by the time-variation in the stimulus itself. These two mechanisms have been called *temporal encoding* and *temporal coding*, respectively (Théunissen and Miller, 1995), though the distinction between them may be artificial. Because of this ambiguity, the approach taken in this chapter is different, in spirit and in substance, from the approach taken in the previous chapter. Specifically, although the information transmitted by neuronal responses is estimated, the primary foci are the examination of interspike-interval (ISI) statistics and the isolation of stimulus configurations that evoke ISIs of different durations. Some of the methods used to analyze ISI distributions are related to the methods introduced in Chapter 3, although in this chapter there is no attempt to track the time-variation of the ISI distribution. At the end of the chapter, a different information-theoretic approach will be introduced; this approach is explained and explored in greater detail in Chapters 6 and 7, where it is applied to a variety of stimuli. Here, its function is confined to quantification of the information conveyed in ISIs of different durations.

Previous work has demonstrated that cortical sensory neurons have high intrinsic temporal precision (Mainen and Sejnowski, 1995; Nowak et al., 1997) and can encode information on the scales of milliseconds and tens of milliseconds (Bura as and Albright, 1999). Three questions arise immediately: (1) What kinds of stimuli are encoded on the different time scales in a neuron's response? (2) How much information is encoded on each time scale? (3) How might this information be decoded by relatively simple components of neurons and neural circuits?

Here, the first two questions are addressed experimentally through a measurement of the responses of V1 neurons to rapidly varying, pseudo-random ("m-sequence") stim-

uli. ISIs fall into three subsets, distinguished on the basis of ISI duration, in a stereotyped manner across neurons. A reverse-correlation procedure is used to generate receptive-field (RF) maps from the full responses as well as from response subsets that only contain spikes that succeed ISIs of particular durations. Finally, information theory is used to quantify the rate and efficiency with which full responses and response subsets convey information about the visual stimulus.

The results indicate that spikes in different ISI subsets are preferentially fired in response to different visual stimuli. In particular, spikes preceded by ISIs shorter than 3 msec, which occur during periods of very high firing rate, tend to be evoked by stimuli that have several subregions of opposite contrast covering the neuron's RF. Each of these spikes also tends to convey more stimulus-related information than the average spike. On the other hand, spikes preceded by ISIs longer than 38 msec are often fired in response to spatially uniform stimuli that reverse contrast over time; each of these spikes is less informative than average.

Pure firing-rate codes can generate subset RF maps and yield information rates similar to those of actual neuronal responses, even though the structure of the ISI distribution is different. By itself, however, this result says nothing about whether the visual cortex uses a firing-rate or temporal code to transmit stimulus-related information (the third general question raised above). Nonetheless, the existence of stereotyped ISI durations in V1 (described in this chapter), together with recently described synaptic and dendritic machinery (such as depression, facilitation, and coincidence detection) that can selectively increase or decrease the importance of particular spikes in shaping a postsynaptic response, suggests that real-time decoding of neuronal signals may rely on known biophysical mechanisms that are specifically sensitive to ISI duration.

Methods

Recording and stimuli. Data-collection methods are as described in Chapters 2, 3, and 4. The precision of spike-time collection is 0.1 msec for 135 neurons and 3.7 msec

(one frame of the visual display) for 36 additional neurons. Spike sorting (see Chapter 2) for the 36 neurons in the second group was performed during the experiments with hardware window discriminators; spike isolation was monitored throughout the experiments but could not be verified offline at the end of the experiment. Of the 171 neurons, 99 (32 simple, 60 complex, and 7 unclassified (Skottun et al., 1991)) responded with firing rates higher than 3 spikes/sec and had significantly modulated RF maps (see below); these neurons are the source of the data presented in this chapter.

The stimuli used in this chapter look to human observers like randomly and rapidly flickering checkerboards. In fact, however, the stimuli are highly structured and deterministic. They consist of a grid in which the temporal sequence of the luminance levels (0 or 300 cd/m^2) in each of 249 pixels (typically 16 x 16 arc-min) is determined by a pseudo-random, binary m-sequence (Sutter, 1992; Victor, 1992; Reid et al., 1997). The same 4095-step m-sequence is used in each pixel, but the starting position in the sequence is different. Because the minimum offset between any pair of pixels is 237 msec, and because an m-sequence is uncorrelated with temporal shifts of itself, systematic correlations of the response with the stimulus sequence can be uniquely assigned to particular checks, provided that latencies beyond 237 msec are rejected as non-physiologic. The m-sequence is advanced simultaneously in all pixels at a stimulus frame rate of 67.6 Hz, so that the luminance of each stimulus pixel has a 50% chance of changing every fourth frame of the 270.3 Hz visual display. This stimulus frame rate has been shown to evoke good RF maps from cortical neurons (Reid et al., 1997). The entire m-sequence stimulus, which lasts 60.6 sec, is repeated eight to 16 times. Presentations are interleaved with a presentation of the contrast-inverted stimulus, which eliminates spurious effects of residual correlations in the stimulus on the RF maps (Sutter, 1992). Standard and contrast-inverted sequences within a repeat are separated by a period of 23 sec during which a uniform field at the mean luminance is presented; repeats are separated by 18 sec.

RF map calculation. Cross-correlation of an evoked spike train with the m-sequence stimulus—a process also known as *spike-triggered averaging*—yields a detailed map of the neuron's spatiotemporal RF (see Fig. 5.4). This map essentially represents the average stimulus preceding each spike, and it is rendered here as a series of con-

tour plots that depict spatial snapshots sequential in time. Each map shows the average change in contrast in each of the stimulus pixels that significantly modulate the response for at least two consecutive 3.7 msec time bins ($p<0.01$ in each bin: the estimated impulse response, averaged across trials, is at least 2.6 standard errors of the mean from zero), as well as for surrounding pixels. Contour maps are smoothed by cubic-spline interpolation.

To the degree that a neuron is a linear system, its derived RF map can also be considered to depict its *spatiotemporal impulse response*—that is, the average response to an incremental flash of light at each spatial position (Victor, 1992). This interpretation of the RF map depends on the spatiotemporal correlation structure of the m-sequence stimulus. Of course, V1 neurons are not linear systems, and there are often higher-order components of the RF map that makes the impulse-response interpretation imprecise. In this situation, the RF maps are the linear functions that best fit the full response, and the spirit of the impulse-response interpretation is retained in the normalization of the RF maps. Thus, contour heights represent, at each time frame, the change in firing rate induced by a luminance step in a particular stimulus pixel averaged over the displayed time window of 14.8 msec.[4]

Information-rate calculation. A modification of the *direct method* is used to calculate the rate of information transfer in the m-sequence responses (Ruyter van Steveninck et al., 1997; Strong et al., 1998). This method is quite different from the metric-space method used in Chapter 4, and it is explained in detail in Chapter 6. Briefly, stimulus-related information is taken to be the total entropy, which quantifies the neuron's repertoire of possible responses, minus the noise entropy, which represents the variability across repeated trials of the same stimulus. The total entropy is derived from the probability distribution of spike counts in short bins across all time and trials in the response.[5] The noise entropy is derived from the probability distribution of spike counts across trials

[4] To obtain absolute firing rates for the full response, for which the RF map is denoted by **f** in equations (5.1) and (5.2), add the displayed values to the mean firing rates, which are given in the caption to Fig. 5.4. To obtain absolute firing rates for the subset responses (**s**; see Results), multiply the RF map values by the fraction of spikes in the appropriate subset and add the product to the subset's mean firing rate. The quantities required for these calculations are provided in the figure caption.

[5] The use of short time bins is dictated by the amount of available data, as discussed in Chapter 6, and means that some fraction of the total information is ignored in the calculation.

at a fixed time in the response, and is then averaged across time. Because only a limited amount of data is available, the entropy estimates are subject to a downward bias, which is corrected by methods detailed in Chapter 6.

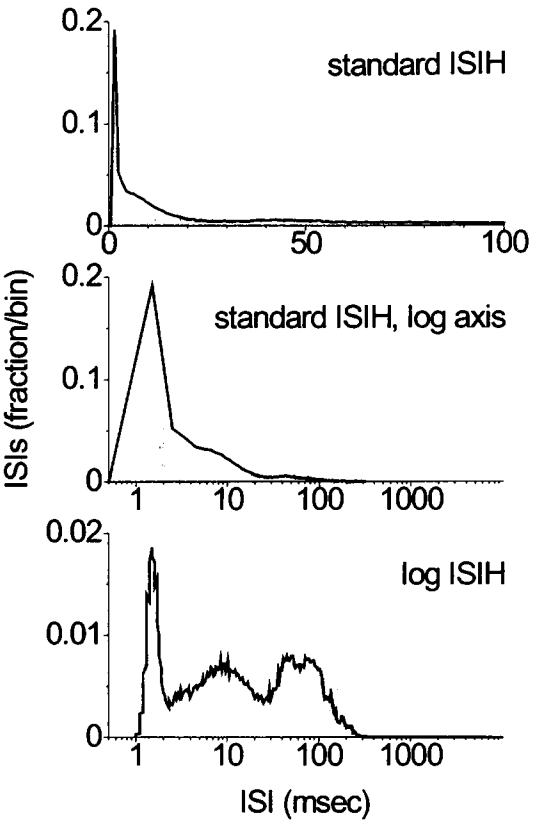

Figure 5.1. *Construction of the log-ISIH.* Each panel shows the ISIHs of the m-sequence response (solid line) and the response to a uniform field at the same mean luminance (shaded) of a simple cell (44/9t) that fired 21.0 spikes/sec; 35,581 ISIs total. **Top**: standard ISIH (1 msec bins). **Middle**: standard ISIH plotted on a logarithmic time axis. **Bottom**: log-ISIH consisting of 300 bins spaced logarithmically between 1 msec and 10 sec, which bracketed the distribution of ISIs found in the data. Each bin is 3.1% wider than the previous one. A relatively smooth histogram is obtained by applying a uniformly distributed, random timing jitter to each ISI, the magnitude of which is at most half the data collection resolution of 0.1 msec. The log-ISIH makes it easy to distinguish three separate ISI peaks, which are not readily visible in the standard ISIH.

Results

Interspike intervals. The data analyzed in this chapter are the responses of 99 V1 neurons in anesthetized macaque monkeys. The stimuli—multiple repeats of pseudo-random (m-sequence) checkerboards—contain a wide variety of spatial and temporal patterns. None of these patterns dominates the stimulus, but some of them are typically effective stimuli for these neurons. These stimuli are therefore well suited to probe a neuron's ability to convert spatial information into spike trains.

Fig. 5.1 shows ISIHs constructed from the response of a simple cell both to the m-sequence stimulus (solid line) and to a uniform field of the same mean luminance (shaded gray region). The top panel is the standard ISIH, in which ISIs are collected into equal bins of width 1 msec. For both the m-sequence and uniform-field responses, the standard ISIH features a prominent peak at very short ISIs, which decays rapidly. The peak is higher for the m-sequence response

than for the uniform-field response, but it has approximately the same width. Following the initial peak and rapid decay, the m-sequence ISIH shows a secondary, slower decay that lasts from approximately 5 msec until 20 msec; this secondary decay is almost entirely absent from the uniform-field response. Finally, both ISIHs begin a very slow decline, sometimes referred to as the *long tail* (Gerstein and Mandelbrot, 1964; Smith and Smith, 1965), at about 20 msec.

The short-ISI features are displayed more prominently when the standard ISIH is plotted on a logarithmic time scale (middle panel). The similar shape of the initial peak and the different secondary decays for the two responses are clearly evident in this plot, as well. However, detail is obscured both at short and long ISIs—at short ISIs because the binning is relatively coarse, and at long ISIs because there are relatively few spikes per bin. Both difficulties are addressed by presenting the data in yet another way (bottom panel). Here, the histogram is constructed with logarithmic time bins so that the duration of each successive bin is a fixed multiple of the previous one. This *log-ISIH*—a new way of looking at ISI data—highlights features such as prominent peaks or shoulders that are at best barely visible in the standard ISIH. The log-ISIH of the m-sequence response has three distinct peaks (corresponding, in the standard ISIH, to the initial rapid peak, the secondary decay, and the long tail), whereas the log-ISIH of the uniform-field response has only two peaks (the initial rapid peak and the long tail). The comparison between the log-ISIHs of the m-sequence response and the uniform-field response indicates that only some of the log-ISIH features are due to m-sequence stimulation. The similarity of the prominent short-ISI peak suggests that this feature, in particular, is for the most part intrinsic to the neuron.

In Fig. 5.2, the log-ISIH of the m-sequence response from Fig. 5.1 is replotted (top panel) together with log-ISIHs from three additional neurons. Each log-ISIH is manually subdivided into its component peaks; the number of peaks varies from neuron to neuron. Thus, the first neuron has three peaks, the second and third have two peaks (which differ in position), and the fourth neuron has only a single peak. Peaks are gray-scale-coded according to the relative position of the maximum: black for short, light gray for medium, and gray for long ISIs. The bottom panel is a log-histogram of the estimated

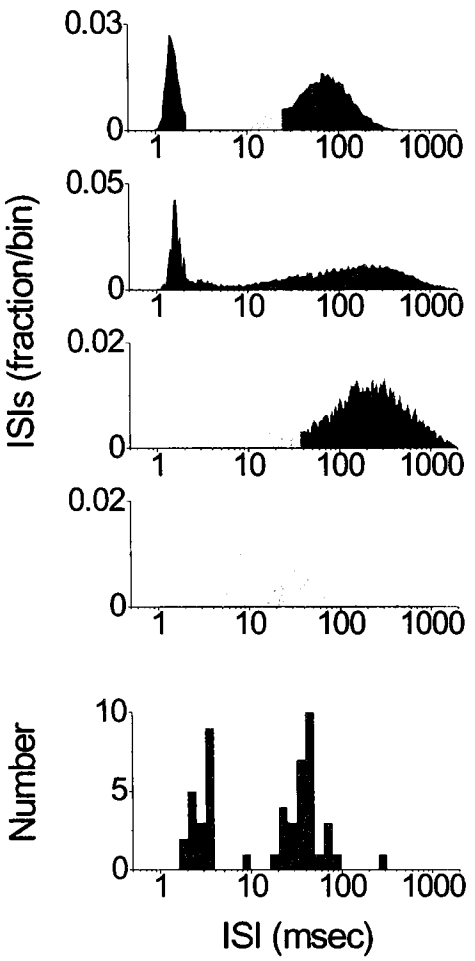

Figure 5.2. *Division of log-ISIHs into component peaks.* Top four panels are the responses of different neurons to the m-sequence stimulus. Log-ISIHs consist of 200 bins spaced logarithmically between 0.1 msec and 2 sec, so that each bin is 5.1% wider than the previous one. Each log-ISIH is divided, by eye, into its component peaks, which are then grayscale-coded according to the position of the maximum (black for < 3 msec, light gray for 3-38 msec, and dark gray for >38 msec). **Top**: same simple cell as in Fig. 5.1; three peaks. **Second panel**: simple cell (38/6); 5.9 spikes/sec; 11,343 ISIs; two peaks (short and long). **Third panel**: simple cell (34/12t); 5.2 spikes/sec; 9,997 ISIs; two peaks (medium and long). **Fourth panel**: complex cell; 29.9 spikes/sec; 57,943 ISIs; one peak (medium). The boundary positions across 66 neurons are collected into a log-histogram (**bottom panel**), which reveals a tightly clustered, bimodal distribution.

boundaries between ISI peaks; it summarizes data from 66 neurons and includes 19 short/medium boundary points and 32 medium/long boundary points. Across the 99 responses subjected to this analysis, 34 had a single peak, 46 had two distinct peaks, and 19 had three peaks. There was no significant difference between simple and complex cells in terms of the number of log-ISIH peaks (χ^2 test).

Two very surprising results emerge from this analysis. First, the summary histogram is bimodal. This indicates that there are at most three log-ISIH peaks in any neuron's response, and that the peak boundaries are stereotyped. Moreover, neurons with only two log-ISIH peaks (the majority) have boundary points that match either the short/medium or medium/long boundary of the three-peak neurons, rather than some intermediate value. The medians of the two boundary points are 3.0 msec (90% range: 1.8 to 3.7 msec) and 38.1 msec (90% range: 20.3 to 64.8 msec). The second surprising feature of the summary histogram is that the two modes are quite sharp, which indicates that log-ISIH peaks are remarkably consistent across neurons. Because the boundary points do not correspond to temporal features of the stimulus, which occur in multiples of 14.8

msec, or to the frame time of the visual display, which is 3.7 msec, they are likely to reflect some aspect of the intrinsic biophysical hardware of cortical neurons and/or their connections.

The consistency of the boundary points across neurons also raises the possibility that the subsets delimited by those boundaries play distinct roles in information encoding. To address this possibility, the ISIs in each neuron's response are divided into three subsets according to the median boundary points across neurons, regardless of the number of peaks in the particular neuron's log-ISIH. ISIs less than 3 msec are considered to be part of the *short-ISI subset*, ISIs between 3 and 38 msec to be part of the *medium-ISI subset*, and ISIs longer than 38 msec to be part of the *long-ISI subset*. Averaged across neurons, 10% of spikes fall into the short-ISI subset, 45% into the medium, and 45% into the long. Box plots showing the distributions across neurons of the percentage of spikes in each subset are shown in Fig. 5.5 (top).

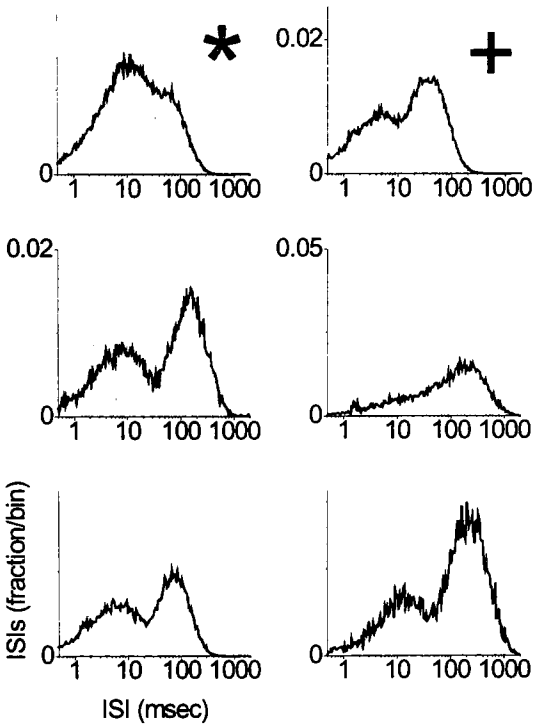

Figure 5.3. *Real vs. exchange-resampled log-ISIHs.* Log-ISIHs of real spike trains (shaded, same parameters as in Fig. 5.2) are compared with those obtained from an exchange-resampling procedure that preserves the firing rate modulation and distribution of spikes per trial from the original data (solid lines). If the original data are consistent with a modulated Poisson process, then the original and resampled log-ISIHs in all panels should superimpose. **Top row, left panel (*)**: complex cell (33/1); 31.7 spikes/sec; 61,556 ISIs. **Top row, right panel (+)**: simple cell (39/9); 34.7 spikes/sec; 67,269 ISIs. **Middle row, left panel**: simple cell (35/1); 10.0 spikes/sec; 19,283 ISIs. The other three log-ISIHs are from Fig. 5.2.

The first step in demonstrating that the individual subsets may play important roles in information encoding is to show that they are not epiphenomena of the firing rate modulation. To do this, an *exchange resampling* procedure (Victor and Purpura, 1996) is applied. Exchange resampling exactly preserves the firing rate modulation and the num-

ber of spikes in each response trial but randomizes the relationship between consecutive spikes. The spike times themselves are drawn at random, without replacement, from the entire set of actual spikes. Log-ISIHs of exchange-resampled responses are shown as solid lines in Fig. 5.3, superimposed upon the log-ISIHs of real responses. If the ISI structure is completely determined by fast firing rate modulations and slow variations in responsiveness, the log-ISIH of real and resampled spike trains should superimpose. In reality, this superposition test succeeds for the longest ISIs but fails for short and medium ISIs. One cause of the failure is likely to be the presence, in real neurons, of a refractory period on the order of 1 msec. However, a refractory period alone cannot account for features such as a sharp notch between short and medium-ISI peaks, or for differences in peak height as seen in the top left panel.

RF maps. For each neuron, RF maps or spike-triggered average stimuli (see Methods) are derived from the full response and from response subsets, where the spikes in each response subset are preceded by ISIs from a single ISI subset (short, medium, or long). Fig. 5.4 shows example RF maps from two neurons, depicting representative examples of the changes that are seen across ISI subsets. RF maps are shown as a series of interpolated contour plots, each averaged over 14.8 msec of the response, at different time lags; taken together, they depict the dynamics of the RF. Red signifies on-subregions of the RF, in which bright stimuli are associated with a higher firing rate than dark stimuli, and blue signifies off-subregions. The ratio of the RF map scales of two subsets is equal to the ratio of the numbers of spikes in the two subsets.

In general, RF maps derived from spike subsets differ from RF maps derived from the full responses, and from each other, in both shape and amplitude. Fig. 5.4A shows the RF map of a complex cell that primarily displays both shape and amplitude changes. This cell fired an average of 31.7 spikes/sec in response to the stimulus, and its log-ISIH is shown in Fig. 5.3 (*). The RF map derived from all spikes (top row) is dominated by the off-subregion, though there is evidence of a weaker on-subregion. When only short-ISI spikes are considered (second row), however, the on-subregion is selectively enhanced in two ways: (1) its duration, measured with a resolution finer than what is shown in Fig. 5.4A, is 42 msec instead of 19 msec; and (2) its peak amplitude, after correcting for the

different numbers of spikes, is 2.9 times larger than the corresponding amplitude in the all-spikes RF map. The relative amplitude of the off-subregion is also enhanced in the short-ISI RF map, but only by a factor of 1.8 and not at every time lag. The shape of the short-ISI RF map suggests that these spikes may selectively encode stimuli that are characterized by strong spatial opponency.

A second prominent RF map change in Fig. 5.4A is found in the long-ISI subset (bottom row). For this subset, the single RF subregion has a biphasic time course, changing from off to on around

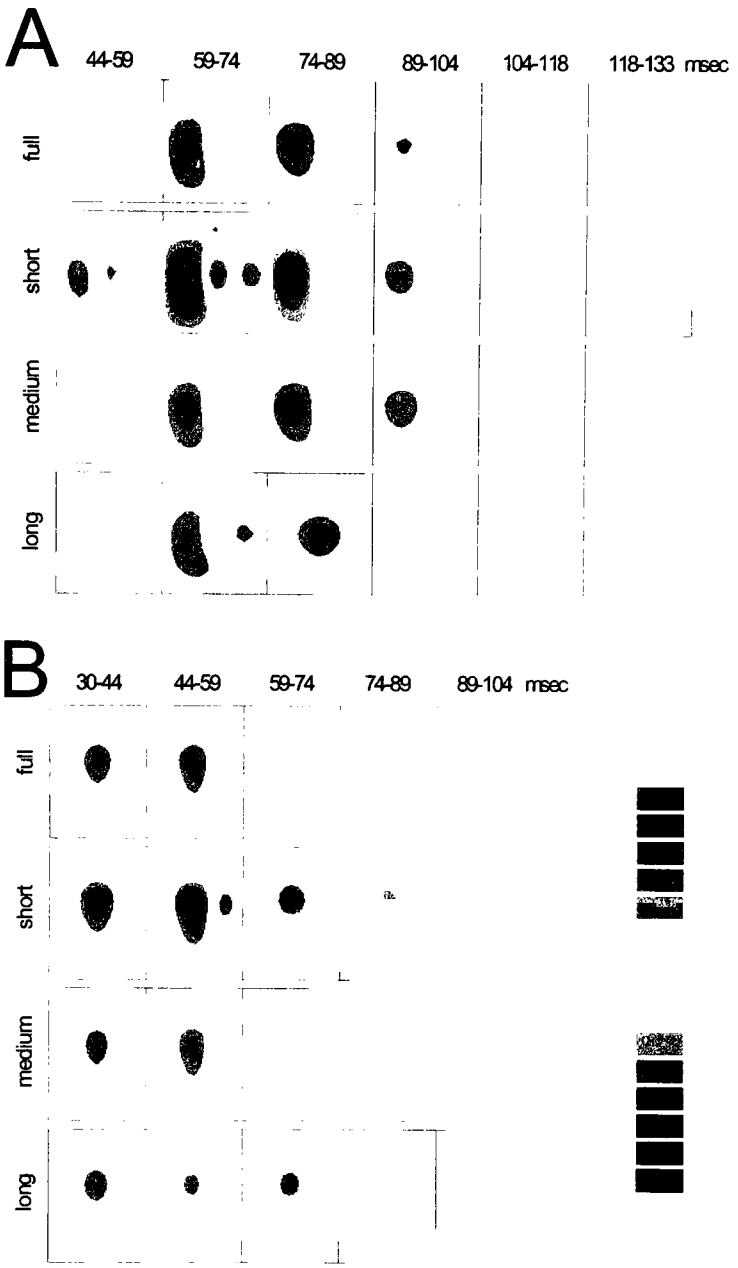

Figure 5.4. *Receptive-field maps.* **A**: *complex cell* (33/1); average firing rate 31.7 spikes/sec; spatial extent 1°22' x 1°22'; information rate 23.8 bits/sec or 0.75 bits/spike; efficiency 16.4%; log-ISIH in Fig. 5.1B. Scale bar: -4.9 to 4.9 spikes/sec. **B**: *simple cell* (39/9); average firing rate 34.7 spikes/sec; spatial extent 1°22' x 1°39'; information rate 26.4 bits/sec or 0.76 bits/spike; efficiency 17.0%, log-ISIH in Fig. 5.1C. Scale bar : -12.9 to 12.9 spikes/sec. In each panel, the **top row** is the RF map of the full response, the **second row** is the RF map of the short-ISI subset (10% of spikes in **A**, 24% in **B**), the **third row** is the RF map of the medium-ISI subset (68% and 50%), and the **bottom row** is the RF map of the long-ISI subset (22% and 26%). Values on the scale bar refer to the RF map of the full response only (**f** in equations (5.1) and (5.2)); to obtain values for any of the ISI subsets (**s**), multiply the full-response values by the fraction of spikes in that subset.

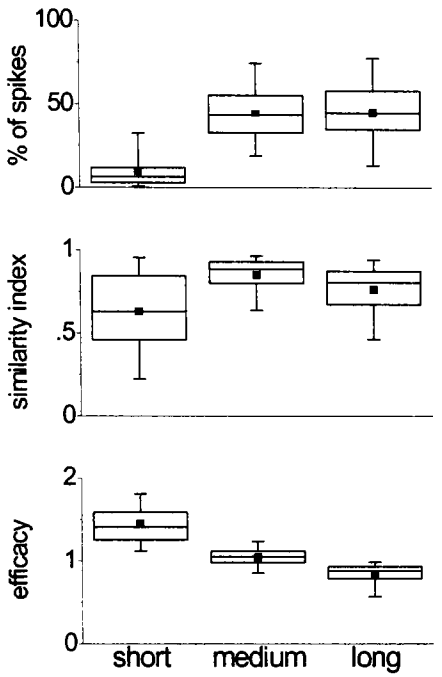

Figure 5.5. *Across-neuron distributions of RF-map indices.* Data from 66 neurons. Box-plot whiskers show the 5th and 95th percentiles, and box boundaries the 25th and 75th percentiles. The horizontal line divides each box at the median, and the square represents the mean. **Top**: percentage of spikes with preceding ISIs < 3 msec (short), 3-38 msec (medium), and >38 msec (long). **Middle**: similarity index (or correlation coefficient) between the subset and full-response RF maps; values can be between -1 and +1. **Bottom**: efficacy, or average contribution made by spikes in each subset to the overall RF-map properties. An efficacy of 1 means that the spikes in that subset contribute as much as expected, given their number.

89 msec, even while the RF maps of other subsets are still dominated by an off-subregion. The time course of the long-ISI RF map, and the fact that it is for the most part spatially uniform at each time lag, suggests that these spikes primarily encode temporal features in the stimulus, and that they typically follow, by at least 89 msec, a period during which the neuron is inhibited. Ignoring the ISI structure and treating all spikes equally obscures the effects of this inhibition.

Fig. 5.4B is the RF map of a simple cell that fired an average of 34.7 spikes/sec in response to the m-sequence stimulus; the log-ISIH of its response is shown in Fig. 5.3 (+). The short and medium-ISI RF maps show evidence of large amplitude changes but less evidence of large shape changes. In other words, scaling these subset RF maps by some factor (called the *efficacy*—see below) would make them look very similar to the all-spikes RF map. The long-ISI RF map, however, primarily exhibits a shape change: the off-subregion is relatively diminished between 30 and 59 msec, and the on-subregion is relatively enhanced between 59 and 74 msec.

Two related indices can be used to quantify these shape and amplitude changes. To calculate these indices, the RF maps are treated as vectors in space and time, with dimension equal to the product of the number of stimulus pixels and the number of time bins. The *similarity index* (DeAngelis et al., 1999) is defined as the Pearson correlation coefficient between vectors derived from two different RF maps:

$$similarity\ index = \frac{\langle \mathbf{s}, \mathbf{f} \rangle}{\|\mathbf{s}\|\|\mathbf{f}\|} \tag{5.1}$$

where **f** is the "vectorized" RF map of the full spike train, **s** is the vectorized RF map of the subset spike train, <...,...> denotes the dot product, and ||...|| denotes the norm, or the square root of the vector's dot product with itself. The similarity index has a value close to 1 for maps that have similar shapes (e.g., 0.98 for the full and short-ISI RF maps in Fig. 5.4B), 0 for maps that are nearly orthogonal, and –1 for maps that have opposite polarities (on-subregions become off-subregions and vice versa). The similarity index is an omnibus measure that averages over both space and time and may therefore dilute the effects of prominent but localized RF map changes. Thus, even a qualitatively large shape change such as occurs in the short-ISI RF map of Fig. 5.4A has a similarity index of 0.83, which is near the 75th percentile of short-ISI similarity indices across neurons (Fig. 5.5, middle). At the other extreme, the shape change seen in the long-ISI RF map of Fig. 5.4A, among the most dramatic in the sample, has a similarity index of 0.60.

A related measure, the *efficacy*, quantifies amplitude changes. The efficacy is the factor by which the amplitude of the subset RF map must be scaled to best match the full RF map in a least-squares sense, after correcting for differences in the number of spikes:

$$efficacy = \left(\frac{n_f}{n_s}\right)\left(\frac{\langle \mathbf{s}, \mathbf{f} \rangle}{\|\mathbf{f}\|^2}\right) \tag{5.2}$$

where n_f is the number of spikes in the full spike train, n_s is the number of spikes in the subset spike train, and other variables are as in equation (5.1). An efficacy greater than 1 signifies that subset spikes play a larger role in the generation of the all-spikes RF map than is expected from the number of spikes in the subset, whereas an efficacy smaller than 1 signifies the opposite. An efficacy less than 0 indicates that the RF map polarity must be flipped and then scaled to obtain the best match. The short-ISI RF map of Fig. 5.4B has a particularly high efficacy of 1.77, signifying a large contribution from those spikes to the overall RF properties of that neuron. The medium-ISI RF map of Fig. 5.4A has an efficacy of 1.02, indicating that those spikes contribute the expected amount to the overall RF map.

Across all neurons, similarity indices are in the range 0.5 to 1 (Fig. 5.5, middle), meaning that shape changes, although present, are moderate. On the other hand, the efficacy distributions of the three subsets are distinct and largely non-overlapping (Fig. 5.5, bottom). This indicates that short-ISI spikes make the largest contribution to the all-spikes RF maps, long-ISI spikes make the smallest contribution, and medium-ISI spikes contribute about as much as expected given their frequency.

Rate coding and temporal coding. To test whether the shape and amplitude changes of RF maps are consistent with a rate-code model of neuronal firing, real data are compared with data derived from the exchange-resampling procedure. Exchange resampling exactly preserves the firing-rate modulation in the original data, as described above. In fact, it yields the only ensemble of spike trains that match the time-varying firing rate of real data and are also rigorously consistent with a firing-rate model. Each spike train is exchange-resampled 200 times, and the similarity indices and efficacies of the subset RF maps of the resampled spike trains are calculated. The same ISI subset definitions are used for real and resampled spike trains, despite the fact that the log-ISIHs are different (Fig. 5.3). Similarity indices and efficacies of the real responses are compared with the means and distributions obtained from exchange-resampled responses.

Fig. 5.6 shows a series of scatter plots that describe the relationship between similarity indices (left column) and efficacies (right column) for real and resampled responses. Each point represents a different neuron. The similarity index or efficacy of a real RF map is plotted along the horizontal axis, and the corresponding mean value from 200 exchange resamplings is plotted along the vertical axis. Results from different subsets are in different rows. In all panels, the cloud of points lies near the line of equality, where real and resampled RF maps have the same similarity indices or efficacies. The largest deviation is in the top-right panel, which reveals that the short-ISI efficacies tend to be higher for exchange-resampled RF maps than for real RF maps. This indicates that short-ISI spikes play an even larger role in generating the RF properties of rate-code-generated spike trains than of real spike trains.

On a neuron-by-neuron basis, the similarity indices and efficacies of the real spike trains tend to fall in the tails of the distributions of similarity indices and efficacies of resampled spike trains. In Fig. 5.6, solid points represent cases where the real data fall within either 2.5% tail of the distribution of resampled data. However, the behavior of the two indices is qualitatively—and, for the most part, quantitatively—similar for real and resampled RF maps, suggesting that the rate-code model does a reasonable job explaining the indices (though not, as described above, the ISI statistics).

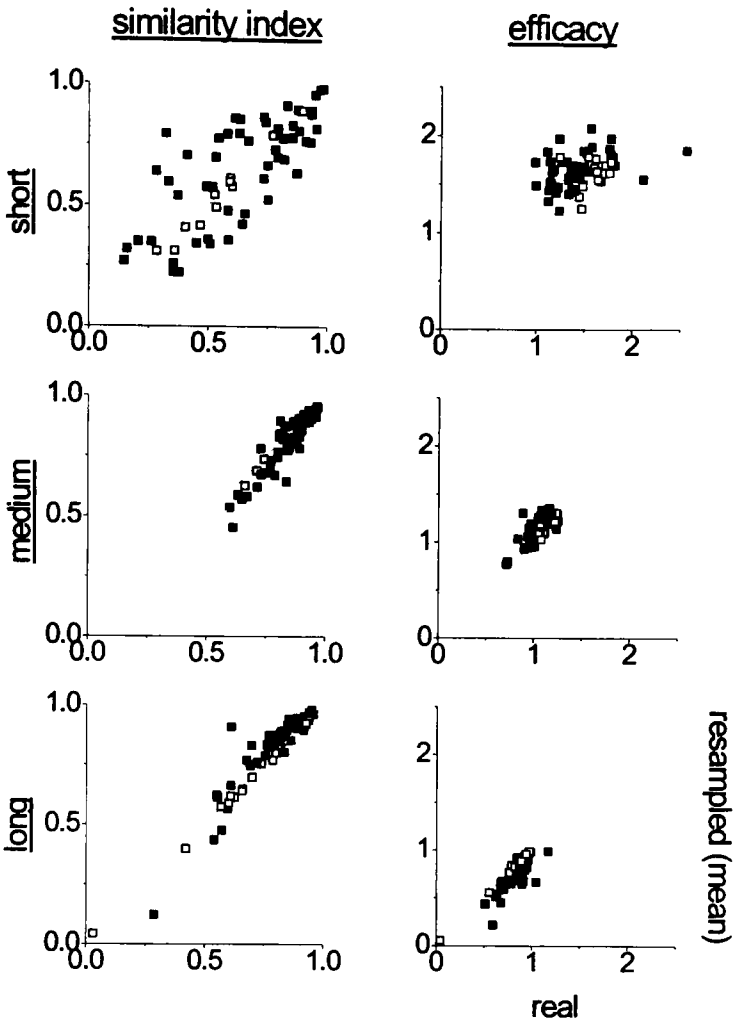

Figure 5.6. *RF map changes: real vs. exchange resampled.* In each panel, the similarity index (left column) and efficacy (right column) for real data are plotted on the horizontal axis, and the mean value of the same parameter derived from 200 exchange resamplings of the real data are plotted on the vertical axis. Data are from 66 neurons, and each point represents a different neuron. The exchange-resampling procedure tests the hypothesis that the observed RF map changes could have been generated by a neuron that fires spikes according to a rate code with exactly the same firing-rate modulation as the original response. Filled squares represent subsets for which the real data are significantly different from the resampled data—that is, points that fall significantly off the diagonal in each panel ($p<0.05$, two-tailed direct comparison).

Information rates. The rates at which stimulus-related information is conveyed by spikes in the full spike train and in each ISI subset are calculated by the *direct method* (see Chapter 6 for a full explanation). Because the direct method relies on the response statistics and does not involve making an explicit model of the stimulus-response relationship, the results of the information calculation are complementary to the RF-map analysis. Whereas

the RF-map analysis presents a picture of the stimulus-encoding characteristics of spikes in the different subsets, the information-theoretic calculations only reveal whether a particular ISI subset conveys more or less information about the stimulus. The method makes limited assumptions about the nature of the neural code, but it does not provide insight into the stimulus features about which information is conveyed.[6]

It is important to point out that because the information calculation is based on spike counts in single short time bins (3.7 to 29.8 msec), the information-rate estimates presented here ignore longer-term correlations in the full response and in the ISI subsets. Obtaining reliable estimates of the full information rates requires far more data than are available, as discussed in Chapter 6. The information rates reported here should be considered instantaneous measures that are not necessarily (and not likely to be) additive over time (DeWeese and Meister, 1999). Thus, an estimated information rate of 20 bits/sec does not actually mean that the neuron transmits 20 bits of information every second.

The results of the information calculation are shown in Fig. 5.7. Not surprisingly, the full spike trains convey the most stimulus-related information per second (top), more than in any ISI subset. Short-ISI spikes, which are rarest, convey information at the slowest rates. On a neuron-by-neuron basis, the median sum of the information rate across the three ISI subsets is larger than the median information rate in the full spike train ($p<0.0001$, Wilcoxon signed-ranks test), a finding that has at least two possible explanations. First, because the calculation does not capture the full amount of information in a spike train, the information rate may be systematically overestimated to a greater extent in the subset spike trains than in the full spike train. Second, in any given neuron, the three subset spike trains may actually convey redundant information, which would render the information rates non-additive across subsets. This is likely to be the case, because the process of selecting a particular ISI subset necessarily places constraints on the other subsets. These constraints and redundancies are not taken into account in the calculation of information rates.

---

[6] In Chapter 6, the direct method is adapted to make possible the calculation of information about specific stimulus attributes, but the procedure is not used in this chapter.

On a bits/spike scale (information rate is divided by firing rate for each ISI subset), short-ISI spikes, which convey the least information per second, can be the most informative (Fig. 5.7, middle). The median short-ISI spike conveys 0.61 bits of information about the stimulus, compared to a median of 0.51 bits/spike conveyed by the full response. Medium-ISI spikes convey approximately the same amount of information per spike as short-ISI spikes (median: 0.65), whereas long-ISI spikes convey less information (0.30). Subset type has a significant effect on information rate ($p<0.0001$, Kruskal-Wallis non-parametric ANOVA).

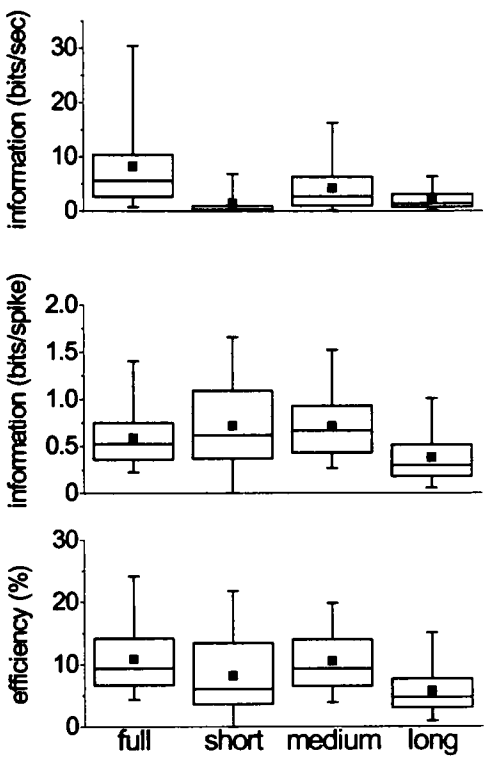

Figure 5.7. *Across-neuron distributions of transmitted information values.* Data from 83 neurons. Box plots as in Fig. 5.5. The full response (leftmost distribution in each panel) is compared with each of the three ISI-subset responses. **Top**: transmitted information, estimated in units of bits/sec. **Middle**: transmitted information, bits/spike. **Bottom**: efficiency, or percentage of the total signal entropy (see Methods) that is used to transmit stimulus-related information.

Bits/spike distributions are similar to the distributions of the efficiency of information transmission. *Efficiency* is defined as information rate divided by total entropy rate. (Ruyter van Steveninck et al., 1997) The median efficiency of all spikes (9.4%) is higher than the median efficiency of short-ISI spikes (6.0%, $p<0.0001$, Wilcoxon signed-ranks test) and long-ISI spikes (4.7%, $p<0.0001$) but indistinguishable from the median efficiency of long-ISI spikes (9.3%, $p=0.28$).[7]

[7] The information-rate estimates presented here are different from those presented in (Reich et al., 2000), which are calculated on the basis of time windows 14.8 msec long, each containing four time bins. Although the information-rate estimates from single time bins, presented in this chapter, are likely to suffer less from small-sample bias (see Chapter 6, Methods), it is nonetheless possible that the qualitative behavior of the four-letter-word information-rate estimates—substantially higher information per spike and efficiency for the short-ISI subset—is closer to reality.

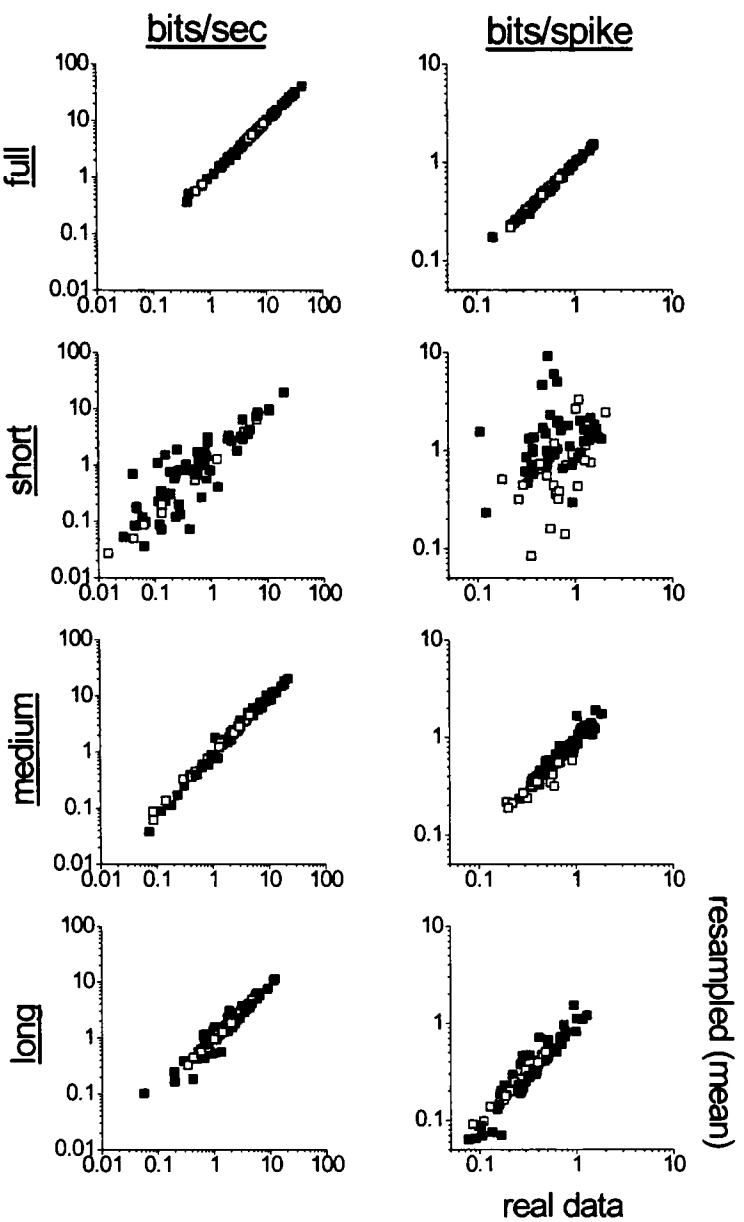

Figure 5.8. *Information: real vs. exchange resampled.* Information rates are plotted in bits/sec (left column) and bits/spike (right column). For each neuron and ISI subset, real data are plotted on the horizontal axis, and the mean value of 40 exchange resamplings on the vertical axis. Data are from 83 neurons. Filled symbols represent subsets for which the real data are significantly different from the resampled data—that is, points that fall significantly off the diagonal in each panel ($p<0.05$, two-tailed direct comparison).

Fig. 5.8 shows the within-neuron comparison between real and exchange-resampled information rates. Each spike train is resampled 40 times (rather than 200, as above, because the computational demands of information calculation are greater than the computational demands of RF-map construction and comparison). The results resemble the results for the similarity index and efficacy (Fig. 5.6). In general, whether measured on a bits/sec or bits/spike scale, and regardless of the particular ISI subset, information rates in real spike trains are similar to information rates in exchange-resampled spike trains: most of the points in Fig. 5.8 lie along the lines of equality. Thus, the qualitative differences in information rates from one ISI subset to the next are expected for neurons that use a rate code, so long as the rate modulations are similar to those in real data. Nevertheless, nearly all points fall significantly off the lines of equality ($p<0.05$, two-tailed direct comparison,

solid squares vs. open squares), indicating that, on a neuron-by-neuron basis, information rates deviate significantly from the predictions of a rate code.

Discussion

The results presented in this chapter provide new evidence for temporal "multiplexing" of stimulus-related information in V1 responses (McClurkin et al., 1991; Victor and Purpura, 1996). Specifically, this means that different spikes within a single response can convey messages about different stimulus features. There are at least three novel aspects of the present results: (1) they identify, through the log-ISIH peaks and in the form of ISI subsets, the relevant time scales in the responses of V1 neurons to rapidly varying stimuli; (2) they provide, through the RF maps, a direct picture of the RF properties of the different spike subset; and (3) they suggest that the information encoded in individual spike times can be decoded by classifying each spike on the basis of the duration of the immediately preceding ISI.

In a pioneering study of the H1 neuron of the blowfly (Ruyter van Steveninck and Bialek, 1988), a Gaussian white-noise stimulus was used to evaluate the probability distributions of stimuli that typically precede arbitrary temporal response patterns consisting of up to three spikes and two ISIs. The authors show that the mean stimulus changes gradually with the duration of the ISI before each spike. Too few data are obtained from each V1 neuron to make possible an identical analysis on the responses presented in this chapter. Although it is likely that the same result holds in monkey V1 as in fly H1—that small changes in ISI boundary points yield small changes in RF maps—the existence and consistency across neurons of the three ISI subsets appears novel and leads to a hypothesis that the grouping of ISIs into subsets similar to the ones described here reflects natural modes of information processing. It is tempting to relate the ISI subsets to phenomena such as oscillatory responses (Gray et al., 1989)—the typical interval in the medium-ISI subset corresponds to a frequency of about 40 Hz—but the results do not provide any evidence that the two phenomena are related.

Bursts. Short-ISI spikes seem to be especially important in the transmission of visual information. These short-ISI spikes likely correspond to the class of spikes known as *bursts* (Connors and Gutnick, 1990), although the correspondence cannot be conclusively established from extracellular recordings alone. In the thalamus, mechanisms of burst production are well understood (Jahnsen and Llinás, 1984; Sherman, 1996), and the relevance of bursts for information encoding and transmission in the lateral geniculate nucleus (LGN), in particular, has been studied (Mukherjee and Kaplan, 1995; Reinagel et al., 1999; Usrey and Reid, 1999). For example, (Reinagel et al., 1999) found that burst spikes can convey 1.5 to 3 times as much information as non-burst spikes.

In the visual cortex, where the mechanisms of burst production are less well understood, several lines of evidence suggest that short-ISI spikes are particularly important for information transmission. Compared to other spikes, short-ISI spikes are differently tuned to certain stimulus attributes (Cattaneo et al., 1981; Legéndy and Salcman, 1985; Livingstone et al., 1996; DeBusk et al., 1997) and tend to be more likely to evoke a postsynaptic response (Alonso et al., 1996; Lisman, 1997; Snider et al., 1998). Moreover, short-ISI spikes are reliable (Victor et al., 1998) in that they are fired at the same time on multiple repeats of a stimulus. The results in this chapter reveal that short-ISI spikes are effective, too: they make a larger-than-expected contribution to RF properties. The RF maps of short-ISI spikes tend to have regions of stark spatial opponency, which implies that these spikes may preferentially extract features such as bars and lines from visual stimuli.

Encoding and decoding. The processes of encoding and decoding information are logically and physiologically distinct. It is possible, for example, that V1 neurons encode information into their rapidly modulated firing rates by means of a Poisson spike generator, consistent with a rate-coding model, even if they decode information by measuring presynaptic ISI durations. In this paper, real responses are compared to exchange-resampled responses, which are considered to be rate-coded because the spike times are determined only from the firing rate. Because exchange resampling exactly preserves spike times from the real data, the firing-rate fluctuations are equivalent to those in the real data.

According to one definition (Théunissen and Miller, 1995; Borst and Théunissen, 1999), information that is carried on time scales more rapid than the time scale of stimulus fluctuations is considered to be *temporally encoded.* The results in this chapter show that temporal encoding is indeed prevalent in V1. But by this definition, even exchange-resampled spike trains are temporally encoded (Fig. 5.8). What distinguishes real and exchange-resampled spike trains most convincingly, then, is not the RF maps or estimated information rates but rather the ISI distributions, as revealed by the log-ISIH. This is consistent with the analysis presented in Chapter 3.

The consistent presence of distinct ISI subsets across neurons, and the different types of visual information that can be extracted by examining spikes from the various subsets in isolation, suggest that ISI decoding may be an important feature of V1 neurons, as it has been shown to be in much simpler systems such as the visceral ganglia of *Aplysia* (Segundo et al., 1963). To accomplish this type of decoding, neurons need not do anything more sophisticated than be sensitive to the durations of individual ISIs. Such sensitivity can be embodied in a single synapse and does not require averaging across stimulus repeats, stretches of time that may be long compared to the time scale of firing-rate modulation, or a large population of neurons that carry similar information.

Although the present results provide support for the hypothesis that ISI decoding plays a role in information transfer in the visual cortex, they are also consistent with other decoding schemes that do not make use of ISIs. One example is the estimation of firing rates by averaging across many neurons that convey similar information (Shadlen and Newsome, 1998) (see Chapter 7). Cortical microstimulation (Salzman et al., 1992) influences both firing rate and ISI structure and is thus consistent with both views of neural coding. In general, ISI-based decoding is more efficient than rate decoding when responses are more regular than Poisson processes (Knight et al., 1979). However, a direct experimental resolution of the roles of different decoding schemes would require manipulating the ISI structure of neural activity without changing the average firing rate, and observing the effects of this manipulation on an animal's behavior. In mammalian cortex, the design and execution of such experiments are challenges to current techniques. But it is interesting to note that such manipulations do indeed affect olfactory discrimination in

the locust (Stopfer et al., 1997) and gustatory perception in the rat (Di Lorenzo and Hecht, 1993).

Real synapses may accomplish ISI decoding by means of processes such as short-term, real-time synaptic modification, including synaptic depression and facilitation (Gerstner et al., 1997; Markram et al., 1998; Goldman et al., 1999), and dendritic nonlinearities, including coincidence detection (Abeles, 1982; Mel, 1994; Margulis and Tang, 1998; Yuste et al., 1999). These processes can selectively weight individual spikes based on the durations of single ISIs (Maass and Zador, 1999), and their particular form can depend on the types of neurons that are connected by each synapse (Thomson, 1997; Reyes et al., 1998). Synaptic facilitation tends to enhance the efficacy of short-ISI spikes, whereas synaptic depression tends to diminish it, thereby increasing the relative efficacy of medium and long-ISI spikes (Gerstner et al., 1997). Thus, a primary role of short-term synaptic modification may be to aid in the decoding of information about multiple stimulus features that would be missed if all spikes were treated equally.

Chapter 6

Formal and Attribute-Specific Information in Primary Visual Cortex

Summary

In this chapter, estimates are made of the rates at which V1 neurons transmit stimulus-related information. Three types of visual stimulus are considered—randomly modulated checkerboard patterns, stationary sinusoidal gratings, and drifting sinusoidal gratings. These stimuli have very different spatiotemporal structures. The overall rate of information transmission, called *formal information*, is obtained by the direct method (Ruyter van Steveninck et al., 1997). The highest information rates are found in the responses of simple cells to drifting gratings (median: 10.3 bits/sec, 0.92 bits/spike); responses to randomly modulated stimuli and stationary gratings transmit information at significantly lower rates. Simple cells tend to transmit information at higher rates, and over a larger range, than do complex cells.

By an extension of the direct method, formal information rates are parsed into *attribute-specific* components, which estimate the rates at which information is transmitted about contrast and spatiotemporal pattern. Contrast-specific information rates vary across neurons—about 0.3 to 2.1 bits/sec or 0.05 to 0.22 bits/spike—but depend little on the stimulus type. Spatiotemporal pattern-specific information rates, however, depend strongly on the type of stimulus and neuron (simple or complex). The remaining information rate, typically between 10% and 32% of the formal information rate for each neuron, cannot be unambiguously assigned to either contrast or spatiotemporal pattern. This indicates that some information concerning these two stimulus attributes is *confounded* in the responses of single neurons in V1. A model that considers a simple cell to consist of a linear spatiotemporal filter followed by a static rectifier predicts higher information rates

than are found in real neurons and completely fails to replicate the performance of real neurons in generating the confounded information.

Introduction

Recent studies of the responses of visual neurons to stimuli with rich temporal structure, such as flickering checkerboard patterns and drifting gratings that abruptly change direction, point to overall information transmission rates of between 5 and 100 bits/sec (Ruyter van Steveninck et al., 1997; Berry and Meister, 1998; Bura as et al., 1998; Reinagel and Reid, 2000). The sensory systems analyzed in these studies range from blowfly lobula plate to primate cortex. The information calculations are based on the direct method (Ruyter van Steveninck et al., 1997; Strong et al., 1998), which estimates the overall rate of information transmission in a set of responses to a single stimulus.

Earlier studies on neurons in primary visual cortex, based on slowly fluctuating stimuli, report information rates an order of magnitude lower (Tolhurst, 1989; Richmond and Optican, 1990; Heller et al., 1995; Victor and Purpura, 1996; Gershon et al., 1998; Mechler et al., 1998b) (Chapter 4). These studies use a variety of methods other than the direct method to calculate the information rates in responses to sets of stimuli that vary along some particular attribute, such as contrast or spatial pattern. All of these methods estimate information as a measure of the degree to which responses can be clustered into the appropriate stimulus classes.

Comparing these and similar results, (Bura as and Albright, 1999) argue that neurons, especially cortical neurons, more effectively convey information about stimuli with rich temporal structure than about stimuli with simpler structure. This argument is incomplete, however, because the two sets of studies take qualitatively different approaches to measuring transmitted information, both in terms of the richness of the stimuli and in terms of the analysis method. It is therefore impossible to draw conclusions about the types of stimulus that evoke the highest information rates from such a comparison. In this

chapter, the results of these two sorts of studies are linked through analysis of the responses of V1 neurons to a battery of stimuli of different spatiotemporal structure. Information is estimated in a uniform fashion, by a variation on the direct method. The major finding is that the overall rate of information transmission—which is dubbed *formal information*—does vary with stimulus type, but that responses to rapidly modulated stimuli do not necessarily convey the most information, particularly in the case of simple cells.

Next, a distinction is drawn between formal and *attribute-specific information rates*. Formal information concerns all aspects of the response that depend on the stimulus. Attribute-specific information concerns only aspects of the response that allow the discrimination between stimuli that differ in some particular attribute, such as contrast, in the face of variation in other attributes, such as spatiotemporal pattern. The attribute-specific information rate is a measure of the degree to which responses to different stimuli cluster according to a particular stimulus attribute and is thus more comparable to the information measured in the second type of study mentioned above. Since each type of stimulus is presented at multiple contrasts, the formal information rate can be parsed into attribute-specific components relating to contrast and spatiotemporal pattern. Here, spatiotemporal pattern refers to a broad category of stimulus attributes that includes temporal fluctuations as well as variations in spatial phase.

Overall, the results indicate that information about contrast is transmitted at a significantly slower rate than information about spatiotemporal pattern, although not for every type of neuron and stimulus. The rates of contrast-specific, but not spatiotemporal pattern-specific, information transmission depend little on stimulus type. Contrast-specific information rates estimated by the direct method are very similar to contrast-specific information rates estimated by a method based on computing the distances between pairs of spike trains (Chapter 4).

Together, contrast- and spatiotemporal pattern-specific information rates account for less than the full formal information rate—typically 68% to 90%. This indicates that a significant portion of the information in V1 responses relates to a *confounded* representation of contrast and spatiotemporal pattern, by which the spatiotemporal pattern of the

stimulus is encoded in a contrast-dependent fashion and the contrast of the stimulus in a spatiotemporal pattern-dependent fashion. An observer who is only aware of the portion of a single neuron's response that contains the confounded information cannot draw conclusions about contrast or spatiotemporal pattern in isolation. The possibility still exists—and is addressed in Chapter 7—that the confounded information may be parsed into individual components by considering the simultaneous responses of additional neurons.

The final section of the chapter asks whether a basic model of a V1 simple cell can account for the results. This model consists of a linear spatiotemporal filter, which is derived from the responses of real neurons, followed by a static rectifier and a Poisson spike generator. The responses of such a model to the same stimuli that are presented to real neurons transmit formal information at rates comparable to, but higher than, those of real responses. However, unlike in the responses of real neurons, all of the formal information in model responses can be parsed into attribute-specific components: the model does not confound the encoding of contrast and spatiotemporal pattern. The results of Chapter 5 indicate that this discrepancy is not due to differences in the underlying dynamics of spike generation, which do not strongly determine information rates. Instead, the discrepancy occurs because this basic model lacks certain features of real cortical neurons, such as contrast gain control (Ohzawa et al., 1982), contrast normalization (Albrecht and Geisler, 1991; Heeger, 1992), and pattern gain control (Carandini et al., 1997a) by which variations in one stimulus attribute can affect the encoding of another.

Methods

The data in this chapter come from recordings of individual neurons in the primary visual cortices of sufentanil-anesthetized macaque monkeys. The detailed experimental procedures are described in Chapters 2-5. Three types of stimulus are used: flickering checkerboards modulated by m-sequences; drifting sinusoidal gratings; and transiently presented, stationary sinusoidal gratings. All stimuli are presented on a Tektronix 608 CRT with a mean luminance of approximately 150 cd/m^2 and a frame rate of 270.329 Hz.

Stimuli. ***M-sequences.*** The principles and methodology of the m-sequence checkerboard stimuli have been extensively described (Sutter, 1992; Reid et al., 1997). In the experiments reported here, two such stimuli are used: a 12^{th}-order m-sequence (4095 stimulus frames) modulating 249 stimulus checks, and a 9^{th}-order m-sequence (511 frames) modulating 25 checks. In both cases, each frame lasts for 14.8 msec (four monitor refreshes), so that the total 12^{th}-order stimulus lasts 60.6 sec and the 9^{th}-order sequence lasts 7.6 sec. Individual checks typically span 16' x 16' of visual angle and are arranged in a square. The size and orientation of the array are sometimes adjusted to better match the neuron's spatial-frequency preferences, but only in cases where the adjustment is expected to produce a dramatically stronger response.

The 249-check ("long") stimulus is surrounded by a black circular aperture, and the 25-check ("short") stimulus is surrounded by a uniform field at its mean luminance. In both stimuli, every check is modulated by the same m-sequence, but the starting point in the sequence varies from check to check. The minimum offset between starting points is 237 msec (64 samples of the m-sequence). The use of m-sequences in this way ensures that there are essentially no pair-wise correlations in time within individual checks, or in space across checks, that are relevant to the neuron's response. The long stimulus is presented at a single contrast (1), and the short stimulus is presented at each of five geometrically spaced contrasts (0.0625, 0.125, 0.25, 0.5, and 1). Both standard and contrast-inverted (reversed dark and light checks) sequences, each repeated 12 to 16 times, are presented in the long stimulus. Standard and inverted sequences within a repeat are separated by a period of 23 sec during which a uniform field at the mean luminance is presented; repeats are separated by 18 sec. For the short stimulus, no inverted sequences are presented. Contrasts are presented in increasing order, separated by uniform-field presentations lasting 10 sec, and the entire of set of contrasts is repeated 25 to 100 times with 10 sec between repeats.

Drifting gratings. "Optimal" drifting sinusoidal gratings with spatial frequency, temporal frequency, and orientation chosen to maximize either firing rate (for complex cells) or response modulation at the driving frequency (for simple cells) (Skottun et al., 1991), are used. For simultaneously recorded neurons, the gratings are optimized for at

least one of the cells, usually the one with the most distinct extracellularly recorded waveform (since this is the neuron most easily monitored during the experiment). The parameter choices based on the response of this cell are likely to be similar to the parameters that would have been chosen for the other simultaneously recorded neurons (DeAngelis et al., 1999), as is occasionally verified empirically. Gratings are presented for 4 sec at each of six geometrically spaced contrasts (0, 0.0625, 0.125, 0.25, 0.5, and 1). The entire set of contrasts is presented five to eight times, with the order of contrast presentation randomized within blocks. For 4 Hz gratings, this yields 80 to 128 stimulus cycles at each contrast. Within each block, grating presentations at different contrasts are separated by presentation of a uniform field of the same mean luminance for 8 sec, and blocks are separated by presentation of the uniform field for 13 sec.

Stationary gratings. Stationary sinusoidal gratings are presented at the same spatial frequency and orientation as the drifting gratings. The spatial phase of the stationary gratings is the one that maximizes the neuron's firing rate in response to stationary gratings of unit contrast. For each neuron, either seven geometrically spaced (0, 0.03125, 0.0625, 0.125, 0.25, 0.5, 1) or nine arithmetically spaced (0, 0.125, 0.25, 0.375, 0.5, 0.625, 0.75, 0.875, 1) contrasts are presented in increasing order. Gratings replace a uniform field of the same mean luminance for a period of 237 msec, after which the uniform field reappears for a minimum of 710 msec. The amount of time between grating presentations increases as a function of the contrast of the preceding grating. For example, the amount of off-time following the 0.5 contrast presentation is 2.84 sec, and following the 0.875 contrast presentation, 4.26 sec. This strategy is used to approximate a uniform state of contrast adaptation (Sclar et al., 1989) prior to the presentation of each different grating. The entire series of contrasts is usually presented 100 times. Only spikes occurring between 30 and 300 msec after stimulus onset are analyzed here.

Information rates. Extensions of the direct method of calculating information rates (Ruyter van Steveninck et al., 1997; Strong et al., 1998) are used to evaluate the responses to all three types of stimulus. This method is based on a comparison of the response variability across time to the response variability across trials. The underlying principle behind the approach is that the portion of variability that cannot be explained by

intrinsic variations in the response to a particular stimulus must represent stimulus-related information.

Direct method. The straightforward application of the direct method, diagrammed in Fig. 6.1A, evaluates what is called the *formal information rate*. The spike train recorded during each trial is divided into time bins, and the spike counts in each bin are tabulated. These spike counts are considered letters in the neuron's response alphabet. Several letters in a row constitute a word, and each word has a probability, possibly stimulus-dependent, of being "spoken" by the neuron. In this paper, one-letter words (single time bins) are chosen because the data sets are not typically large enough to obtain reliable multi-letter-word information estimates. Based on others' results in different systems (Strong et al., 1998; Reinagel and Reid, 2000), as well as on calculations for a limited number of neurons from which large amounts of data were collected, it is estimated that information values are likely to differ by at most 25% in the two cases, depending on the particular

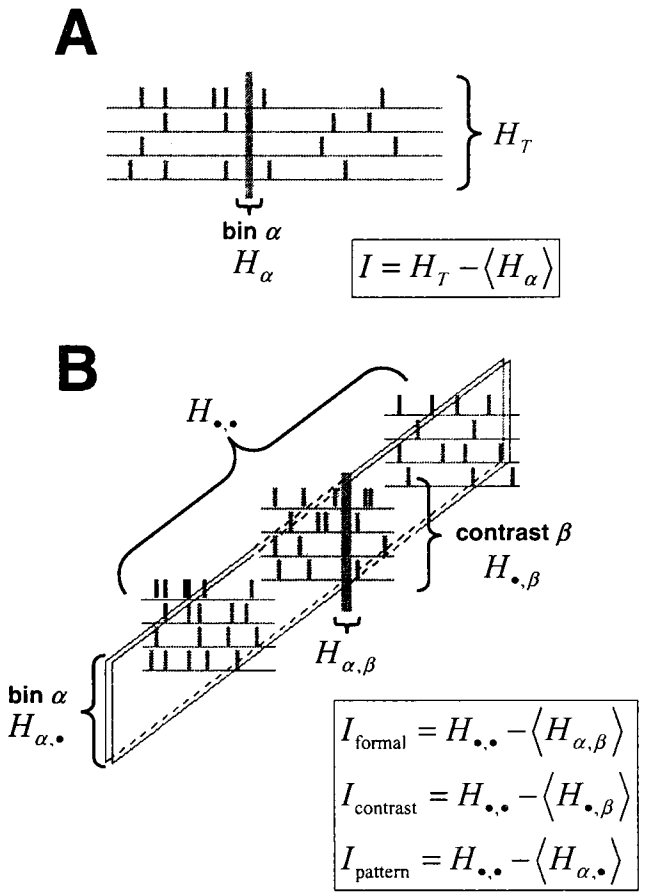

Figure 6.1. *Direct method of information-rate estimation.* Stimulus-related information I is calculated as the difference between the response variability across time (*total entropy*, H_T or $H_{\bullet,\bullet}$) and the response variability across trials or stimulus conditions (*noise entropy*). **A**: straightforward application of the direct method to multiple trials of a single stimulus, in which the noise entropy is calculated as the average entropy in each bin (H_α) (Ruyter van Steveninck et al., 1997). **B**: modification of the direct method for calculation of attribute-specific information in an experiment with multiple contrasts. The total entropy $H_{\bullet,\bullet}$ is unchanged, but the noise entropy depends on the particular type of information rate being estimated (formal or attribute-specific).

neuron and stimulus (estimates from responses to some stimuli may be more affected than estimates from responses to others). However, the qualitative results (formal vs. attribute-specific information and amount of confounded information) presented in this chapter are not likely to change greatly.

Calculations are performed at a variety of bin (letter) sizes, ranging from 0.9 to 59.2 msec. The bin size that yields the highest information rates is chosen. This choice is justified because the actual information rate must increase as the bin size decreases (Strong et al., 1998), even though information-rate estimates may not increase because of limitations set by the amount of available data. It is important to emphasize that the quantities estimated here are information rates for brief samples of the responses (single time bins or letters), not the total information contained in an extended response. In general, the conversion between information rate estimates and information estimates over extended responses is subadditive, in part because the information encoded at different times in the response may be redundant (DeWeese and Meister, 1999).

From the set of binned spike counts, two quantities are extracted. The first, called the *total entropy* (H_T), is a measure of the response variability across time—that is, the uncertainty in spike count across all bins. H_T is calculated from the distribution of spike counts in individual bins, across all time and trials (represented by the light gray rectangle in Fig. 6.1A). A direct estimate of the probabilities p_n that n spikes are observed is obtained directly from the spike count statistics. Shannon's formula (Cover and Thomas, 1991) is then applied to calculate the entropy:

$$H = -\sum_{n=0}^{\infty} p_n \log_2 p_n \tag{6.1}$$

The second quantity extracted is the *noise entropy*, which is a measure of the response variability across trials at a fixed time—that is, the uncertainty in spike count in bins at a particular time relative to the stimulus. Unlike the situation with total entropy, which is derived from all bins taken together, the number of noise-entropy estimates is equal to the number of time bins in a single trial. The set of bin-specific entropies $\{H_\alpha\}$ is obtained from the distribution of spike counts at fixed time bins α across trials (dark gray in Fig. 6.1A) and Shannon's formula. The transmitted information is taken to be the difference

between the total entropy and the averaged noise entropy: $I = H_S - \langle H_a \rangle$. Information values are calculated in bits, which are normalized by the bin size to obtain bits/sec and by the total spike count to obtain bits/spike.

An important caveat of the direct method is that it is only sensitive to fluctuations *within* an analyzed response. If there is little variation in the local spike-count distribution during the course of a response, then the direct method yields a low information rate. Such is the case for complex cells when the stimulus is a drifting sinusoidal grating at fixed contrast. Since complex cells respond to these stimuli primarily by elevating their discharge rates (Skottun et al., 1991), the direct method only detects information if the analyzed response includes both background and stimulus-driven firing, because the appearance of the stimulus causes a change in firing rate. In this chapter, the unit-contrast drifting-grating responses, unlike the unit-contrast stationary-grating responses, do not include any period of background firing.

Thus, the direct method will only yield substantial information rates if the stimulus sufficiently modulates firing patterns in a neuron's response. For this reason, the direct method has, in the past, always been applied to the responses of neurons to rapidly varying stimuli, which typically evoke a wide range of firing rates that fluctuate over time (Ruyter van Steveninck et al., 1997; Bura as et al., 1998; Reinagel and Reid, 2000). Nonetheless, there is no *a priori* reason to limit the application of the direct method to rapidly varying stimuli, and the method can yield useful results even when applied to other sorts of responses. In neither case, however, can the results be taken to be an estimate of the neuron's *channel capacity*, or maximum information transmission rate (Cover and Thomas, 1991); instead, they only reflect the information rates transmitted in response to the specific stimuli considered here.

Attribute-specific information rates. The formal information rate calculated by the direct method evaluates the overall rate of information transfer about all time-varying aspects of the stimulus. It does not evaluate the rates at which information about individual stimulus attributes is transmitted without the confounding influence of other attributes. These latter quantities are referred to as the *attribute-specific information rates*, and

an extension of the direct method makes it possible to calculate them. In this chapter, the focus is on contrast and spatiotemporal pattern, but the idea can be applied to any situation in which two (or more) attributes are varied independently. As used here, *spatiotemporal pattern* is an omnibus term that refers to aspects of the stimulus that do not change as the contrast is varied. Here, the stimuli are completely defined by contrast and spatiotemporal pattern: indeed, each stimulus is defined by a spatiotemporal pattern and a contrast value by which it is multiplied.

As shown in Fig. 6.1B, attribute-specific information is calculated by a procedure that elaborates on the one depicted in Fig. 6.1A. The total entropy $H_{\cdot,\cdot}$ is an overall quantity that represents the uncertainty in spike count across all bins in the entire data set—that is, across time, trials, and contrasts—and is used in both formal and attribute-specific information calculations. When restricted to a single contrast, $H_{\cdot,\cdot}$ reduces to H_T, the total entropy from Fig. 6.1A. (The dot (•) notation is used to denote inclusion of either all time or all contrasts in the information calculations.) To obtain the overall noise entropy used in the calculation of formal information rates, all the estimates of the noise entropy $H_{\alpha,\beta}$, taken from spike counts measured on different trials at fixed times (α) and contrasts (β) (dark gray region), are averaged together. This is a direct extension of the procedure used when only a single stimulus type is presented, as in Fig. 6.1A, in which $H_{\alpha,\beta}$ is equivalent to H_α.

Attribute-specific information rates are obtained by averaging over one or another stimulus attribute, and over trials. Intuitively, by ignoring the value of one stimulus attribute, stimulus variations along the ignored attribute are considered to be a source of "noise." This noise potentially adds variability to the spike count and could reduce the ability of stimulus-induced variations in spike count to transmit information about the non-ignored attribute. To obtain the contrast-specific noise entropy $H_{\cdot,\beta}$, spike counts recorded at contrast β, regardless of time bin or trial number (medium gray), are used to calculate contrast-specific noise entropies that are then averaged across contrast. This essentially represents the uncertainty in spike count at each contrast, averaged across all contrasts. The contrast-specific information is then the difference between the total en-

tropy and the averaged contrast-specific noise entropy: $H_{\bullet,\bullet} - \langle H_{\bullet,\beta} \rangle$. To obtain the spatiotemporal pattern-specific noise entropy $H_{\alpha,\bullet}$, entropies derived from spike counts at time α, regardless of contrast (light gray), are averaged across time bins. This corresponds to the uncertainty in spike count at each time relative to the stimulus, averaged across all such times. The spatiotemporal pattern-specific information is the difference between the total entropy and the averaged spatiotemporal pattern-specific noise entropy: $H_{\bullet,\bullet} - \langle H_{\alpha,\bullet} \rangle$. The sum of the two pattern-specific information rates cannot exceed the formal information, except for measurement errors, and equality can only hold under circumstances in which the two attributes are independently represented. Proof of this statement and further background concerning attribute-specific information can be found in the Appendix.

Strictly speaking, because of the subadditivity of information rates, and because contrast (unlike spatiotemporal pattern) is constant during the course of the stimulus, the notion of a "contrast-specific information rate" is a bit artificial. For example, since only five contrasts are presented in the experiment, the total contrast-specific information must eventually saturate, since it cannot achieve a value larger than $\log_2(5)$=2.32 bits. The contrast-specific information rates reported in this chapter, therefore, should be considered to be typical values in each time bin, even though the actual amount of contrast-specific information that is transmitted may vary from bin to bin. Of course, it would be possible to design a stimulus in which contrast, rather than or in addition to spatiotemporal pattern, is varied dynamically. The present results predict the information rates that would be measured in such an experiment, although the dynamics of the system under study might render this prediction poor.

Bias in the information-rate estimates. Because only a limited amount of data is available, the estimates of the total and noise entropies are both subject to a downward bias. This is a generic property of information estimates from limited data sets (Miller, 1955; Carlton, 1969). Since the transmitted information is the difference between these two entropies, the resulting information rate will be either underestimated or overestimated depending on the relative magnitude of the bias in the two entropy estimates.

When the probabilities of each word are directly estimated from the observed probabilities, an asymptotic estimate of the bias is $(k-1)/2N\ln(2)$, where k is the number of distinct, observed words (here, spike counts per bin) and N is the total number of observations (Panzeri and Treves, 1996; Victor, 2000). Because N is large for the total entropy (number of bins times number of trials), the correction is quite small (for m-sequence responses, about 0.01%). On the other hand, in the calculation of the individual bin-specific noise entropies, N can itself be small (as low as 16 for m-sequences), making the correction term much larger (sometimes on the order of 10% or more).

In many cases, particularly with short bins, only one distinct spike count—zero—is observed. These bins contribute an entropy of 0 to the averaged noise entropy, even with the asymptotic bias correction (because k=1). Because these bins are so common, and because entropy is a logarithmic function of probability, as in equation (6.1), the noise entropy is potentially severely underestimated. This can result in the overestimation of transmitted information. The problem of the zero-count bins is addressed by assuming that the noise entropy varies slowly when the number of spikes is very low. This assumption leads to an *ad hoc* procedure for grouping several consecutive bins together in order to generate a single estimate of the bin-specific noise entropy. Specifically, when a bin with no spikes in any trial is encountered, subsequent bins are considered until one is found that has at least one spike. For these m bins, the noise entropy is calculated as described above, applying the analytic bias correction with N, the number of observations, equal to m times the number of trials. This value of the noise entropy is then assigned to each of the bins that are grouped together in this way. The final value of the noise entropy is again the average of the individual bin-specific entropies, where some of those entropies have been calculated by grouping several bins together. This grouping occurs most commonly in the calculation of the formal noise entropy, for which the number of observations is simply equal to the number of trials. The effective number of trials in the calculation of the pattern-specific noise entropy is higher because time bins are grouped together across contrasts. The effective number of trials in the calculation of the contrast-specific noise entropy is vastly higher because time bins and trials are grouped together at a single contrast, so that bins with only one value of spike count are never encountered.

The assumption that the noise varies slowly when the number of spikes is very low makes it possible to obtain accurate information values for synthetic spike trains that are examples of modulated (inhomogeneous) Poisson processes with fewer than half the number of trials that would otherwise be required (simulations not shown). Indeed, for long stimuli (such as the m-sequences), accurate information estimates can sometimes be obtained with as few as four trials of the stimulus—typically, at least 16 are available. For briefer stimuli (such as the gratings), more trials are required. Because real neuronal responses are not examples of modulated Poisson processes (Chapter 3), an additional criterion is imposed to eliminate data sets that do not contain enough trials. This criterion requires that the formal information rates obtained from half the data, randomly chosen, be within 10% of the information rate obtained from the full data set. This requirement is quite strict and eliminates between 20% and 50% of the neurons recorded with each stimulus type. The neurons that are retained tend to convey information at slightly higher rates.

Finally, in order to obtain an estimate for the scatter of individual information estimates, the jackknife procedure is used (Efron and Tibshirani, 1998). Specifically, for each data set, 1/16 of the trials is sequentially removed, and information rates are recalculated; 1/16 is chosen because only 16 trials are available for some m-sequence data sets. From the resulting distribution of information rates, the standard error of the information rate obtained from the full data set is estimated. The jackknife estimate of the standard error is:

$$\sigma_J = \sqrt{\frac{N-1}{N}\sum_{i=1}^{N}\left[\hat{I}_i - \left\langle \hat{I} \right\rangle\right]^2} \tag{6.2}$$

where N is the number of jackknife samples (16) and $\hat{I}_i$ is the i^{th} jackknife information-rate estimate.

Simple-cell model. To help explain the findings about formal and attribute-specific information rates—in particular, the confounded information rate (see Results)—a simple-cell model is constructed. This model considers a simple cell to consist of a linear spatiotemporal filter the output of which is subject to a static nonlinear rectification

and a Poisson spike generating mechanism (Carandini et al., 1996). Rather than assuming any particular form for the linear filter, first-order kernels derived from responses of real V1 neurons to unit-contrast m-sequence stimuli are used directly. The calculation of these kernels, and their normalizations, are discussed in Chapter 5 and (Sutter, 1992; Reid et al., 1997). The subset of 11 neurons modeled here includes five simple and six complex cells. Complex cells, which often yield significant linear kernels (Chapter 5), are modeled here in exactly the same way as simple cells. However, because only linear kernels are used in the model, and because the model does not include any full-wave rectification (Movshon et al., 1978a), the model neurons derived from complex cells respond like simple cells. Although only neurons with robust linear kernels or receptive-field maps are modeled, it has been verified that these neurons have firing and information rates not significantly different from the firing and information rates of the entire population of simple cells, for unit contrast responses (Kolmogorov-Smirnov test, $p>0.05$).

To derive the parameters of the rectification (threshold and linear gain), the linear response is first predicted by convolving the first-order kernel with the unit-contrast m-sequence stimulus. Then, the constant offset and linear gain are found that, when applied to the predicted linear response histogram, yield the best least-squares fit to the histogram of the observed unit-contrast m-sequence response for the neuron being modeled. The model neurons are presented with the same three stimuli that are presented to real V1 neurons: m-sequences, stationary sinusoidal gratings, and drifting sinusoidal gratings. Kernels and stimuli are binned at 1.8 msec resolution (half of a single display frame for the actual visual stimuli). Because the spatiotemporal patterns in grating stimuli—particularly drifting gratings—change faster than the spatiotemporal patterns in m-sequences (even though such changes are not necessarily detected by the neuron), the gratings are integrated over each m-sequence check and time bin before being presented to the model neurons.

The model yields a response histogram that serves as the modulation envelope of an inhomogeneous Poisson process, which is then used to determine the spike times in each trial. For the purpose of calculating information rates, the response histogram and the Poisson assumption are sufficient, because knowledge of the firing rate in each bin

yields the full spike-count probability distribution in that bin. Thus, for model responses, an *exact* value for the information in one-letter words can be obtained. In practice, these exact information rates are extremely close to the ones obtained by applying the method described above for real data; indeed, this similarity constitutes the primary justification for the application of the bias-correction methods.

Results

Formal information rates. Fig. 6.2 shows the responses of a complex cell in monkey V1 to the three different stimuli used in this chapter, each presented at unit contrast. Response histograms, obtained by averaging the number of spikes across all trials in consecutive 7.4 msec bins and then normalizing by the bin width, are presented atop raster diagrams that show the spike times following stimulus onset in each trial. The top panel presents results from a flickering checkerboard stimulus, in which the time course of contrast modulation in each check is determined by a binary m-sequence (see Methods). The stimulus lasts 60.6 sec and is repeated 14 times; here, only spike times that occur between 32 and 33 sec after stimulus onset are shown.

A striking feature of these responses is that spike firing tends to be clustered at particular times, which presumably follow transient changes in the stimulus. However, since the stimulus changes every 14.8 msec (67.6 times during the displayed one-second period), it is clear that not all stimulus transitions are followed by a consistent change in firing probability—that is, only some of the changes, such as the one that causes firing shortly before 32.5 sec, are effective in driving the neuron. This event-like firing in response to stimuli of this sort has been noted in other species and visual areas as well (Bair and Koch, 1996; Berry et al., 1997; Ruyter van Steveninck et al., 1997). As estimated by the direct method (see Methods), the response to the m-sequence conveys 13.4 bits/sec (0.75 bits/spike) of stimulus-related information, which is within the range reported for neurons in area MT of alert macaque monkeys in responses to similar stimuli (Bura as et al., 1998).

Fig. 6.2 also shows the responses of the same neuron to 100 presentations of a stationary sinusoidal grating of optimal orientation and spatial frequency. The grating appears at time 0 and disappears 237 msec later (times marked by solid vertical lines). The response histogram reveals three distinct firing levels. High firing rates begin abruptly 40 to 45 msec after the stimulus is presented and reappear at about 310 msec, after the grating is removed.[8] In between, the firing rate decays to a lower level that is still higher than the baseline. The high firing rates that occur during the response transients resemble, in terms of peak rate and duration, the brief periods of high firing rate in the m-sequence response. However, the neuron spends a great deal more time firing spikes at the lower rate than at the higher rate. This results in a lower information rate of 4.3 bits/sec (0.14 bits/spike), despite the fact that the mean firing rate is substantially

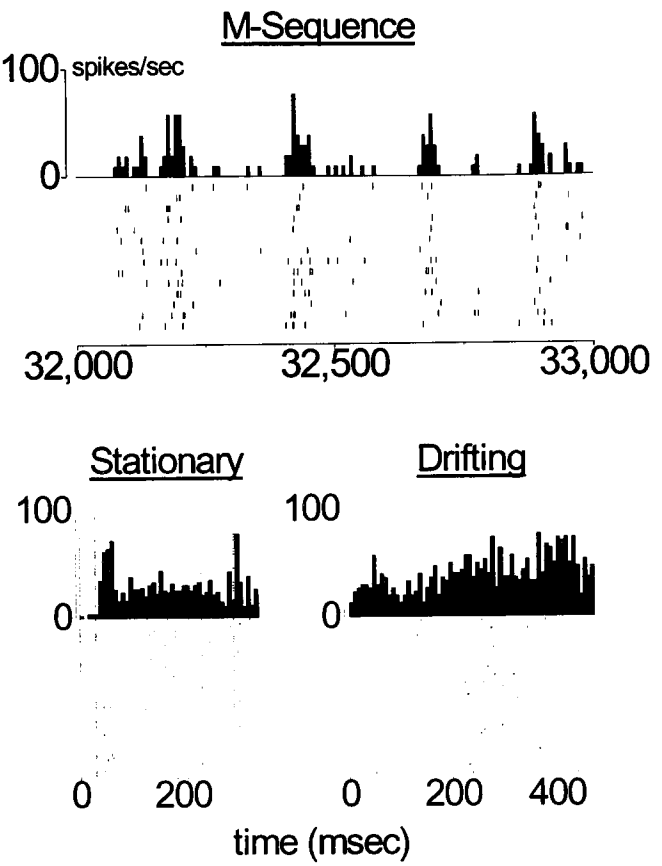

Figure 6.2. *Responses of a representative complex cell to three types of stimulus at unit contrast.* Cell 44/9s. Each panel shows response histograms (7.4 msec bins) atop raster diagrams that depict the spike times fired during each trial. **Top**: flickering binary checkerboard pattern modulated by an m-sequence. The entire sequence lasts 60.6 sec. Only spikes that occur in the 1 sec period between 32 and 33 sec after the start of the stimulus are shown. 14 trials. **Bottom-left**: stationary sinusoidal grating presented at the cell's optimal orientation, spatial frequency, and spatial phase. The stimulus appears at time 0 and is removed 237 msec later; solid vertical lines mark these times. Spikes occurring between 30 and 300 ms after stimulus onset (vertical dashes) are analyzed. 100 trials. **Bottom-right**: drifting sinusoidal grating presented at the cell's optimal orientation, spatial frequency, and temporal frequency. 40 trials.

[8] Information-rate calculations on stationary-grating responses in this chapter only consider spikes that occur from 30 to 300 msec following stimulus onset (dashed lines), meaning that the off response of this neuron is effectively ignored.

higher than in the m-sequence response (29.8 vs. 10.5 spikes/sec). Although both total and noise entropy are higher in the stationary-grating response, the noise entropy, which reflects the spike count variability across trials at particular times, is subject to a proportionately greater increase than the total entropy. Thus, the variability in firing is larger for the grating response than for the m-sequence response, and the stimulus-induced modulation is relatively small.

The third stimulus is a sinusoidal grating, again presented at the optimal orientation and spatial frequency, that drifts at 2.1 Hz. For this neuron, the responses to 40 cycles of the drifting grating were recorded. The analysis treats individual cycles of the grating as separate stimulus trials. As is the case for most complex cells, the response to the drifting grating is only weakly modulated (Skottun et al., 1991), and the most prominent feature is an elevation of the mean firing rate (compare the average response level in the bottom-right panel to the response level just after stimulus onset in the bottom-left panel). There are no periods of abruptly increased firing rate because, unlike the m-sequence and stationary-grating stimuli, the drifting-grating stimulus contains no temporal transients. The information rate in the drifting-grating response is 4.4 bits/sec (0.11 bits/spike). This is in the top 15% of information rates measured in bits/sec in the sample of 43 complex-cell responses to drifting gratings, but near the median information rate measured in bits/spike. The firing-rate elevation that is such an important feature of the responses of complex cells to drifting sinusoidal gratings does not contribute to the information rate calculated by the direct method, which reflects only reproducible modulations in the spike count probability during the course of the response. Such slow modulation, superimposed on a relatively high overall firing rate of 41.5 spikes/sec, is evident in the response histogram.

Fig. 6.3 and Table 6.1 summarize the results of similar experiments performed on the population of recorded V1 neurons, all at unit contrast. Panels A-F summarize across-neuron results, and panel G shows within-neuron comparisons for the neurons that convey information at non-zero rates in response to at least two of the stimuli. As mentioned in Methods, ignoring long time correlations in the information-rate estimates may affect estimates from some stimuli more than from others; as such, the reported values must be

considered to be estimates of typical information rates in brief time bins and not of the actual information rates in the responses.

The results reveal that differences between simple and complex cells are most pronounced in the responses to drifting gratings, which evoke the highest information rates in simple cells but the lowest information rates in complex cells. For simple cells (Fig. 6.3A-C), the unexpected finding is that information rates are higher in drifting-grating responses than in m-sequence responses (filled triangles in Fig. 6.3G are typically above the line of equality). This is evidence against the hypothesis that high information rates in V1 neurons are more likely to be evoked by stimuli that change rapidly in time than by stimuli that change slowly, as might be the case in the motion-sensitive area MT (Bura as et al., 1998; Bura as and Albright, 1999). Complex-cell responses to drifting gratings behave differently: the open triangles in Fig. 6.3G are typically below

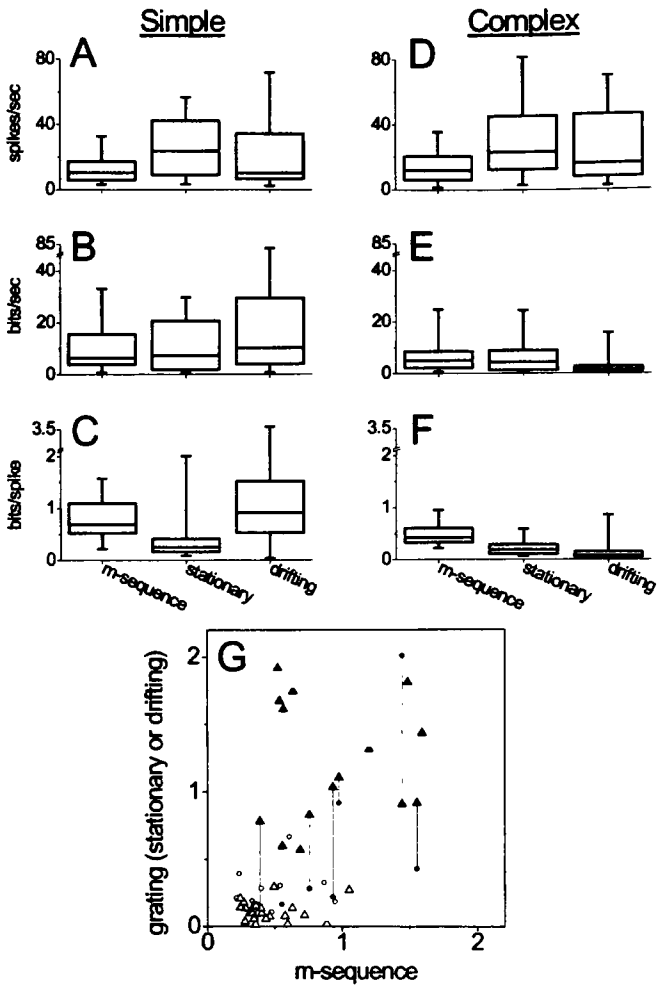

Figure 6.3. *Summary across neurons of firing and information rates in responses to three types of stimulus.* All stimuli are presented at unit contrast. Median values and population sizes are given in Table 6.1. **A-C**: simple cells. **D-F**: complex cells. **A, D**: mean firing rate, spikes/sec. **B, E**: information rate, bits/sec. **C, F**: information rate, bits/spike. Boxes represent the 25% to 75% range of the data, whiskers represent the 5% to 25% and 75% to 95% range, and horizontal lines represent the medians. **G**: information rates (bits/spike) for neurons that convey significant amounts of information in response to at least two stimuli. *Filled symbols*: simple cells. *Open symbols*: complex cells. *Circles*: stationary gratings. *Triangles*: drifting gratings. Lines connect points corresponding to neurons that convey significant amounts of information in response to all three stimulus types. Data from individual cells, plotted in **G**, follow the population trends evident in **A-F**.

STIMULUS	TYPE	*N*	MEDIAN		
			SPIKES/SEC	BITS/SEC (*RE*)	BITS/SPIKE
m-sequence	simple	11	10.38	6.41 (0.038)	0.69*
	complex	72	11.97†	4.99 (0.053)‡	0.42§
stationary	simple	12	23.69	7.29 (0.16)	0.25*
	complex	26	23.15†	4.38 (0.20)‡	0.19§
drifting	simple	25	10.14	10.28 (0.13)	0.92*
	complex	43	16.16†	1.58 (0.33)‡	0.08§

Table 6.1. *Summary of unit-contrast firing and information rates.* Medians of distributions plotted in Fig. 6.3. Median values that differ significantly between simple and complex cells are shaded gray ($p<0.05$, Mann-Whitney U test). Medians of distributions that vary significantly with stimulus type are denoted by symbols (*, †, ‡, §; $p<0.05$, Kruskal-Wallis non-parametric ANOVA). *N*: number of neurons in data set. *RE*: relative error, or standard error derived from jackknife resampling (see Methods) divided by information rate from the full data set.

the line of equality. As discussed above, this is because complex cells, unlike simple cells, transmit very little information about the aspect of the drifting-grating stimulus—its spatial phase—that varies during the course of the response. Table 6.1 shows that stimulus type has a significant effect on firing and information rates for complex cells, but only on information rates in bits/spike for simple cells. However, it is likely that the lack of significance for simple cells is simply due to the smaller number of simple cells in the sample.

Attribute-specific information rates. The direct method is usually applied only to neurons' responses to rapidly modulated stimuli, such as the unit-contrast m-sequence responses in Figs. 6.2 and 6.3. Most stimulus sets used in neurophysiology experiments can be classified along two (or more) attributes: one or more features, such as contrast and spatial phase, that are explicitly varied from one stimulus presentation to the next; and the time course of the stimulus itself. For responses to these sorts of stimuli, straightforward implementation of the direct method measures only the overall rate of information transmission, which in this chapter and the next is called the *formal information rate*, and not information rates for individual stimulus features, which are called *attribute-specific information rates*. Earlier methods, based on firing rates (Tolhurst, 1989), principal components (Optican and Richmond, 1987), stimulus reconstruction (Bialek et al., 1991), and time structure of individual responses (Victor and Purpura, 1996; Panzeri and Schultz, 2000), are expressly designed to measure attribute-specific information.

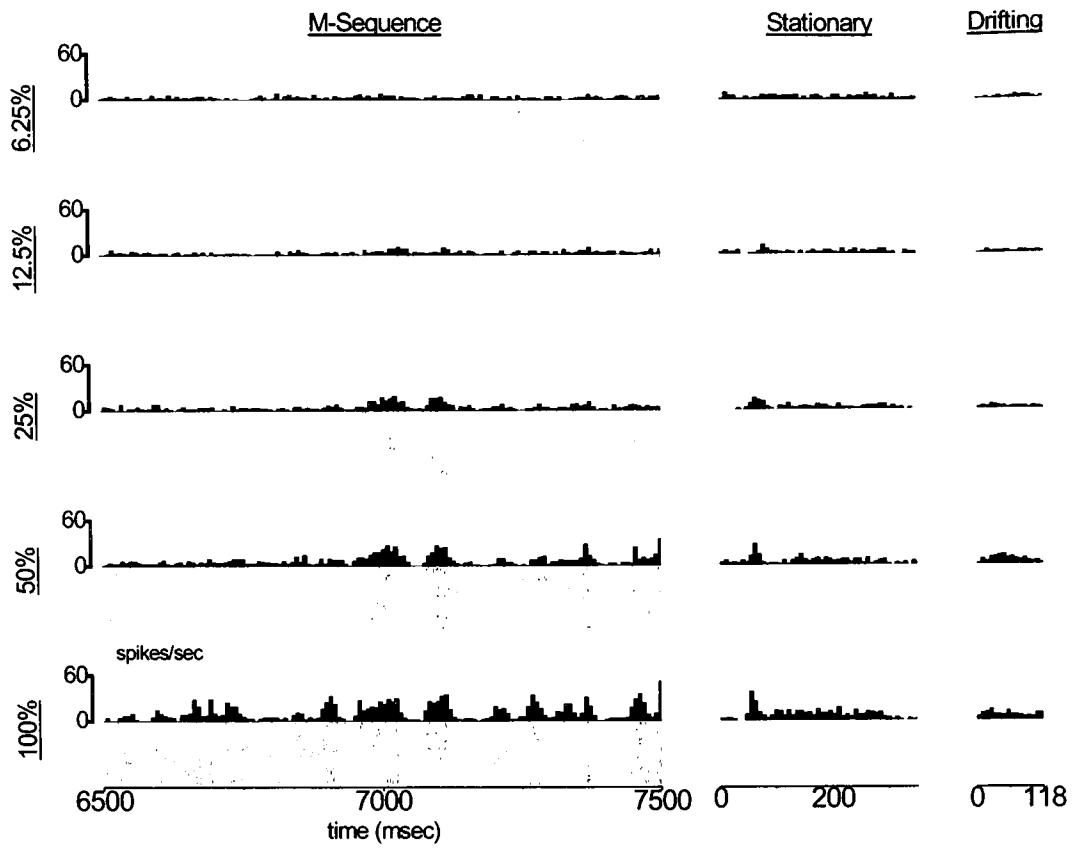

Figure 6.4. *Responses of a representative complex cell to three types of stimulus at five contrasts.* Raster diagrams and response histograms (7.4 msec bins). Cell 45/2s. Contrasts 0.0625, 0.125, 0.25, 0.5, and 1. **Left column**: 1 sec snippets of the responses to 25 presentations of a 7.6 sec m-sequence checkerboard stimulus. **Middle column**: responses to 100 presentations of a stationary-grating stimulus, showing the first 350 msec after stimulus onset. **Right column**: responses to 192 cycles of a sinusoidal grating drifting uniformly at 8.4 Hz.

The experiments in this chapter focus on contrast and spatiotemporal pattern. Formal information rates are parsed into components specific to these stimulus attributes, as described in Methods. Responses to five contrasts (0.0625, 0.125, 0.25, 0.5, and 1) are analyzed. The m-sequence stimulus used in the contrast experiments is shorter, running for 7.6 sec (instead of 60.6 sec) and modulating 25 checks instead of 249.

Fig. 6.4 shows the responses (raster diagrams and histograms) of another complex cell, which has a maintained discharge (response to a uniform field at the mean luminance) of 2.4 spikes/sec. For all three types of stimulus, the responses generally become more reproducible as contrast increases. Some features, such as the spikes that occur

around 7100 msec in the m-sequence response, or the transient firing-rate elevation at the beginning of the stationary-grating response, are particularly reproducible and give rise to peaks in the histograms.

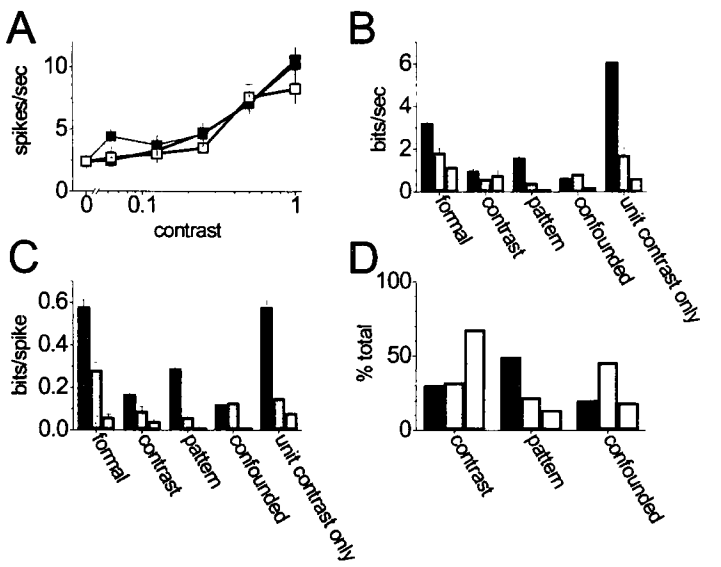

Figure 6.5. *Firing and estimated information rates for the responses of a representative complex cell.* Data are for the neuron from Fig. 6.4. **A**: mean firing rate (spikes/sec) as a function of contrast: m-sequences (solid), stationary sinusoidal gratings (gray), and drifting sinusoidal gratings (white). Error bars represent two standard errors of the mean. **B**: information rate (bits/sec). **C**: information rate (bits/spike). Error bars in **B** and **C** are estimates of the standard error derived from jackknife resampling (see Methods). **D**: information rate (% of total). All five contrasts are used for these calculations, except for the bars labeled *unit contrast only* in **B** and **C**.

Fig. 6.5A plots the neuron's firing rate as a function of contrast for all three stimulus types. Clearly, the curves are quite similar, to within the error bars (95% confidence limits of the mean). As shown in Figs. 6.5B-C, the formal information rate is 3.2 bits/sec (0.58 bits/spike) in the m-sequence response, 1.8 bits/sec (0.28 bits/spike) in the stationary-grating response, and 1.2 bits/sec (0.06 bits/spike) in the drifting-grating response. Responses to lower contrasts can contribute extra information as well as extra variability, so that including them in the information calculations can both raise and lower formal information rates. In this example, including those contrasts decreases the formal information rate measured in bits/sec (but not bits/spike) in the m-sequence responses (black bars), increases the formal information rate measured in bits/spike (but not bits/sec) in the stationary-grating responses (gray bars), and leaves virtually unchanged the formal information rates in the drifting-grating responses (white bars).

To isolate the relative amount of information transmitted about different aspects of the stimuli, the direct method is modified by changing the definition of noise entropy while leaving unchanged the definition of total entropy (see Methods). For these stimuli,

information is conveyed about contrast and spatiotemporal pattern. Spatiotemporal pattern refers to aspects of the stimuli that affect the response variation across time at a fixed contrast. The results are displayed in Fig. 6.5B-C, together with error estimates derived from jackknife resampling (see Methods).

The information rates due to contrast and spatiotemporal pattern alone do not sum to the full formal information rate, as they would if the two stimulus attributes were encoded independently (see Appendix). Instead, for this neuron, the sum of the two attribute-specific information rates fails to account for 19% to 46% of the formal information rate, depending on stimulus type. The information not accounted for by the attribute-specific information rates is called *confounded*. The presence of confounded information means that the dynamics of contrast- and spatiotemporal pattern-encoding are interdependent. The confounded information cannot be used to determine either contrast or spatiotemporal pattern based on the response of this neuron alone.

The results of the contrast and spatiotemporal pattern analysis over the population of neurons are shown in Fig. 6.6 and Table 6.2. Here, data from simple and complex cells are combined for m-sequence and stationary-grating responses because no significant differences among the information-rate distributions were found, probably due in part to the limited number of simple cells in the sample. Overall, as with unit-contrast responses, drifting gratings evoke the highest formal information rates in simple cells but the lowest formal information rates in complex cells. The same is true for spatiotemporal pattern-specific information rates.

The strikingly high formal and spatiotemporal pattern-specific information rates in simple-cell drifting-grating responses are explained by the fact that simple cells are exquisitely sensitive to spatial phase (Hubel and Wiesel, 1962; Movshon et al., 1978b; Victor and Purpura, 1998), which is the only aspect of the stimulus that varies at fixed contrast. The spatial-phase variation causes the firing rate to be deeply modulated, which results in high spatiotemporal pattern-specific, and hence formal, information rates. This is emphatically not the case for complex cells, as discussed above. On the other hand, simple and complex cells transmit the same amount of contrast-specific information in

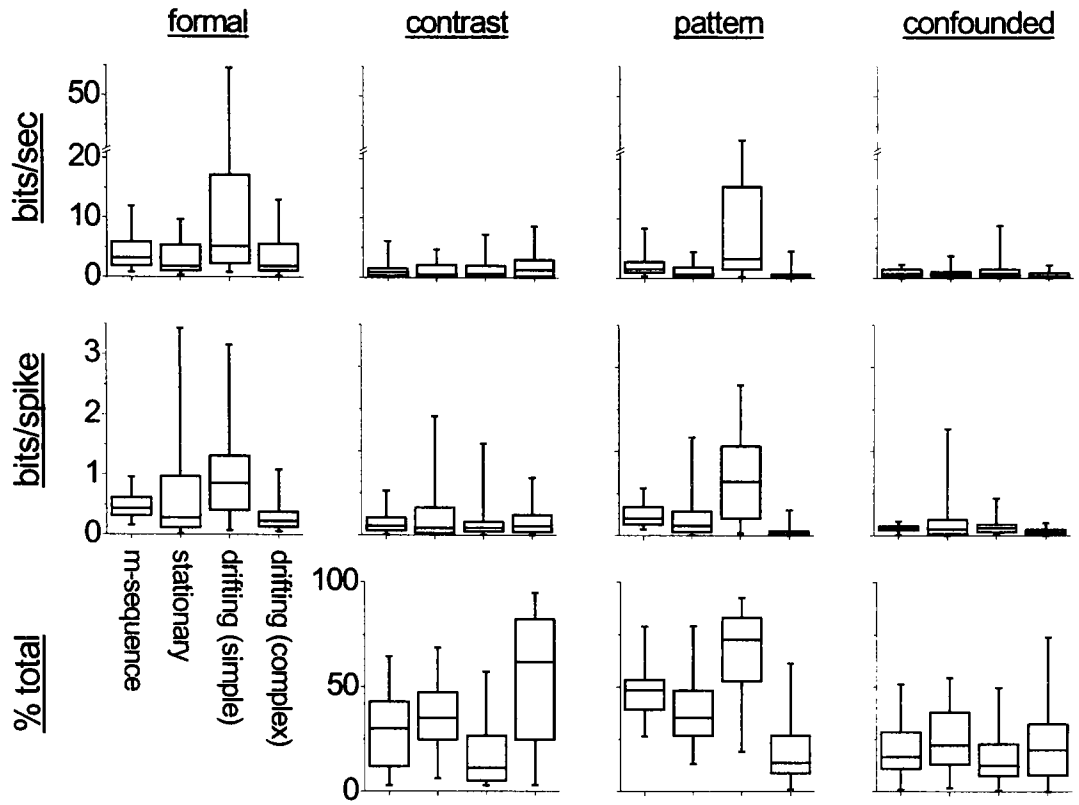

Figure 6.6. *Summary across V1 neurons of formal and attribute-specific information rates.* Data taken from 44 neurons (m-sequences); 50 neurons (stationary gratings); 31 simple cells and 118 complex cells (drifting gratings). Data from simple- and complex-cell responses to m-sequences and stationary gratings are combined because their distributions are not significantly different from one another. Box plots as in Fig. 6.3. **Top row**: information rate, bits/sec. **Middle row**: information, bits/spike. **Bottom row**: attribute-specific information rate as a percentage of the formal information rate. **First column**: formal information rate. **Second column**: contrast-specific information rate. **Third column**: spatio-temporal pattern-specific information rate. **Fourth column**: confounded information rate.

response to drifting gratings. Indeed, the contrast-specific information-rate distributions are relatively independent of stimulus type, measured either in bits/sec or bits/spike (p>0.05, Kruskal-Wallis ANOVA).

For the neuron presented in Figs. 6.4 and 6.5, the attribute-specific information rates do not account for all the formal information in the neuron's response. This is also the case for the population results (Fig. 6.6, right column). The median confounded information—the portion of information that cannot be resolved into contrast- or spatio-temporal pattern-specific components—represents a substantial fraction of the formal in-

STIMULUS	TYPE	*N*	MEDIAN INFORMATION RATE										
			BITS/SEC				BITS/SPIKE				%		
			F	*C*	*SP*	*CNF*	*F*	*C*	*SP*	*CNF*	*C*	*SP*	*CNF*
m-sequence	all	35	3.23	0.83	1.50	0.68	0.46	0.14	0.21	0.086	30.3	45.8	16.6
stationary	all	29	1.84	0.59	0.63	0.74	0.28	0.087	0.12	0.08	35.2	35.2	22.1
drifting	simple	27	5.18	0.69	3.31	0.76	0.85	0.085	0.64	0.092	11.3	72.6	12.5
	complex	57	1.86	1.31	0.30	0.34	0.23	0.11	0.028	0.045	62.0	13.8	20.0

Table 6.2. *Summary of five-contrast formal and attribute-specific information rates.* Medians of distributions plotted in Fig. 6.6. Simple and complex cell results were indistinguishable for m-sequence and stationary-grating responses and are therefore combined. Medians of distributions that vary significantly with stimulus type are shaded gray ($p<0.05$, Kruskal-Wallis non-parametric ANOVA, simple and complex-cell drifting-grating responses separately compared with m-sequence and stationary-grating responses). *N*: number of neurons in data set; *F*: formal; *C*: contrast-specific; *SP*: spatiotemporal pattern-specific; *CNF*: confounded.

formation. Indeed, across all neurons and stimulus types, the confounded information rate typically accounts for 10% to 32% (interquartile range) of the formal information rate.

Information rates in a model V1 simple cell. The model described here considers the responses of a V1 simple cell to derive from a linear spatiotemporal filter followed by a static rectifier and a Poisson spike generator (see Methods). The shape and size of the filter, as well as the parameters of the rectifier, are drawn from the responses of actual neurons to the long m-sequence stimulus at unit contrast. Similar models have been used in the past to describe the responses of visual neurons to various kinds of stimuli (Carandini et al., 1996), even though it is well known that V1 neurons, even simple cells, display many features that are not captured in the model, including nonlinearities of spatiotemporal summation (Movshon et al., 1978a) and contrast response (Albrecht and Hamilton, 1982).

Fig. 6.7 shows the responses of a real simple cell (top row) and its corresponding model neuron (bottom row) to the three types of stimulus, each presented at unit contrast. Despite the rudimentary nature of the model, it successfully captures many of the features of the real data. In particular, the model replicates the location of the peaks in the m-sequence response and the existence of a transient period of elevated firing rate at the beginning of the stationary-grating response. Notable differences, especially in the grating responses, do exist. For example, the transient portion of the stationary-grating response

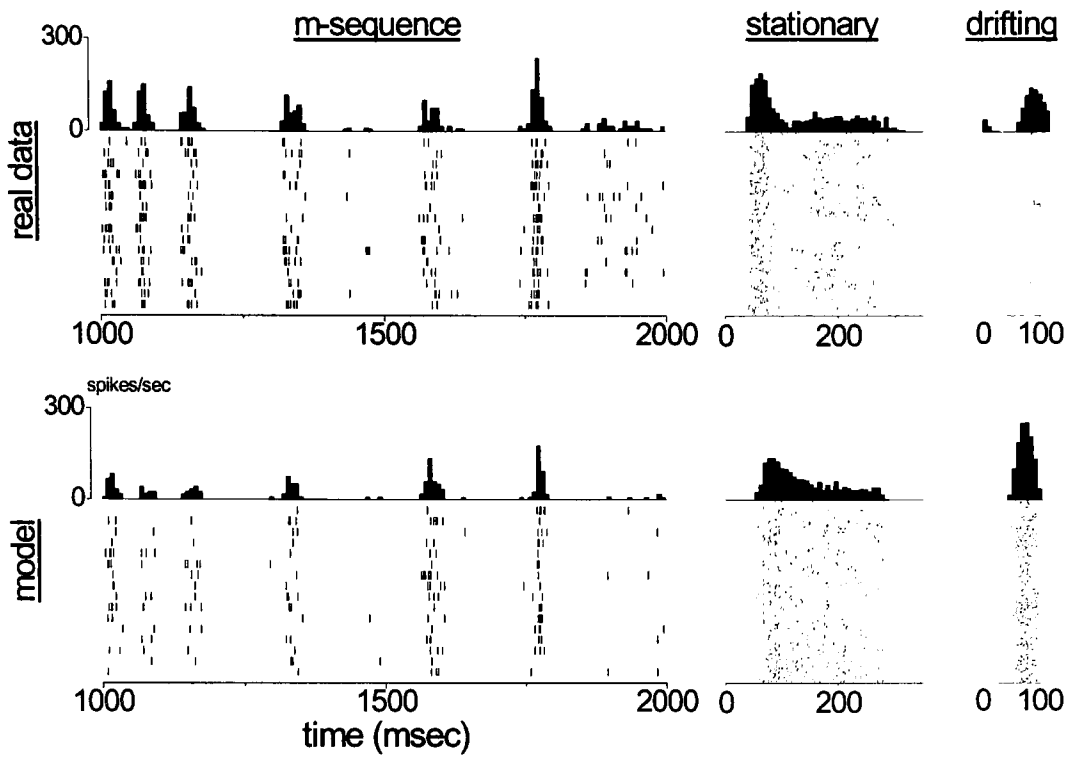

Figure 6.7. *Comparison of real and model simple-cell responses.* Cell 35/1s, unit-contrast stimuli. Model parameters are derived from the m-sequence response of the real neuron. The model consists of a linear spatiotemporal filter, a static rectifier, and a Poisson spike generator (see Methods). Each panel shows response histograms (7.4 msec bins) atop raster diagrams. **Left**: flickering binary checkerboard pattern modulated by an m-sequence. The entire sequence lasts 60.6 sec, and only the spikes that occur in the 1 sec period between 1 and 2 sec after the start of the stimulus are shown. 16 trials. **Middle**: stationary sinusoidal grating presented at the cell's optimal orientation, spatial frequency, and spatial phase. The stimulus appears at time 0 and is removed 237 msec later. 100 trials. **Right**: drifting sinusoidal grating presented at the cell's optimal orientation, spatial frequency, and temporal frequency (8.4 Hz). 256 trials.

is briefer in the real data than in the model, and the drifting-grating response is narrower in the model than in the real data. Moreover, the distinct, brief period of very low firing rate that immediately follows the real neuron's transient response to the stationary grating is absent from the model's response. These differences are not surprising given that the model's parameters are fit to the linear part of the m-sequence response, and not to the grating responses.

The similarities and differences between real and model responses are also reflected in the rates of information transmission. The m-sequence response of the real neuron transmits information at a rate of 15.5 bits/sec (1.6 bits/spike), compared to 19.2 bits/sec (2.5 bits/spike) for the model. In response to the stationary and drifting-grating

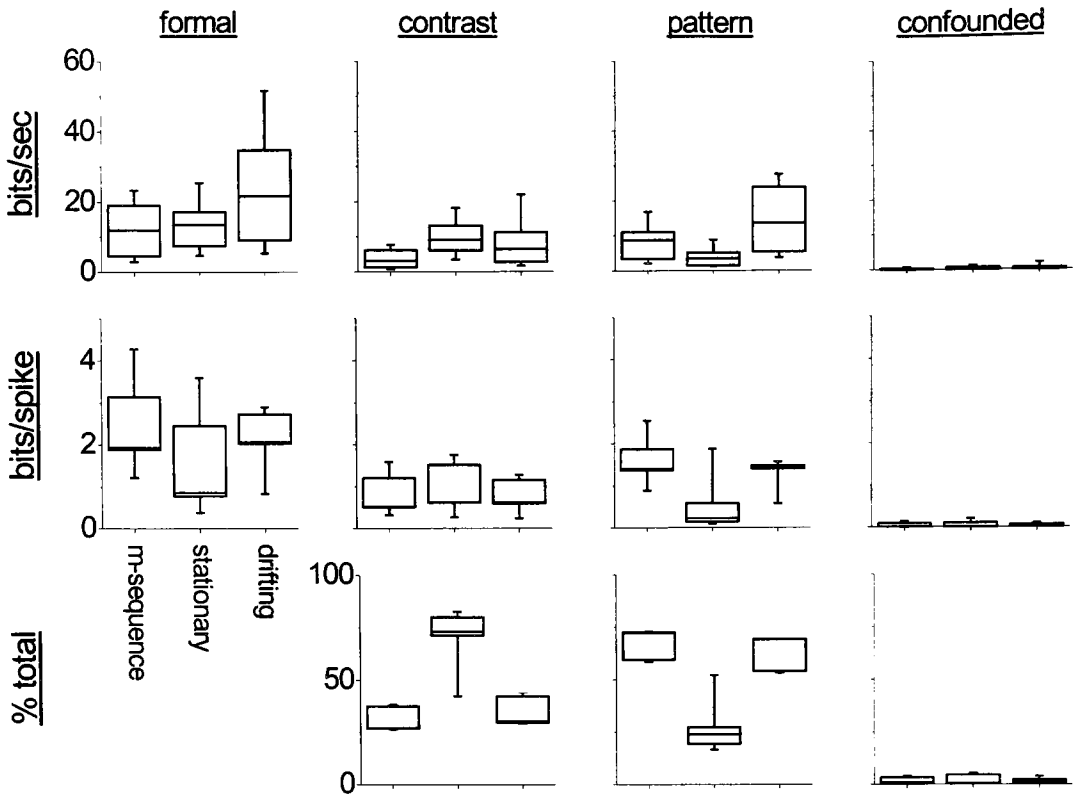

Figure 6.8. *Summary across 11 model neurons of information rates in responses to three types of stimulus.* Details as in Fig. 6.6.

stimuli, the real neuron transmits 24.2 bits/sec (0.43 bits/spike) and 49.4 bits/sec (0.92 bits/spike) of information, respectively, whereas the model neuron transmits 20.8 bits/sec (0.42 bits/spike) and 90.8 bits/sec (1.2 bits/spike). It should be noted that the information in this real neuron's response to the drifting grating is at the top of the range of such information rates, even among simple cells (Fig. 6.3).

Fig. 6.8 and Table 6.3 show that these results generalize to the population of 11 neurons modeled here; the results should be compared with Fig. 6.6 (noting the sometimes-different vertical scales) and Table 6.2. Typically, the information rates of model responses are higher than the information rates of real responses. As is the case with the simple cell modeled in Fig. 6.7, and with real simple cells, the model neurons convey the most information about drifting gratings. The low pattern-specific information rates in the modeled stationary-grating responses correspond to the low pattern-specific information rates in the responses of real neurons to these stimuli. However, unlike in real neurons,

STIMULUS	MEDIAN INFORMATION RATE										
	BITS/SEC				BITS/SPIKE				%		
	F	*C*	*SP*	*CNF*	*F*	*C*	*SP*	*CNF*	*C*	*SP*	*CNF*
m-sequence	13.1	3.5	9.5	0.10	1.9	0.52	1.4	0.015	27.1	72.4	0.77
stationary	13.6	9.1	3.5	0.077	0.86	0.63	0.23	0.004	73.2	26.5	0.55
drifting	21.7	6.4	13.6	0.23	2.1	0.62	1.4	0.022	30.1	69.2	1.3

Table 6.3. *Summary of five-contrast formal and attribute-specific information rates for 11 model neurons.* Medians of distributions plotted in Fig. 6.8. Compare to the data in Table 6.2. Medians of distributions that vary significantly with stimulus type are shaded gray ($p<0.05$, Kruskal-Wallis nonparametric ANOVA). Abbreviations as in Table 6.2.

the contrast-specific information conveyed by the model neurons does depend significantly on the type of stimulus.

The most striking difference between real and model responses is in the confounded information rate, which is nearly zero in model responses. In model responses, contrast and spatiotemporal pattern can be independently determined from the response time course and depth of modulation. The prominence of the confounded information in real responses suggests that an interaction between the coding of contrast and spatiotemporal pattern constitutes one of the primary differences between real and model neurons.

Discussion

The experiments described in this chapter are designed to evaluate the ways in which the spatiotemporal features of different visual stimuli affect the rates at which V1 neurons transmit information. The first stimulus is based on pseudo-random m-sequences and appears as a rapidly modulated checkerboard pattern, in which the spatial pattern changes every 14.8 msec. The second stimulus is a sinusoidal grating at the cell's optimum orientation, spatial frequency, and spatial phase; it appears abruptly and is removed 237 msec later. The third stimulus is a sinusoidal grating that drifts at the cell's optimal temporal frequency and has the same orientation and spatial frequency as the stationary grating. V1 simple cells typically transmit information at the highest rates in response to high-contrast, drifting-grating stimuli. The same stimuli evoke the lowest information rates in complex-cell responses because those responses are not modulated in time.

Do cortical neurons transmit information at high or low rates? The results presented here are surprising in light of the fact that a comparison of information rates in a variety of neural systems suggests that stimuli that change rapidly in time drive neurons to encode information at rates often more than an order of magnitude higher than the corresponding rates for slowly changing stimuli (Bura as and Albright, 1999). Based on the formal and attribute-specific information rates calculated for V1 neurons in response to different types of stimulus, it is likely that the major cause of this discrepancy is not simply the rate of variation of the stimulus, as some have argued (Bura as and Albright, 1999). Moreover, the cause does not lie in the number of transient changes in a stimulus, because such transients are absent from drifting gratings (which evoke the highest information rates in simple cells) but present in both m-sequences (which evoke the highest information rates in complex cells) and stationary gratings.

Instead, the results suggest that the magnitude of a measured information rate has a complicated dependence on the type of attribute-specific information that is being calculated and the sensitivity of the neuron under study to that stimulus attribute. For example, a stimulus that consists of rapid changes in spatial phase, such as a drifting grating, evokes high spatiotemporal pattern-specific information rates in simple cells and low spatiotemporal pattern-specific information rates in complex cells, but a stimulus that consists of rapid changes in luminance, such as an m-sequence checkerboard, evokes indistinguishable contrast-specific information rates in simple and complex cells.

Information rates and channel capacity. The information rates for complex cells can be compared to estimates of the channel capacities of complex cells in the supragranular layers of alert monkeys, which range from 6.7 to 8.5 bits/sec (Wiener and Richmond, 1999). The channel capacity is a measure of the maximum information rate that a communications channel can transmit. In response to the stimuli used here, which differ from the stimuli used by (Wiener and Richmond, 1999) to estimate channel capacity, complex cells transmit information at approximately half of this channel capacity. Since the stimuli in this chapter are not expressly designed to evoke information rates that approach the channel capacity, this result seems rather impressive.

Relevance of confounded information for visual processing. The use of the direct method to calculate both formal and attribute-specific information rates reveals a hitherto overlooked aspect of the information conveyed by V1 neurons. A substantial fraction of the formal information rate (typically 10% to 32%) cannot be attributed to either contrast or spatiotemporal pattern alone. This portion of the information, which is called *confounded*, arises from an interdependence of contrast and spatiotemporal pattern in generating neuronal responses. Confounded information is not present in the model responses, where there is no such interdependence (see Appendix). In other words, the amount of confounded information quantifies the effects of changes in the spatiotemporal profile of the stimulus on the contrast response and sensitivity functions of V1 neurons (Maffei and Fiorentini, 1973; Tolhurst and Movshon, 1975; Albrecht, 1995; Gawne et al., 1996b). Such changes may be mediated by variations in the "adaptive state" of a neuron under different spatiotemporal stimulus conditions, particularly between rapidly modulated stimuli, like the m-sequence checkerboard, and gratings (Gaska et al., 1994).

Another way in which contrast and spatiotemporal pattern can potentially interact is through a refractory period that is intrinsic to a neuron's spike generating mechanism but influences the responses to both stimulus attributes. However, it is not likely that refractory periods contribute significantly to the confounded information reported here. This is because exchange resampling (see Chapters 3 and 5), which effectively eliminates the refractory period while preserving the overall rate modulation and distribution of spike counts per trial, does not reduce, and may even increase, the amount of confounded information. Whatever its basis, the finding that a substantial portion of the total information is confounded means that downstream neurons cannot use all of the information provided to them to draw conclusions about either one of those stimulus attributes in isolation. But perhaps the task of the visual system is not simply to decompose stimuli into components relating to contrast and spatiotemporal pattern. For example, if it were possible to calculate attribute-specific information along the visual system's preferred axes, all of the confounded information might disappear.

Alternatively, or perhaps in addition, it is possible that the messages that correspond to the confounded information are separable into contrast- and spatiotemporal pat-

tern-specific components if the concurrent responses of other neurons to the same stimulus are considered (deCharms, 1998). It is known, for example, that spikes that are synchronous across two cat LGN neurons can convey additional information beyond what could be obtained from each neuron's individual response (Dan et al., 1998), and similar results have been obtained in various cortical systems (Vaadia et al., 1995; Riehle et al., 1997; Maynard et al., 1999). However, it is important to point out that simply averaging together the responses of redundant neurons, or even neurons that have some degree of correlated variability but identical average responses (Shadlen and Newsome, 1998), would not help to disambiguate the confounded information. Thus, whether concurrent decoding of responses of a cluster of neurons can reduce the amount of confounded information is an issue that must be resolved experimentally. It is relatively straightforward to do this by an extension of the analysis methods used in this chapter to the responses of multiple neurons recorded together; this is the focus of Chapter 7.

What is expected from a simple-cell model? The degree to which a model of V1 simple cells can replicate the experimental results was examined. This model is quasi-linear and therefore fails to account for many of the interesting nonlinearities displayed by V1 neurons, particularly complex cells (Movshon et al., 1978a) but also, to some extent, simple cells (DeAngelis et al., 1993; Mechler et al., 1998a). In particular, the model does not account for contrast-related nonlinearities (Dean, 1981; Albrecht and Hamilton, 1982; Carandini et al., 1997b). The model significantly overestimates the magnitude of formal and attribute-specific information rates, in particular the contrast-specific information rates in stationary-grating responses (compare Fig. 6.8 to Fig. 6.6). Most significantly, the model responses contain no confounded information, in stark contrast to the prominent confounded information found in real responses.

In Chapter 5, a spike-train resampling technique is used to show that, for V1 neurons, the details of spike generation do not have a large effect on the magnitude of formal information rates. That result, together with additional resamplings done in connection with the results in this chapter (not shown), indicates that the discrepancies between real and model information rates (including confounded information) are not likely to be due to the assumption of a Poisson spike-generating mechanism in the model. Moreover,

these discrepancies are also not likely to be due to cell-to-cell variation in the shape of the linear filter or kernel, since such variation is similar for real and model responses and in both cases has little impact on information rates.

Instead, the discrepancies between real and model responses almost certainly relate to the fact that real responses to stimuli that differ only in contrast are not simply related by a scaling factor, but rather depend strongly on factors such as spatiotemporal pattern and the level of adaptation (Ohzawa et al., 1982; Bonds, 1991; Albrecht, 1995). It is possible that a single mechanism—nonlinear suppression that is sometimes called *contrast normalization* (Albrecht and Geisler, 1991; Heeger, 1992)—can account for all of these discrepancies, but only if the mechanism is sensitive to the spatiotemporal parameters of the stimulus and can affect the dynamics of the response. Indeed, such mechanisms are known to exist in the retina (Shapley and Victor, 1981), lateral geniculate nucleus (Sclar, 1987), and primary visual cortex (Reid et al., 1992). The sensitivity can be intrinsic to the suppressive mechanism itself or, alternatively, might be derived from a pooling of the responses of other V1 neurons with different stimulus-response properties.

Appendix

In this appendix, it is shown that there is no confounded information if and only if a neuron's response depends independently on the two stimulus attributes (in the experiments, contrast and spatiotemporal pattern). Stimuli are assumed to be defined by an independent choice of a stimulus s_1 out of a set S_1 (corresponding to the first attribute) and a stimulus s_2 out of a set S_2 (corresponding to the second attribute). A corollary of this demonstration is that there is no confounded information if and only if, for each possible response r, there is no mutual information between the conditional distributions $\{S_1|r\}$ and $\{S_2|r\}$. This is in turn equivalent to the statement that, for each response r, the conditional probability $p(s_1,s_2|r)$ is a separable function of s_1 and s_2: $p(s_1, s_2 \mid r) = p(s_1 \mid r)p(s_2 \mid r)$.

For notation, the conventions of (Cover and Thomas, 1991) are adopted. The mutual information between two variables X and Y is:

$$I(X;Y) = H(X) + H(Y) - H(X,Y), \tag{6.A1}$$

where $H(X)$ and $H(Y)$ are, respectively, the entropies of X and Y, and $H(X,Y)$ is the joint entropy of X and Y. For example, if $x \in X$ is distributed according to the probabilities $p(x)$, then:

$$H(X) = -\sum_{x \in X} p(x) \log p(x).$$

By the Chain Rule for entropies:

$$H(X,Y) = H(X) + H(Y \mid X) \tag{6.A2}$$

where $H(Y \mid X)$ denotes the conditional entropy of Y given X and is defined as:

$$H(Y \mid X) = -\sum_{x \in X} p(x) \sum_{y \in Y} p(y \mid x) \log p(y \mid x). \tag{6.A3}$$

Here, $p(y \mid x)$ is the conditional probability of y, given the occurrence of x.

The *confounded information* C is defined as:

$$C = I(R; S_1, S_2) - I(R; S_1) - I(R; S_2) \tag{6.A4}$$

where S_i is the i^{th} set of stimuli and R is the set of responses. Substitution of (6.A1) and (6.A2) into (6.A4), together with the fact that $H(S_1, S_2) = H(S_1) + H(S_2)$ (since S_1 and S_2 are independent), yields:

$$C = H(S_1 \mid R) + H(S_2 \mid R) - H(S_1, S_2 \mid R). \tag{6.A5}$$

Further substitution of (6.A3) into (6.A5) yields:

$$C = \sum_{r \in R} p(r) \left[\begin{array}{c} -\sum_{s_1 \in S_1} p(s_1 \mid r) \log p(s_1 \mid r) - \sum_{s_2 \in S_2} p(s_2 \mid r) \log p(s_2 \mid r) \\ + \sum_{s_1 \in S_1} \sum_{s_2 \in S_2} p(s_1, s_2 \mid r) \log p(s_1, s_2 \mid r) \end{array} \right]. \tag{6.A6}$$

Each $p(r)$ is a probability and therefore nonnegative. The term in brackets in (6.A6) must also be nonnegative, since it is the mutual information of S_1 and S_2, given the occurrence of r, and quantities of mutual information cannot be less than zero (Cover and Thomas, 1991). Thus, C=0 if and only if the term in brackets in (6.A6) is zero for every r that occurs with nonzero probability $p(r)$. But this is true if and only if, for every r, $I(S_1 \mid r; S_2 \mid r) = 0$. This latter condition in turn requires that:

$$p(s_1, s_2 \mid r) = p(s_1 \mid r)p(s_2 \mid r), \tag{6.A7}$$

meaning that $p(s_1, s_2 \mid r)$ is a separable function of s_1 and s_2.

It is straightforward to show that (6.A7) is equivalent to the independence condition, given that the stimulus probabilities are independent, i.e. $p(s_1, s_2) = p(s_1)p(s_2)$. By Bayes's rule:

$$p(s_1, s_2 \mid r) = \frac{p(r \mid s_1, s_2)p(s_1)p(s_2)}{p(r)}. \tag{6.A8}$$

Combining (6.A7) and (6.A8), and the fact that $p(s_i \mid r)p(r) = p(r \mid s_i)p(s_i)$, the independence condition:

$$p(r \mid s_1, s_2) = \frac{p(r \mid s_1)p(r \mid s_2)}{p(r)}, \tag{6.A9}$$

is obtained.

Examples. In the following, eight simple examples are presented to illustrate the concept of confounded information. The examples consist of systems with two independent inputs, each of which can take a value of 0 or 1 with equal probability, and a single output. There are therefore four different input configurations of the stimuli (s_1,s_2): (0,0), (0,1), (1,0), and (1,1). In the first six examples, the output is also binary, whereas in the last two examples, the output can take on more than two distinct values. The example systems are displayed in Table 6.4, together with the corresponding information values.

In the first example, the response is independent of s_1 and is completely determined by s_2. In the second example, the response is independent of s_2 and is completely determined by s_1. Of the 2 bits of entropy in the stimuli, only 1 bit is conveyed as information, but it is transmitted perfectly. Moreover, the response to one stimulus is independent of the value of the other, so that there is no confounded information.

In the third example, the system responds with a 1 if the two inputs are identical, and with a 0 if they are different. There is still 1 bit of information transmitted, but in this case the response to one stimulus depends completely on the value of the second, so that

STIMULI		RESPONSES (conditional probability $p(r \mid s_1, s_2)$)							
s_1	s_2	1	2	3	4	5	6	7	8
0	0	0 (1)	0 (1)	1 (1)	0 (1)	0 (0.5) 1 (0.5)	0 (1)	0 (1)	0 (1)
0	1	0 (1)	1 (1)	0 (1)	0 (1)	0 (1)	0 (0.5) 1 (0.5)	1 (1)	1 (1)
1	0	1 (1)	0 (1)	0 (1)	0 (1)	0 (0.5) 1 (0.5)	0 (0.5) 1 (0.5)	1 (1)	2 (1)
1	1	1 (1)	1 (1)	1 (1)	1 (1)	1 (1)	1 (1)	2 (1)	3 (1)
INFORMATION (bits)									
$I(S_1;R)$		1	0	0	0.31	0.19	0.19	0.5	1
$I(S_2;R)$		0	1	0	0.31	0	0.19	0.5	1
$I(S_1,S_2;R)$		1	1	1	0.81	0.5	0.5	1.5	2
CNF		0	0	1	0.19	0.31	0.12	0.5	0

Table 6.4. *Eight simple systems that feature two binary inputs (stimuli) and different output (response) distributions.* The conditional probability of each possible output is given in parentheses next to the output value itself. Information is calculated by the methods detailed in the Appendix. $I(X;Y)$ refers to the information between two variables X and Y, and *CNF* refers to the confounded information.

all the information is confounded—that is, there is no information transmitted about either stimulus unless the value of the other stimulus is known.

In the fourth example, the system responds only if the values of both stimuli are 1. Here, the response is symmetric in the two stimuli, so that the response to each stimulus alone conveys the same amount of information. However, because the response depends jointly on the values of both stimuli, some of the information—23% of the total—is confounded.

In the next two examples, the system's response again depends in a complicated manner on the two stimuli, so that the confounded information is non-zero. In the fifth example, S_2 acts like a gate: the system responds at random if $s_2 = 0$, and identically reflects s_1 if $s_2 = 1$. Confounded information arises because although no information is conveyed about s_2, the response to s_1 is more informative if s_2 is known. In the sixth example, the system responds at random if the two stimuli are different and reflects their shared value if they are the same. The response thus conveys equal amounts of information about the two stimuli, but there is still some confounded information.

The first six examples illustrate that confounded information can arise even in very simple systems, so long as the conditional response probabilities are separable in the

two stimuli, as in (6.A9). The last two examples demonstrate that this requirement is not equivalent to a requirement that stimulus encoding be linear or additive. The seventh example is a system that simply sums the values of the two stimuli (and thus has three possible responses); the confounded information in this system is 33% of the total, despite the fact that the system is linear. On the other hand, the eighth example is a system that generates distinct responses to each of the four stimuli, but its response is also additive in the sense that the response to any pair of stimuli is the sum of the responses to two pairs of stimuli that add up to the same input. For instance, the response to (1,1) is equal to the sum of the responses to (1,0) and (0,1). As with any system that maps each input to a distinct output, even if it is not additive, the system in the eighth example does not produce confounded information.

Thus, it is not surprising that substantial amounts of confounded information are observed in real neuronal responses: real responses depend in a complicated way on both contrast and spatiotemporal pattern. In the simple examples considered here, in fact, only systems that ignore one or the other stimulus, as in the first two examples, or systems that respond differently to each stimulus pair, as in the eighth example, can feature responses that convey information about both stimuli without confounding them.

Chapter 7

Independent and Redundant Information in Nearby Cortical Neurons

Summary

Tetrode recording techniques, which allow the simultaneous responses of several nearby neurons to be isolated (see Methods), and the direct method of calculating information rates, which can be extended to analyze the responses of such multi-neuronal data, comprise a powerful pair of tools. These tools can be used to address the central question in the thesis: How is stimulus-related information represented in the activity of nearby V1 neurons? In this chapter, a less general version of the question is asked: What is the benefit of keeping track of each neuron's response, compared with the simpler alternative of averaging responses across neurons? The major conclusion is that the information conveyed by groups of nearby neurons is nearly independent, more so when responses of different neurons are kept separate than when they are averaged together. The amount of information loss due to averaging is greater in larger populations. However, for some stimulus attributes, such as contrast, little or no information is lost when responses are averaged together.

Introduction

How do neurons in sensory cortex work together to represent a stimulus? (Mountcastle, 1957) showed that the cortex is organized into functional "columns," so that neurons with similar stimulus selectivities are in close proximity to one another. In V1, responses of nearby neurons convey information about similar stimulus attributes

(Hubel and Wiesel, 1962; DeAngelis et al., 1999). This localization of shared stimulus selectivity makes it possible for experimenters to record population activity that bears a meaningful relationship to visual stimuli.

It has been argued that the large trial-to-trial variability of cortical neurons' responses means that reliable information can only be transmitted by the average activity of groups of similarly tuned neurons (Shadlen and Newsome, 1998), since averaging tends to increase the signal-to-noise ratio. If the background activities of cortical neurons are correlated (Arieli et al., 1996; Lampl et al., 1999; Schiff et al., 1999), however, the ability of averaging to increase the signal-to-noise ratio is limited. On the basis of these considerations, it has been estimated that the number of neurons required to reliably encode a particular piece of stimulus-related information can be substantial—on the order of 100 (Zohary et al., 1994; Shadlen and Newsome, 1998). At the same time, theoretical analysis (Abbott and Dayan, 1998) has revealed that neurons tuned to similar stimuli can convey independent or nearly independent information in their responses, a phenomenon that has been directly demonstrated for pairs of complex cells in visual cortex (Gawne et al., 1996a). These results suggest that although a decoding mechanism that operates by averaging would reduce noise, it would also sacrifice information independently encoded by different neurons.

In this chapter, the direct method is used to calculate information rates under two assumptions: first, that responses are averaged together across neurons, and second, that the neuron of origin of each spike must be considered in the decoding. The data are for the most part those analyzed in previous chapters, but here the fact that many of the neurons were recorded simultaneously is used for the first time. As in Chapter 6, responses to different types of stimulus are considered, and information rates are parsed into formal, attribute-specific, and confounded components. The analysis shows that the information conveyed by nearby neurons is nearly independent, and that a substantial portion of that information is lost if the responses of those neurons are averaged together.

Methods

The data-collection methods are identical to those described in Chapter 6 and are not recapitulated here in detail. The data described in this chapter come from tetrode recordings of two to eight well-isolated, individual neurons. Details of tetrode recording and spike sorting are presented in Chapter 2. As in Chapter 6, stimuli include m-sequence checkerboards as well as drifting and stationary sinusoidal gratings presented at five contrast values: 0.0625, 0.125, 0.25, 0.5, and 1. For each stimulus type, a neuron's response is analyzed if its firing rate is greater than 1 spike/sec and is significantly affected by the stimulus (i.e., mean stimulated firing rate or rate modulation is significantly different from the uniform-field condition).

Responses to two additional sets of stimuli, both involving stationary sinusoidal gratings, are discussed in this chapter. The first set consists of unit-contrast gratings at the optimal orientation and spatial frequency. These gratings are presented at 16 different spatial phases, spaced uniformly around the circle. Each experimental trial consists of 237 msec of the grating followed by 710 msec of uniform-field presentation, repeated eight times. The 16 phases are presented in a different random order during each repeat of the entire stimulus set. For the analysis described here, six to eight repeats are collected, giving 48 to 64 presentations of the grating at each spatial phase.

The second set of new stimuli consists of unit-contrast gratings that vary in spatial phase, orientation, and spatial frequency.[9] The gratings appear for 237 msec and are replaced by a uniform field at the same mean luminance for an additional 1026 msec before the next grating is presented. Gratings are presented at each of four different spatial phases, spaced uniformly around the circle, and at each of two or three orientations and spatial frequencies, chosen to be close the peak of the neurons' tuning curves (determined with drifting gratings). If different neurons in the cluster are found to have distinct orientation or spatial-frequency preferences, the individual preferences of each neuron are in-

[9] Responses of single neurons to stimuli such as these have been previously described and analyzed with the metric-space method, the same method used to calculate transmitted information in Chapter 4 (Victor and Purpura, 1998). The data used in (Victor and Purpura, 1998) were collected with single electrodes and do not overlap with the data used here.

cluded in the stimulus set. Stimuli are presented in pseudo-random order, so that no particular stimulus occupies a privileged position in the order of presentation. Gratings are presented between 64 and 128 times across all neurons, although each grating is presented an equal number of times at a given recording site.

Information rates. From the responses of neurons to multiple repeats of each stimulus, information transmission rates are calculated by a modification of the direct method (Ruyter van Steveninck et al., 1997). As explained in Chapter 6, this analysis is predicated on the notion that changes in a neuron's firing probability on short timescales, quantified by a measure of entropy, can be induced by the presence of a stimulus and by intrinsic neuronal noise (which may be correlated across neurons). The portion of the response modulation due to noise can be estimated by assessing the variability across responses to repetitions of an identical stimulus. The difference between the overall (total) and noise entropies is the information rate, which is normalized to units of bits/sec.

A point made in Chapter 6 regarding the magnitude of the calculated information rates needs to be reiterated here. Information rates obtained through the direct method using single time bins are instantaneous measures—that is, estimates of the average information in brief samples of the response, randomly selected from the time course of the experiment. Because information rates are measured in single, short time bins, longer-term correlations in the response, which may be due both to correlations in the stimulus and the intrinsic dynamics of the neuron, are effectively ignored. Since these correlations may affect the information rate, the total transmitted information over a long period of time cannot be taken as the product of the information rate (measured in short bins) and the time period of interest. That is, information rates are not additive across time (DeWeese and Meister, 1999).

To use the direct method to calculate information rates from the responses of multiple neurons, two strategies, which correspond to different ways of decoding the neural signal, are compared. In the first strategy, a *summed-population code*, spike counts from the various neurons are simply added together in each time bin, and the direct method is applied exactly as described in Chapter 6. The addition of spike counts across neurons

formalizes the hypothesis that stimuli are represented by activity in a local cluster of neurons, and that the times of the spikes but not their neuron of origin are important. In the second strategy, a *labeled-line code*, both the neuron of origin and the spike times are important. Each time bin holds a vector of spike counts, each dimension of which contains the spike count in one of the neurons. The labeled-line code formalizes the hypothesis that postsynaptic neurons process input in a manner that depends on the neuron of origin. For the purposes of using the direct method to calculate information rates for the labeled-line code, entropies are based on the empirical probability distributions of the across-neuron spike count vectors.

Clearly, the information rates for the labeled-line code can be no lower than for the summed-population code, because all of the information in the summed-population code can be recovered by summing across dimensions in the labeled-line code. If knowledge of the neuron that fires a spike is not important for conveying stimulus-related information, then the information rates for the two codes are expected to be identical to within the resolution of the measurement. In the results that follow, information rates for the two codes are compared to each other and to the results that would obtain if information rates calculated separately for each neuron, as in Chapter 6, were simply added together. This final quantity (the *separate sum*) is the overall information rate that would be conveyed by a set of completely independent neurons.

Bias in the information-rate estimates. The methods developed in Chapter 6 for correcting the inherent bias in information estimates from limited data sets are extended to the multi-neuron analysis and used in this chapter. These methods are based on the assumption that the entropy changes little during periods of low population firing rate. Thus, in the calculation of noise entropy, consecutive time bins are grouped together so that each set of time bins contains at least one spike fired by at least one of the neurons. However, the use of multi-dimensional vectors in the labeled-line code raises several additional complications, because the number of degrees of freedom, and hence the bias in the information calculations, is increased relative to the single-neuron case. Simulations reveal that the effects of this bias on the total entropy are negligible, but that they can be substantial with respect to the noise entropy, since the number of samples is smaller. For

this reason, the noise entropy (but not the total entropy) is assumed to be additive across neurons, which effectively means that the noise is assumed to be uncorrelated across neurons in short time windows. Grouping of time bins so that each bin contains at least one spike proceeds separately for each neuron in the cluster. This bias-correction procedure increases (or, at the very least, leaves unchanged) the estimated noise entropy and therefore reduces the estimated information rate (Cover and Thomas, 1991). Overall, the procedure is not expected to induce a substantial systematic error in the information calculations (Panzeri et al., 1999).

There is an additional source of bias in the multi-neuron calculations that is difficult to correct. As discussed in Chapter 6, the information rate can only increase as the time bin is made shorter, and the duration of the time bins is chosen separately for each neuron or group of neurons in order to maximize the estimated information rate. In practice, estimated information rates decrease if the time bins are made too short because the intrinsic bias is not sufficiently corrected on those timescales. Because summing (or averaging) spike counts across neurons increases the effective amount of data (number of spikes in each bin), and may therefore lead to a more accurate estimate of the underlying spike-count distribution and entropy, it is possible that the information rates are more severely underestimated for the labeled-line code than for the summed-population code. This can lead to situations in which the calculated information rate is higher for the summed-population code than for the labeled-line code. If the information rate were calculated completely accurately, as explained above, this would never happen. Therefore, when this situation is encountered, the two codes are considered to be equivalent and the estimated information rate is taken to be the one calculated for the summed-population code. It should be noted that because this source of bias tends to increase the estimate of information rate for the summed-population code relative to the labeled-line code, it is likely to diminish the apparent distinction between information rates for the two codes.

Results

Formal information rates. ***M-sequence responses.*** Fig. 7.1 shows two sets of results from cell pairs in V1. The stimulus is a flickering m-sequence checkerboard at unit contrast. The extracellularly recorded spike waveforms from the two cells in each pair are cleanly isolated and detected on overlapping leads of a tetrode (see Chapter 2), so that each pair consists of nearby, possibly neighboring, neurons. Analysis of the first pair is shown in Fig. 7.1A-D and reveals an excellent example of two nearby neurons that convey largely independent stimulus-related information (see below). This particular pair consists of one simple cell (Fig. 7.1A-B, top) and one complex cell (bottom).[10] Both neurons respond robustly to the stimulus, as can be seen from the sharp and reliable fluctuations in firing rate during a representative one-second segment of the response (Fig. 7.1A). The response histograms indicate that although the two neurons occasionally respond to similar features of the stimulus (double-headed arrow), their responses are different from one another most of the time. For example, the large peak pointed to by the single-headed arrow in the simple-cell response does not appear in the complex-cell response. Receptive-field maps (Fig. 7.1B) obtained by cross-correlation of the stimulus and the spike train response (see Chapter 5) indicate that the two neurons are sensitive to stimuli presented in the same region of space; in particular, the off subregions of the two receptive fields (dark) are spatially superimposed.

The second example in Fig. 7.1 (E-H) comes from the responses of a pair of complex cells that convey information at rates more typical of the population. These two neurons respond somewhat less robustly to the m-sequence stimulus. Nevertheless, clear differences between the responses can be seen in the raster diagrams and response histograms (Fig. 7.1E) and in the receptive-field maps (Fig. 7.1F). As with the simple-cell/complex cell pair, the receptive fields of these complex cells are spatially superim-

[10] Despite the fact that the receptive-field map of the bottom neuron in Fig. 7.1B resembles simple-cell receptive-field maps (Reid et al., 1997), the neuron's response to drifting gratings at the optimal spatial frequency, temporal frequency, and orientation resembles that of a complex cell. Since the simple-cell/complex-cell classification is based on the responses to drifting gratings (Skottun et al., 1991), the neuron is classified as a complex cell.

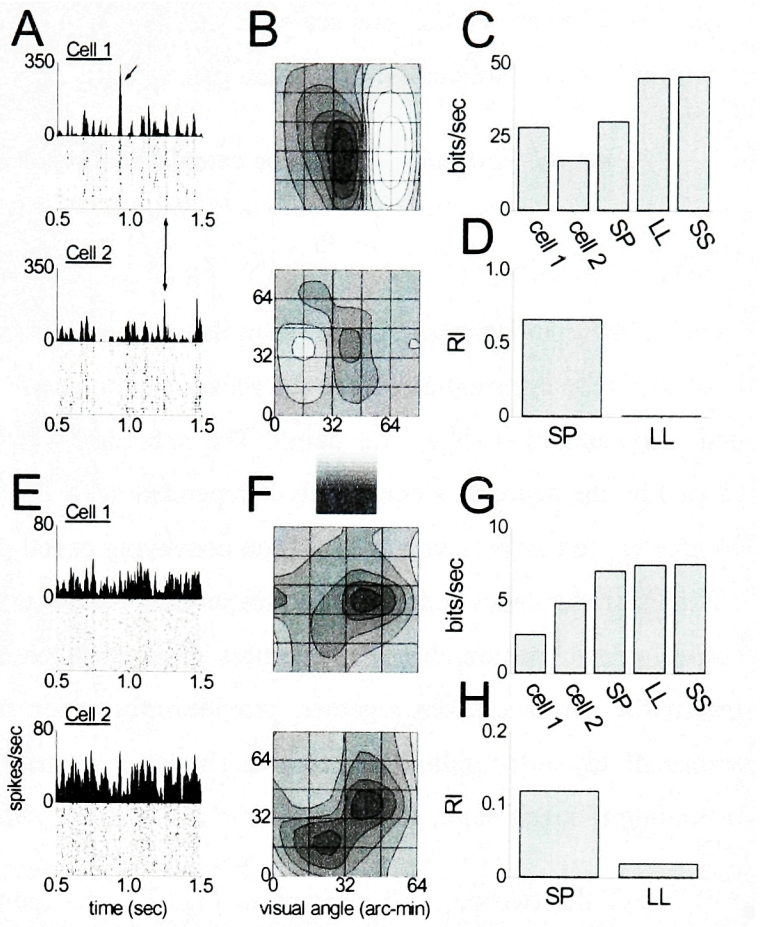

Figure 7.1. *Responses of two pairs of neurons to the unit-contrast m-sequence.* **A-D**: simple cell (top) and complex cell (bottom). **E-H**: two complex cells. **A,E**: response histograms (3.7 msec bins) and raster diagrams that depict the spike times fired during a one-second segment of each trial. **B,F**: contour maps that show snapshots of each neuron's receptive field averaged over a 14.8 msec post-stimulus time window (Chapter 5): 44 to 59 msec in **B**, 74 to 89 msec in **F**. Gray scale: ±5 spikes/sec (stimulus-induced change from background firing rate) in **B**, ±0.4 spikes/sec in **F**. Dark regions correspond to areas in which dark stimuli cause the neuron to fire or bright stimuli inhibit the neuron from firing, and light regions correspond to the opposite situation. **C,G**: Formal information rates, in bits/sec: each cell individually (*cell 1* and *cell 2*), summed-population code (*SP*), labeled-line code (*LL*), and separate sum (*SS*). **D,H**: redundancy indices (RI) for the summed-population and labeled-line codes. See text for explanations of the different codes and the definition of the redundancy index.

posed, but in this case both neurons have a single dominant receptive-field subregion, of the same polarity.

Fig. 7.1 also plots the rates at which information about this stimulus is conveyed by the two pairs of neurons. For the simple-cell/complex-cell pair (Fig. 7.1C), the rate of information transmission for the summed-population code (SP), in which responses from the two neurons are averaged together in each time bin, is only slightly higher than the rate for the first cell alone. On the other hand, the information rate for the labeled-line code (LL), in which the neuron of origin of each spike is considered, is 49% higher than the summed-population-code information rate. It is also nearly identical to the sum of the information rates calculated separately for the two neurons (SS, or *separate sum*). For the pair of complex cells (Fig. 7.1G), the difference between the labeled-line and summed-population information rates is only 5.6%

(note the different vertical-axis scales in panels C and G, which indicate that the overall rate of information transmission is much greater for the simple-cell/complex-cell pair).

A *redundancy index* (RI) can be calculated for each group of neurons as:

$$RI = \left(1 - \frac{I_C}{I_{SS}}\right) \Big/ \left(1 - \frac{1}{N}\right), \tag{7.1}$$

where I_C is the information rate in the multi-neuron code (summed-population or labeled-line); I_{SS} is the information rate in the separate sum; and N is the total number of neurons under consideration ($N = 2$ for pairs). The redundancy index is 0 when the information carried by the neurons is completely independent ($I_C = I_{SS}$) and 1 when the information is completely redundant, with each neuron conveying exactly the same information ($I_C/I_{SS} = 1/N$). The redundancy index can be greater than 1 if the neurons provide contradictory or confusing information about the stimulus when considered together, and it can be negative if the neurons, taken together, provide information that is not available in the response of any of the individual neurons. The latter situation is sometimes called *synergistic coding* (Panzeri et al., 1999; Brenner et al., 2000).

For the two pairs of neurons in Fig. 7.1, the redundancy index is substantially greater for the summed-population code than for the labeled-line code. For the simple-cell/complex-cell pair (Fig. 7.1D), the summed-population redundancy index is more than half of what it would be if the responses of the two neurons were completely redundant. On the other hand, the labeled-line redundancy index is very close to 0, indicating that the responses are nearly independent if the neuron of origin of each spike is taken into account. Thus, in spite of visible correlations such as the double-headed arrow in Fig. 7.1A, a spike fired by one neuron is of little help in predicting the occurrence of a spike fired by the other. For the complex-cell pair (Fig. 7.1H), responses are substantially less redundant for the summed-population code, but they are once again nearly independent for the labeled-line code.

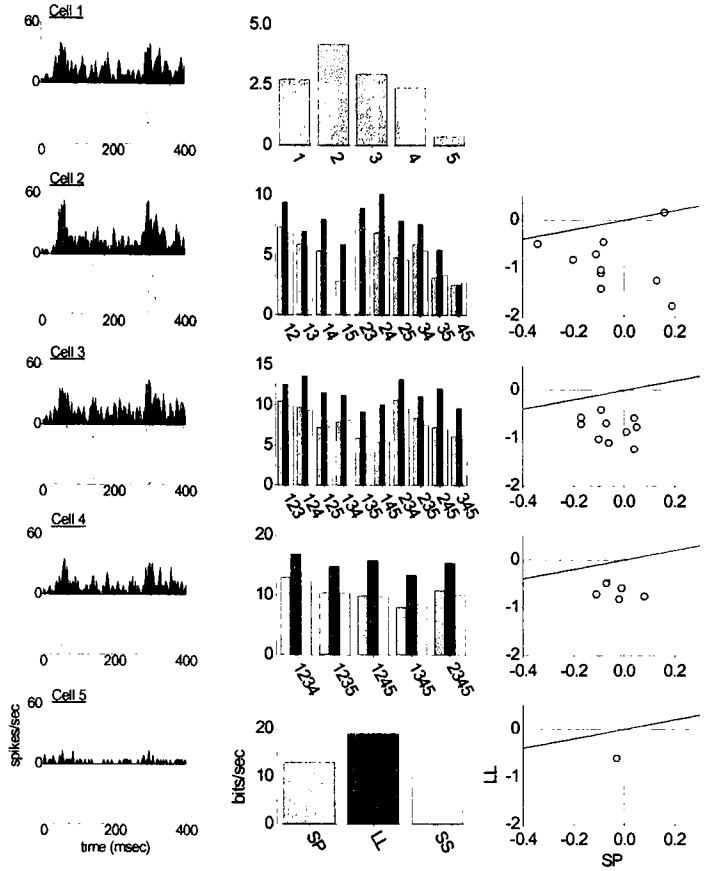

Figure 7.2. *Responses of five simultaneously recorded complex cells to unit-contrast stationary sinusoidal gratings. Cell 1* and *cell 2* are the same neurons as those presented in Fig. 7.1E-H. Spike times between 30 and 300 msec post-stimulus are used in the information-rate calculations. **Left column**: response histograms (3.7 msec bins) and raster diagrams. **Middle column**: formal information rates (bits/sec) for each combination of neurons. **Top row**: individual neurons. **Second row**: pairs of neurons: **Third row**: triplets of neurons. **Fourth row**: quadruplets of neurons. **Fifth row**: quintuplets of neurons. For all combinations with more than one neuron, *gray* gives the information rate for the summed-population code (SP), *black* gives the information rate for the labeled-line code (LL), and *white* gives the information rate for the separate sum (SS), in which the information rates for each individual neuron are added together. **Right column**: redundancy indices for the summed-population code are plotted (on the horizontal axis) against redundancy indices for the labeled-line code. Note the different horizontal- and vertical-axis scales. The diagonal line shows the locus of points at which the redundancy indices for the two codes are identical.

Stationary-grating responses. The responses of the complex-cell pair to stationary-grating stimuli at unit contrast are shown in Fig. 7.2 (cells 1 and 2). The responses of three additional complex cells recorded simultaneously on the same tetrode are also shown. One of these neurons (cell 5) barely responds to the stimulus and therefore conveys very little stimulus-related information by itself. This set of stationary-grating responses shows synergistic coding, a phenomenon that is relatively rare in the data set (see Fig. 7.4). The signature of synergistic coding is higher information rates for either the summed-population or labeled-line codes (or both, as is the case with this set of responses) than for the separate sum, which results in negative values of the redundancy index. In the middle column of Fig. 7.2, information rates are plotted for each combination of the five neurons: separately analyzed single neu-

rons in the top panel, followed by pairs, triplets, quadruplets, and quintuplets of neurons. In each case (except for the singlets, for which there is trivially no across-neuron coding), information rates for the summed-population code (gray), labeled-line code (black), and separate sum (white) are juxtaposed. For most combinations, both the summed-population and labeled-line codes yield higher information rates than the separate sum, although the magnitude of the effect, with a few exceptions (such as pair *45*), is substantially greater for the labeled-line code.

The information rates are converted into redundancy indices in the right column. For each combination of neurons, the corresponding value for the summed-population code (horizontal axis) is plotted against the redundancy index for the labeled-line code (vertical axis). Note that the vertical axis has a much larger range than the horizontal axis. The line of equality is shown in each redundancy-index plot for reference; the fact that labeled-line information rates can be no less than summed-population information rates means that points cannot fall between the line of equality and the horizontal zero-line. The plots reveal that redundancy indices are typically negative for both codes, reflecting the synergistic coding, but significantly more negative for the labeled-line code. With increasing numbers of neurons (lower panels in the right column), redundancy indices tend to cluster more tightly around the vertical zero-line for the summed-population code, meaning that responses analyzed with this code tend to be more independent (and less synergistic). This is not the case, however, for the labeled-line code—strongly synergistic information transmission persists even when all five neurons are analyzed together.

Drifting-grating responses. In Fig. 7.3, the drifting-grating responses of the two pairs of neurons from Fig. 7.1 are analyzed. For the first pair (Fig. 7.3A-C), the parameters of the drifting grating are optimal for the simple cell (top), although the complex cell also responds robustly. The rate of information transfer for the simple cell is high, even relative to the population of recorded simple cells (see Fig. 6.3), and despite the obvious non-stationarity in the response phase across trials. For the simple-cell/complex-cell pair, the summed-population code actually conveys information at a lower rate than does the simple cell alone—that is, averaging in the response of the complex cell leads to some confusion. This may occur if, for example, signal and/or noise in the two responses are

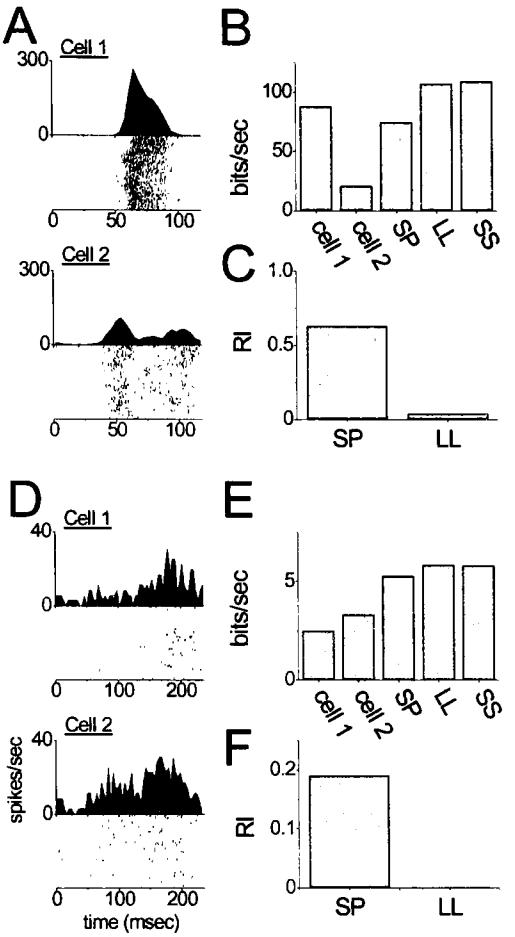

Figure 7.3. *Responses of the two pairs of neurons to drifting sinusoidal gratings.* Same pairs as in Fig. 7.1 Parameters of the gratings are optimized for at least one of the neurons. **A,D**: response histograms (3.7 msec bins) and raster diagrams. **B,E**: formal information rates (bits/sec) for individual neurons and neuron pairs (summed-population, code [SP], labeled-line code [LL], and separate sum [SS]). **C,F**: redundancy indices for the summed-population and labeled-line codes.

uncorrelated or anticorrelated. Keeping track of the neuron of origin of each spike, as in the labeled-line code, yields an information rate that is nearly equal to the sum of the two neurons' separately calculated information rates, resulting in a redundancy index close to zero. Thus, the drifting-grating responses of these two neurons are similar in character to their m-sequence responses (Fig. 7.1A-D), although the information rates themselves are more than twice as high in the drifting-grating responses.

The pair of complex cells (Fig. 7.3D-F) transmits less information in response to drifting gratings than to m-sequences. As discussed in Chapter 6, the direct method, when applied to the response of a complex cell to a drifting grating at a single contrast, yields a low information rate because the response is only poorly modulated in time. This particular pair of neurons has responses that are more modulated than is typically the case for complex cells, so that the information rates estimated by the direct method are unusually high (see Fig. 6.3). For this pair, the changes in information rate and redundancy index for the different codes are similar to the changes observed in the m-sequence responses (Fig. 7.1E-H)—in particular, there is substantial redundancy for the summed-population code but near-independence for the labeled-line code.

Summary of information rates at fixed contrast. Fig. 7.4 is an across-recording-site summary of the information rates conveyed by the summed-population and labeled-line codes in response to the three types of stimulus considered above. The top row shows the number of distinct sites at which neuron combinations of different sizes were recorded. (These are cumulative distributions in the sense that a site at which N neurons are recorded is also considered to be a site at which groups of up to N-1 neurons are recorded.) Since a site with N neurons contributes $\binom{N}{k}$ distinct k-tuples of neurons, so that some sites would be highly over-represented if all neuron combinations were treated equally, the data in Fig. 7.4 are shown as grand medians across recording sites, where each recording site contributes the median across all neuron combinations of a particular size at that site. Across-neuron medians, taken without regard to recording site, are similar in character but are unreliable when the sample size is small.

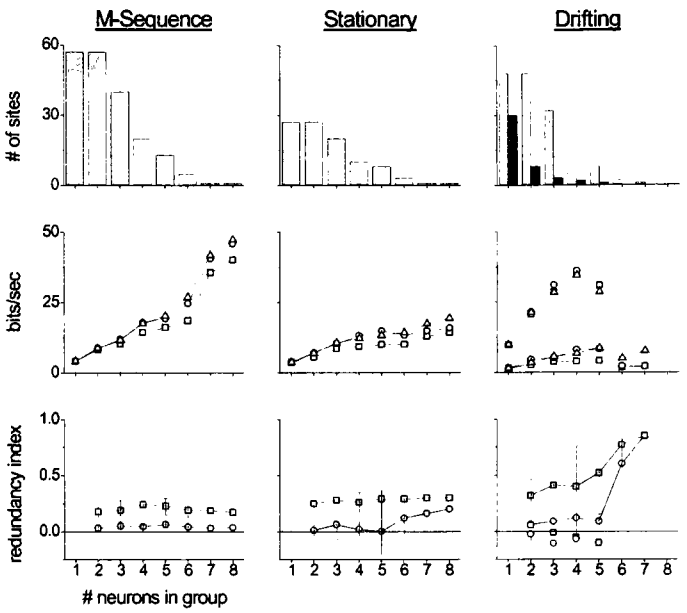

Figure 7.4. *Summary of formal information rates and redundancy indices in response to unit-contrast stimuli.* Medians are first calculated within each recording site; the data in the figure are grand medians across recording sites. **Left column**: m-sequence responses. **Middle column**: stationary-grating responses. **Right column**: drifting-grating responses. Results from simple cells alone are presented as black columns and dotted lines. **Top row**: distribution of number of sites at which at least N (horizontal-axis parameter) neurons are recorded. **Middle row**: median information rates in bits/sec for the summed-population code (□), labeled-line code (○), and separate sum across neurons (Δ). **Bottom row**: median redundancy indices for the summed-population (□) and labeled-line (○) codes. When more than one recording site is sampled, error bars represent 95% confidence intervals of the medians determined from 1000 bootstrap resamplings of the distributions.

The second row of Fig. 7.4 shows median information rates as a function of the number of neurons in each combination. The results indicate that small differences between the summed-population and labeled-line codes for pairwise combinations of neurons can have important implications for coding when modest populations (groups of up to eight neurons) are considered. For the labeled-line code, median m-sequence information rates increase nearly as quickly as the separate sum and more quickly than for the summed-population code. Because raw information rates can vary widely across recording sites, and because only limited numbers of recording sites are available for larger combinations of neurons, the redundancy index, which is normalized separately for each neuron combination (Eq. 7.1), is a more sensitive measure of the differences in information rates between the two codes. For the data in Fig. 7.4, the difference between the redundancy-index curves (bottom row) does not change with the number of neurons being analyzed. This means that the amount of information lost by averaging across populations increases when more neurons are added to the pool, even though the increase may not be reflected in the median information rates because of small sample sizes. For the m-sequence responses, median redundancy indices (bottom row) are about four times higher for the summed-population code than for the labeled-line code. The differences in redundancy indices for the two codes are highly significant: error bars in these and subsequent plots represent 95% confidence intervals of the medians, calculated from 1000 bootstrap resamplings of the distributions. Error bars are not shown in the information-rate plots to maintain clarity of presentation.

For the stationary-grating responses, which convey less information overall, the difference between the labeled-line and summed-population codes is greatest for combinations of five neurons. For larger combinations, labeled-line redundancy indices tend to increase relative to summed-population redundancy indices, although this may in part be a result of undersampling of the large combinations. The strongly synergistic coding signified by the large negative labeled-line redundancy indices for the five neurons of Fig. 7.2 is not typical of the population. However, synergistic encoding is in general far more likely to occur when the neuron of origin of each spike is considered than when responses are averaged across neurons.

The drifting-grating responses transmit information at the lowest rates, in large measure because unmodulated complex-cell responses dominate the sample by approximately three-to-one. Sites at which larger numbers of neurons were recorded tended to be especially dominated by complex cells, although the sample size is not large enough to determine whether this is a general feature of V1 recordings. Information rates from combinations of simple cells exclusively (dotted lines and black columns in the right column) are higher than median m-sequence information rates, consistent with the results of Chapter 6. In the small sample of simple cells, but not in the larger sample, synergistic coding is observed, as revealed by negative redundancy indices especially for the labeled-line code. Median redundancy indices for the two codes, across all recording sites, remain roughly equidistant up to combinations of five neurons and then converge, although the convergence may again be an artifact of undersampling. Thus, for all three types of stimulus, increasingly large amounts of information are lost if responses are averaged together rather than kept separate for each neuron.

Attribute-specific information rates. In Chapter 6, a distinction was drawn between information about particular stimulus attributes, such as contrast and spatiotemporal pattern, and the overall information conveyed by the response (dubbed *formal information*). For the responses depicted in Figs. 7.1-7.4, in which stimuli are all presented at a single (unit) contrast, only formal information can be calculated. However, it is of interest to extend the attribute-specific information to multi-unit recordings, to determine whether the differences between the labeled-line and summed-population codes persist, and if so, how these differences vary with the kind of information considered. In addition, for the responses of single neurons, information confounded between contrast and spatiotemporal pattern represents 10% to 32% (interquartile range) of the formal information (Chapter 6). Does the amount of confounded information increase or decrease when the responses of multiple neurons are analyzed together?

For this analysis, responses to m-sequences, drifting gratings, and stationary gratings are collected at five different contrasts (0.0625, 0.125, 0.25, 0.5, and 1). In addition, responses to unit-contrast stationary gratings at 16 different spatial phases (spaced equally around the circle) are analyzed, as are responses to unit-contrast stationary grat-

ings that vary along three different dimensions (spatial phase, spatial frequency, and orientation). Population summaries of responses to each set of stimuli are presented in turn.

M-sequence responses. Fig. 7.5 presents data from m-sequence stimulation of V1 neurons at five contrasts. As in Fig. 7.4, the median is first taken separately at each recording site, and then the grand medians across sites are plotted against the number of neurons in the combination. When responses to all five contrasts are analyzed, formal information rates (left column, top row) are lower than they are when only the unit-contrast response is analyzed (Fig. 7.4), regardless of the particular code (summed-population, labeled-line, or separate-sum). Again, the summed-population code leads to a lower formal information rate that generally increases with the number of neurons in the combination. The labeled-line code, on the other hand, remains very close to the separate sum for groups of up to four neurons. The drop-off with five and six neurons may be a

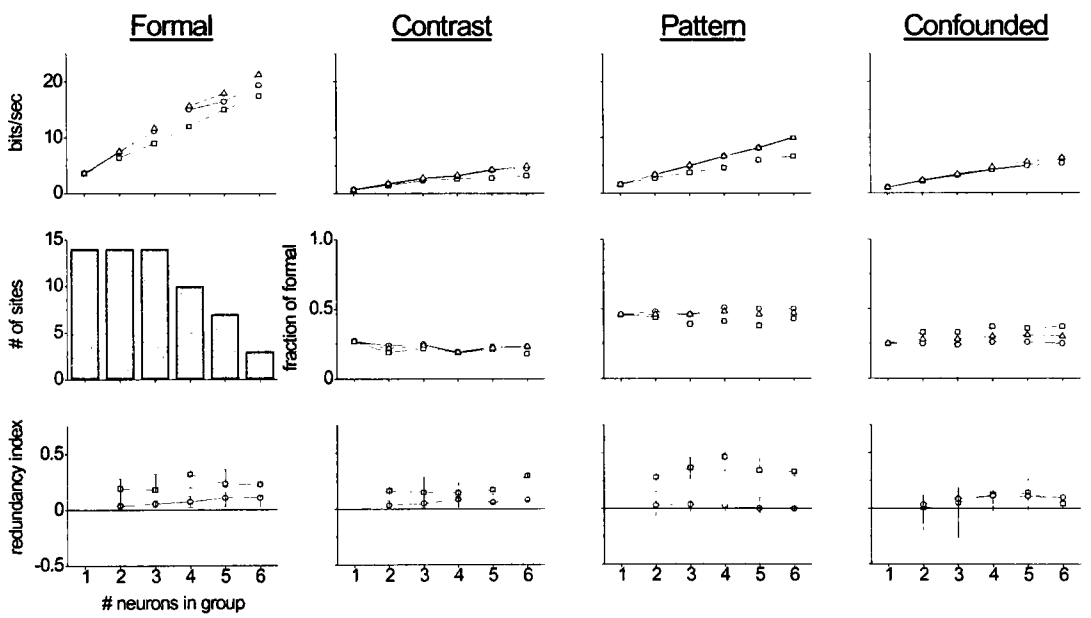

Figure 7.5. *Summary of formal, attribute-specific, and confounded information rates for responses to m-sequences at five contrasts.* Contrasts used: 0.0625, 0.125, 0.25, 0.5, 1. Medians are first calculated within each recording site; the data in the figure are grand medians across recording sites. **First column**: formal information rates. **Second column**: contrast-specific information rates. **Third column**: spatiotemporal pattern-specific information rates. **Fourth column**: confounded information rates. **Top row**: median information rates, bits/sec, for the summed-population code (□), labeled-line code (○), and separate sum across neurons (Δ). **Middle row, first panel**: distribution of number of sites at which at least *N* (horizontal-axis parameter) neurons are recorded. **Middle row, second, third, and fourth panels**: fraction of formal information rate accounted for by information about specific attributes. **Bottom row**: median redundancy indices for the summed-population (□) and labeled-line (○) codes. Error bars as in Fig. 7.4.

result of the fact that large neuron-groups are undersampled; in the larger database of unit-contrast m-sequence responses (Fig. 7.4), the difference between labeled-line and summed-population information rates continues to increase up to groups of six neurons. Redundancy indices (Fig. 7.5, bottom row), which are normalized indicators of population behavior, converge slightly for the two codes as the number of neurons gets larger but remain significantly higher for the summed-population code. This indicates that stimulus-related information is conveyed in the identity of the neuron that fires each spike.

Attribute-specific information rates behave similarly, although the information rates themselves are, as expected, significantly lower. As described in Chapter 6, contrast-specific information reflects the way the overall spike-count distribution changes as a function of contrast, whereas spatiotemporal pattern-specific information reflects the way the spike count distribution at fixed times changes regardless of contrast. The relative magnitude of the two types of attribute-specific information is determined by the relative sizes of these distribution changes—that is, whether the firing statistics change more with contrast or with time. For m-sequence responses, the difference in spatiotemporal pattern-specific information rates between the summed-population and labeled-line codes increases with the number of neurons throughout the entire range. By the same token, the difference between the redundancy indices for the two codes remains constant, and the indices themselves are significantly different. However, the two codes give similar contrast-specific information rates and redundancy indices (overlapping error bars). Thus, averaging spike counts across neurons sacrifices little contrast-specific information but an increasingly large portion of spatiotemporal pattern-specific information as more and more neurons are added to the pool. This suggests that nearby neurons may have similar contrast sensitivities but reliable differences in terms of sensitivity to spatiotemporal pattern. Overall, the fraction of the formal information rate devoted to encoding each stimulus attribute is relatively insensitive to the number of neurons being analyzed.

One hypothesis from the results of Chapter 6 is that the information confounded between contrast and spatial pattern might disappear when the responses of multiple neurons are considered. This is not what is found experimentally. Instead, the confounded

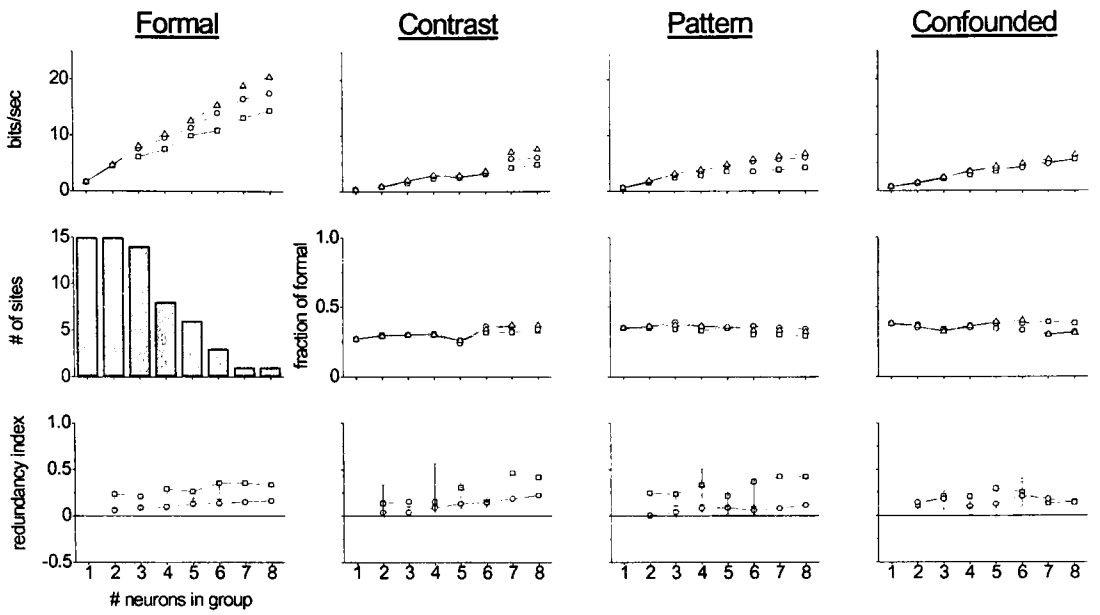

Figure 7.6. *Summary of formal, attribute-specific, and confounded information rates for responses to stationary sinusoidal gratings at five contrasts.* Details as in Fig. 7.5.

information rate, like the formal, contrast-specific, and spatiotemporal pattern-specific information rates, increases in proportion to the number of neurons analyzed together. The fraction of confounded information is largest for the summed-population code and smallest for the labeled-line code, but this is only because the formal information rate for the summed-population code is generally lower. The redundancy indices for confounded information are indistinguishable for the two codes, indicating that whatever is represented by the confounded information does not depend on the identity of the neuron that fires each spike.

Stationary-grating responses. The results for stationary gratings are presented in Fig. 7.6. Here, formal information rates are slightly higher when responses to lower contrasts are included (compare Fig. 7.4, middle column, with Fig. 7.6, left column). The summed-population information rate loses ground to the labeled-line information rate as more neurons are added to the pool, but again only for formal and spatiotemporal pattern-specific information rates. As with m-sequence responses, the contrast-specific and confounded information rates and redundancy indices are largely insensitive to the code. Thus, contrast-specific information in responses to stationary gratings is not likely to be

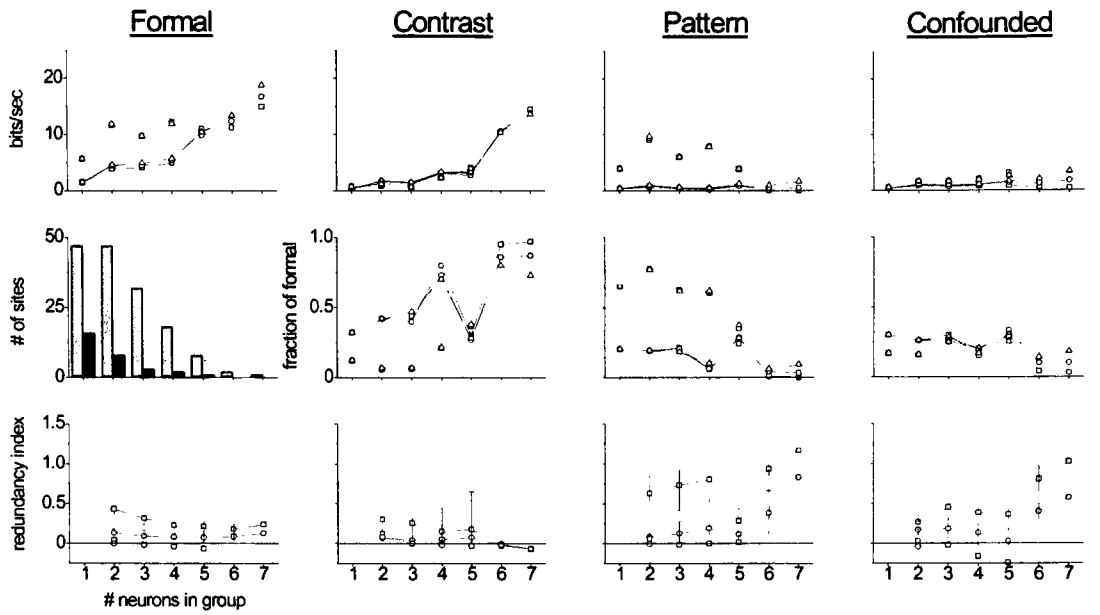

Figure 7.7. *Summary of formal, attribute-specific, and confounded information rates for responses to drifting sinusoidal gratings at five contrasts.* Details as in Fig. 7.5.

encoded by labeled lines, but rather by the average activity over populations of neurons. The fraction of the formal information rate comprised by each type of attribute-specific information is not related to the number of neurons in the combination or to the particular multi-neuronal code.

Drifting-grating responses. Formal and attribute-specific information rates in the drifting-grating responses behave slightly differently (Fig. 7.7). In particular, information rates for the summed-population and labeled-line codes are nearly identical, even with respect to spatiotemporal pattern. This is true for all neurons (solid lines and gray columns), which are dominated by complex cells (see above), and even for combinations of simple cells exclusively (dotted lines and black bars). Despite the similarity of median information rates for the two codes, however, the redundancy indices remain distinct and indicate that information is lost when responses are averaged across neurons. The amount of information loss decreases with larger combinations of neurons, especially for formal and contrast-specific information rates, as indicated by the convergence of the redundancy indices for the two codes. As the number of neurons increases, redundancy indices for both codes decrease for formal and contrast-specific information rates, indicating

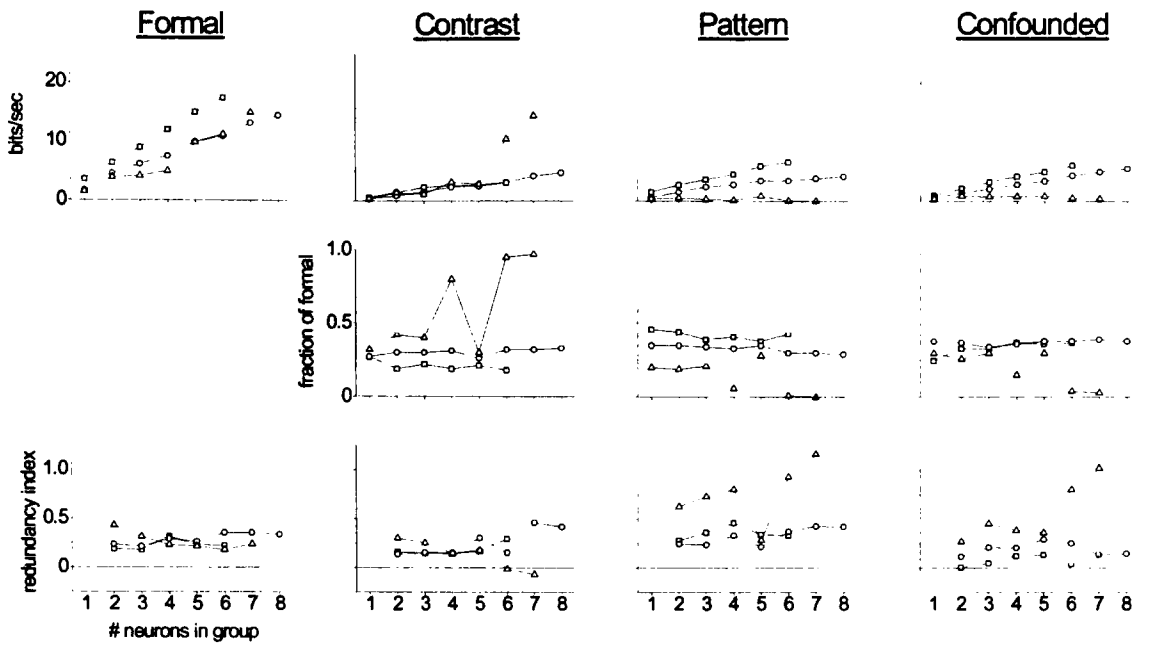

Figure 7.8. *Summary of formal, attribute-specific, and confounded information rates for the summed-population code.* Data taken from Figs. 7.5, 7.6, and 7.7. Medians are first calculated within each recording site; the data in the figure are grand medians across recording sites. Stimuli: m-sequences (□), stationary sinusoidal gratings (○), and drifting sinusoidal gratings (△). **First column**: formal information rates. **Second column**: contrast-specific information rates. **Third column**: spatiotemporal pattern-specific information rates. **Fourth column**: confounded information rates. **Top row**: median information rates, bits/sec, **Middle row**: fraction of formal information rate accounted for by information about specific attributes. **Bottom row**: median redundancy indices. Error bars as in Fig. 7.4.

greater independence across populations of neurons for formal information and for information about contrast. On the other hand, redundancy indices increase with the number of neurons for spatiotemporal pattern-specific and confounded information rates, although not when only simple cells are considered. Consistent with the results of Chapter 6, contrast-specific information is the major contributor to the formal information for all neurons, but spatiotemporal pattern-specific information is the major contributor for exclusive combinations of simple cells.

The similarities and differences among information rates for different stimuli can be better appreciated in Figs. 7.8 and 7.9, in which the data from Figs. 7.5, 7.6, and 7.7 are replotted to facilitate comparisons across stimulus type. In Fig. 7.8, all data are analyzed with the summed-population code—that is, by averaging spike counts across neurons; in Fig. 7.9, the labeled-line code, which keeps track of the neuron of origin of each spike, is used. In both figures, squares (□) represent m-sequence responses, circles (○)

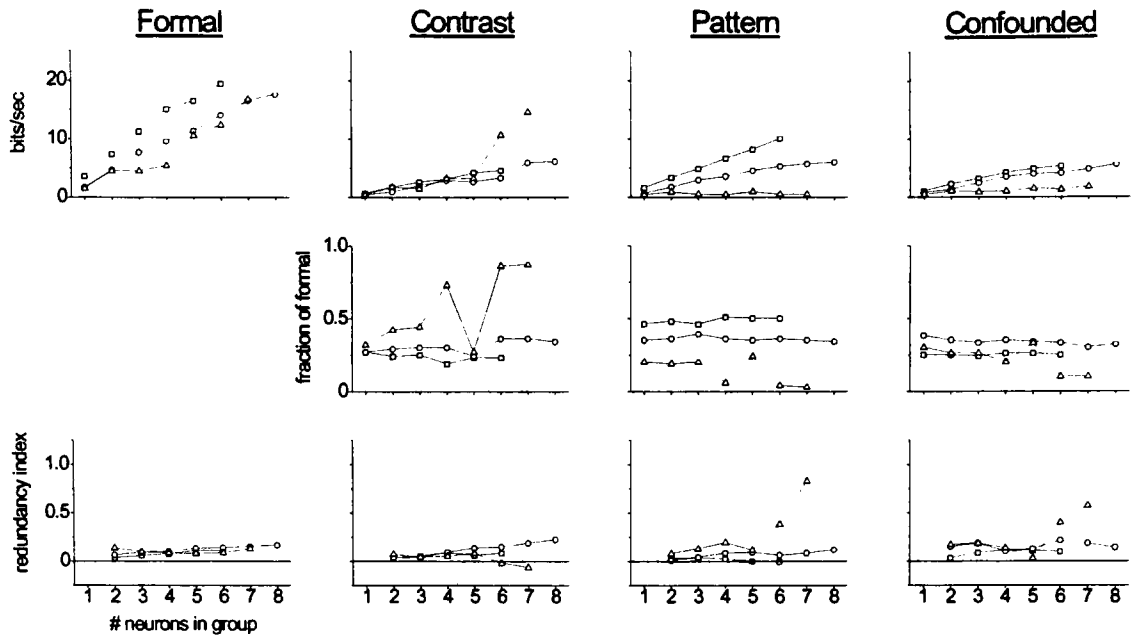

Figure 7.9. *Summary of formal, attribute-specific, and confounded information rates for the labeled-line code.* Details as in Fig. 7.8.

represent stationary-grating responses, and triangles (△) represent drifting-grating responses. These plots reveal that a result obtained from single neurons and discussed in Chapter 6—that the contrast-specific information rate depends little on stimulus type—is borne out when larger groups of neurons are considered (except for the high information rates in the drifting-grating responses of the largest groups of neurons, which are most likely an undersampling artifact). On the other hand, spatiotemporal pattern-specific and confounded information rates depend a great deal on stimulus type for larger groups of neurons, regardless of the code, just as they do for single neurons. Overall, m-sequence responses convey information at the highest rates.

Stationary gratings that vary in spatial phase. In this chapter, two additional sets of stimuli not considered in Chapter 6 are introduced. Both consist of stationary gratings, presented for 237 msec each, that vary along stimulus dimensions other than contrast. The method of determining attribute-specific information is identical to the method used for stimuli that vary in contrast, although the information itself, needless to say, no longer concerns contrast.

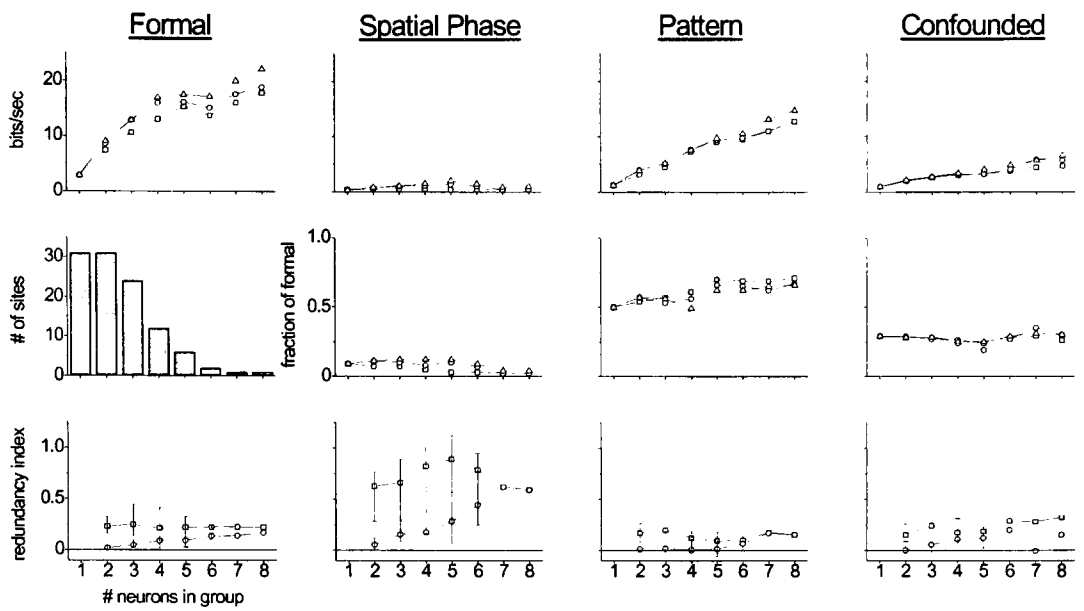

Figure 7.10. *Summary of formal, attribute-specific, and confounded information rates for responses to stationary sinusoidal gratings at 16 spatial phases spaced uniformly around the circle.* Details as in Fig. 7.5.

Fig. 7.10 presents analysis of responses to stationary gratings that vary in spatial phase. V1 neurons—simple cells, in particular—can be exquisitely sensitive to the spatial phase of a stimulus (Hubel and Wiesel, 1962; Movshon et al., 1978b; Victor and Purpura, 1998). Not surprisingly, then, formal information rates are of the same order of magnitude as information rates in responses to stimuli that vary in contrast. The most salient feature of the data presented in Fig. 7.10, however, is that there is evidently very little spatial phase-specific information (second column)—instead, nearly all of the formal information rate is devoted to spatiotemporal pattern or is confounded.[11]

[11] For these responses, and for the responses to stimuli that vary along multiple dimensions (spatial phase, spatial frequency, and orientation), *spatiotemporal pattern*—the term used here to represent aspects of the stimulus that do not vary along the primary axes of the experiment—is a bit of a misnomer. Nonetheless, the nomenclature is preserved for the sake of consistency and in recognition of the fact that spatiotemporal pattern is actually an amalgam of different stimulus features. The finding that information is confounded between spatial phase and spatiotemporal pattern is expected, in part, because spatiotemporal pattern and spatial phase are related stimulus attributes.

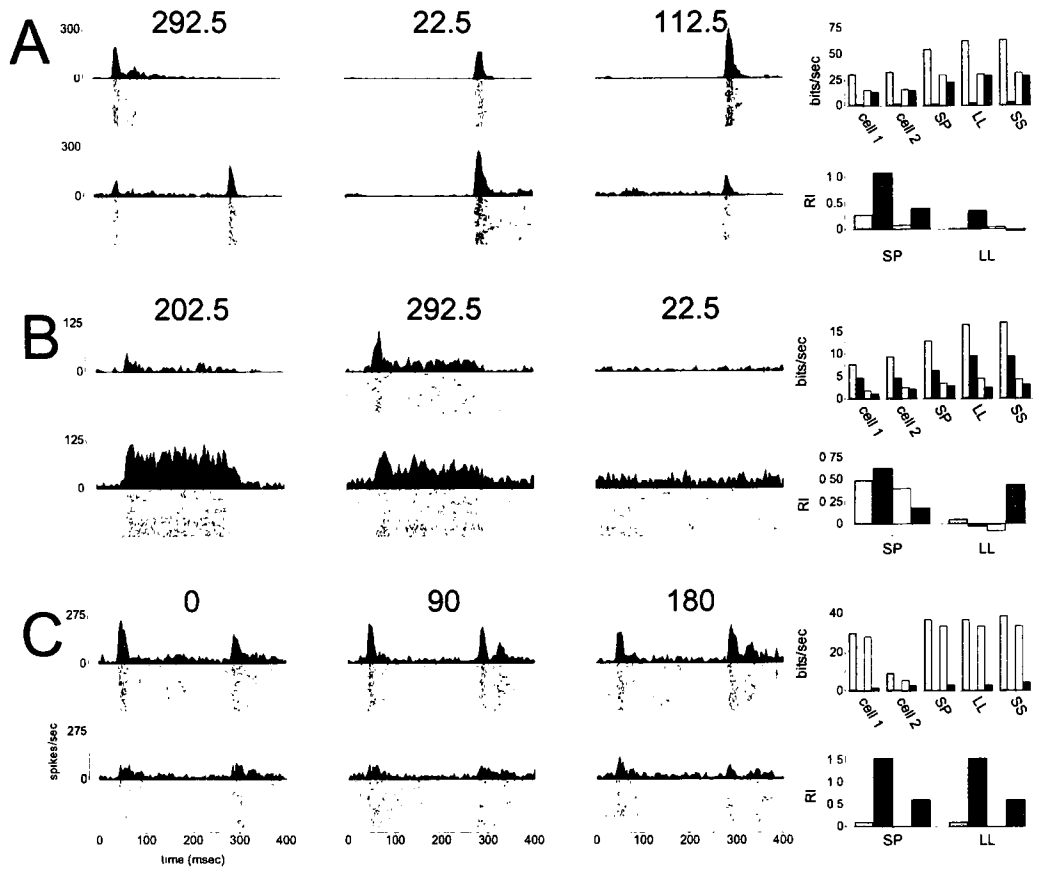

Figure 7.11. *Responses of three pairs of simultaneously recorded neurons to stationary sinusoidal gratings at each of three spatial phases.* Phases selected from 16 in the full stimulus set. Gratings are presented transiently for 237 msec, and spike times that occur between 30 and 300 msec after stimulus onset are analyzed. Displayed spatial phases are 90° apart and are chosen to illustrate typical changes in neuronal responses as the spatial phase is varied. **A**: simple cells. **B**: simple cells. **C**: complex cells. **Columns 1-3**: response histograms (3.7 msec bins) and raster diagrams. **Column 4**: formal information rates (bits/sec) and redundancy indices. Information rates are shown for each cell individually, for the separate sum (SS), and for both multi-neuron codes (summed-population [SP] and labeled-line [LL]). *Light gray*: formal information. *Black*: spatial phase-specific information. *White*: spatiotemporal pattern-specific information. *Dark gray*: confounded information.

A close look into the response patterns of three representative neuron pairs (Fig. 7.11) helps to provide the explanation for the low spatial phase-specific information rates in these responses. For the first pair of neurons (Fig. 7.11A), both simple cells, the effect of changing spatial phase is to alter the relative balance between on and off responses (see Chapter 4), but not the overall firing rate. The spatial phase-specific information rate estimated by the direct method with single time bins is sensitive to the variation in spike-count distribution, averaged across time and trials, that occurs when the spatial phase is

changed. However, the distribution evidently changes little in responses such as these, which therefore yield very low estimates of the spatial phase-specific information rates (right column, black) despite the fact that the neurons clearly encode spatial phase extremely well. Had there been sufficient data to estimate the information rates from much longer time windows (in the terminology of Chapter 6, long words with many, short letters), the on/off response balance would have been captured and the estimated spatial-phase-specific information rates would have been large; indeed, high values of transmitted information about spatial phase are found for these responses when other methods of calculating information are used (Victor et al., 2000). As estimated by the direct method, however, the bulk of the spatial-phase specific information in these responses is most likely captured in the confounded information (right column, dark gray), or is simply missed. This limitation of the direct method more severely affects responses to stimuli that vary along spatial dimensions than along contrast. Since contrast strongly affects the spike-count distributions, and only rarely affects the balance between on and off response components (see Chapter 4), substantial contrast-specific information rates are obtained even from stationary-grating responses (Fig. 7.6, second column).

The second representative pair of neurons (Fig. 7.11B), also simple cells, responds to spatial-phase variations by changing both overall spike count and the balance between the transient (on) and tonic components of the response. The variation in spike-count distribution with spatial phase is captured by the direct method and the resulting spatial phase-specific information rates are relatively large (right column, black). Indeed, for this pair of neurons, the spatial phase-specific information rate is larger than the spatiotemporal pattern-specific (white) and confounded (dark gray) information rates. However, pairs of neurons like these are relatively rare and have little effect on the population medians shown in Fig. 7.10. Far more common are complex cells such as the pair analyzed in Fig. 7.11C, in which the response is virtually insensitive to spatial phase, either in terms of average firing rate or transient/tonic balance. For such neurons, there is virtually no spatial phase-specific information and little confounded information. Rather, nearly all of the formal information rate can be attributed to spatiotemporal pattern (white) and relates to the aspects of the stimulus that cause the transient on/off pattern of firing that is, in this example, spatial-phase independent.

Thus, the direct method with single time bins is likely to detect spatial phase-specific information when spatial phase has a strong effect on average firing rate or spike-count distribution, as in Fig. 7.11B, but not when spatial phase is encoded in some other fashion, as in Fig. 7.11A. In the latter case, most of the spatial phase-specific information is considered by the present analysis to be confounded. Again, this is a limitation of the single time-bin implementation of the direct method for responses such as these, where there are not enough data available to accurately estimate information rates in long time windows. Unlike spatial phase-specific information rates, spatiotemporal pattern-specific information rates are high when the firing rate is strongly modulated in time, regardless of the ways in which those modulations are affected by the spatial phase of the stimulus. Returning to Fig. 7.10, it is clear that the population medians—which discount outliers such as the neurons in Fig. 7.11B—give more weight to responses such as those in Fig. 7.11C.

Also clear from Fig. 7.10 is that higher formal information rates are transmitted by the labeled-line code than by the summed-population code as the number of neurons increases. The effect is less pronounced for the median attribute-specific information rates, especially with respect to spatiotemporal pattern. For the spatial-phase experiments, redundancy indices of formal and attribute-specific information rates for both codes are distinct for small combinations of neurons but converge for larger combinations, suggesting that the benefits of keeping the responses of each neuron separate become smaller when more neurons are considered together. This is less true for stationary gratings that vary in contrast (Fig. 7.6), possibly because contrast can have a profound effect on response latency (Chapter 4).

Stimuli that vary along multiple attributes. Fig. 7.12 shows the results from a series of experiments with stationary-grating stimuli in which spatial phase, spatial frequency, and orientation are all varied. Attribute-specific information rates are calculated from these data, in turn, by collapsing across all but one of those dimensions. The results are similar to those obtained with stationary gratings that vary only in spatial phase (Fig. 7.10) in that the direct method assigns very low spatial phase-, spatial frequency-, and orientation-specific information rates. The explanation for low spatial phase-specific in-

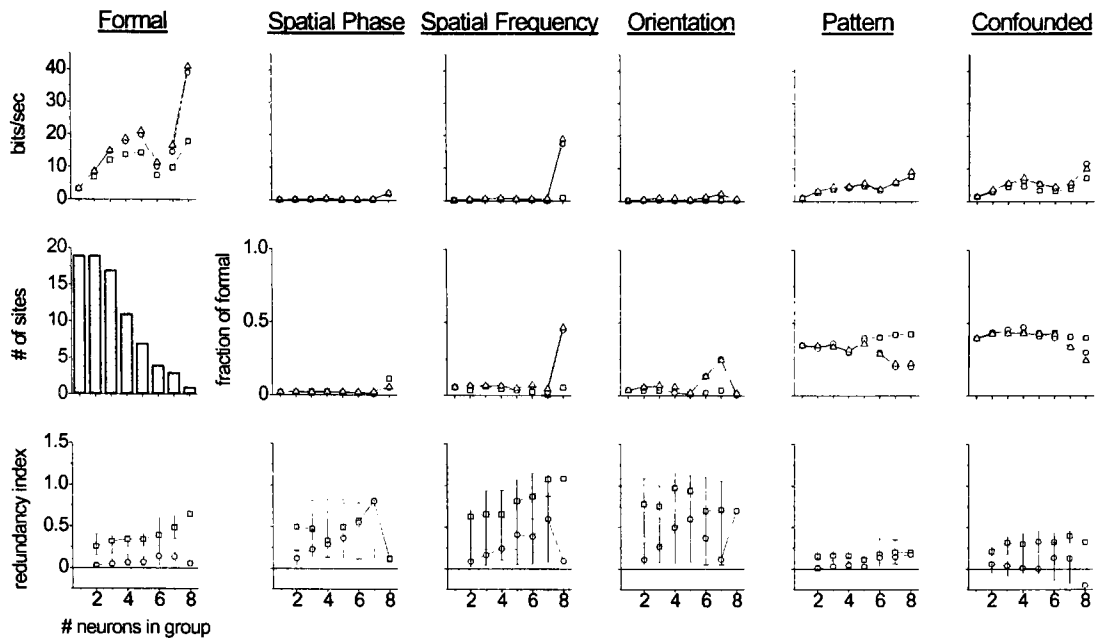

Figure 7.12. *Summary of formal, attribute-specific, and confounded information rates for responses to stationary sinusoidal gratings that vary in spatial phase, spatial frequency, and orientation.* Details as in Fig. 7.5.

formation rates presented above applies directly to these responses as well, but the low spatial frequency- and orientation-specific information rates are more perplexing. It is of course possible—indeed, likely—that information about the four stimulus attributes (spatial phase, spatial frequency, orientation, and spatiotemporal pattern) is confounded in these responses. At the same time, with a complex data set such as this one, it is difficult to disambiguate the effects of information that is truly confounded from the effects of limited sample size and the fact that the measurement of information rates in short time windows is particularly inappropriate for some stimulus attributes.

The summed-population code sacrifices a substantial amount of formal information, the proportion of which increases with the number of neurons, as evidenced by more-or-less parallel redundancy indices for the summed-population and labeled-line codes. Redundancy indices for attribute-specific and confounded information are also higher for the summed-population code, indicating that higher rates of information about the different attributes are conveyed through the labeled-line code. Overall, redundancy indices tend to be close to zero for the labeled-line code when the information rates them-

selves are not too small (i.e., formal, spatiotemporal pattern-specific, and confounded information rates), meaning that responses are nearly independent across populations of neurons. Again, the problems of estimating attribute-specific information rates with the short-time-window direct method for this data set mean that these results can only be considered tentative.

Grand summary of redundancy indices. Median redundancy indices for every recording site, stimulus type, and group size are plotted in Fig. 7.13. In each panel, the summed-population redundancy index is plotted against the corresponding value for the labeled-line code. The redundancy index for the formal information rate (Fig. 7.13A) includes data from all five stimulus types, whereas the attribute-specific redundancy indices (Fig. 7.13B-D) only include data from the experiments in which stimuli were presented at multiple contrasts. For formal, contrast-specific (Fig. 7.13B), and spatiotemporal pattern-specific (Fig. 7.13C) information rates, redundancy indices for the labeled-line code cannot be larger than for the summed-population code, as discussed above. This means that no points can fall above the line of equality. Since the confounded information rate is the difference between the formal information rate and the sum of the contrast- and spatiotemporal pattern-specific information rates, there is no such constraint on its associated redundancy indices (Fig. 7.13D).

Median values for each set of data are plotted as large, open circles in Fig. 7.13. Median values are always significantly closer to zero (independent information transmission) for the labeled-line code than for the summed-population code ($p<0.0001$, Wilcoxon signed-ranks test). However, even for the summed-population code, information conveyed by neighboring neurons is better described as independent rather than synergistic (redundancy index < 0), redundant (~ 1), or confusing (>1). The median difference between labeled-line and summed-population redundancy indices is 0.19 for the formal information rate, 0.086 for the contrast-specific information rate, 0.35 for the spatiotemporal pattern-specific information rate, and 0.057 for the confounded information rate. These results indicate that a substantial amount of formal and attribute-specific information is lost if the neuron of origin of each spike is ignored (the summed-population code) rather than utilized (the labeled-line code) in the decoding process. The effect is most no-

ticeable with information about spatiotemporal pattern. Overall, at 17.7% of the recording sites for the labeled-line code, and at 10.7% for the summed-population code, formal information is encoded synergistically. At the other extreme, no recording site produced evidence of confused or contradictory formal information across neurons for the labeled-line code, and only 1.7% of the sites did so for the summed-population code.

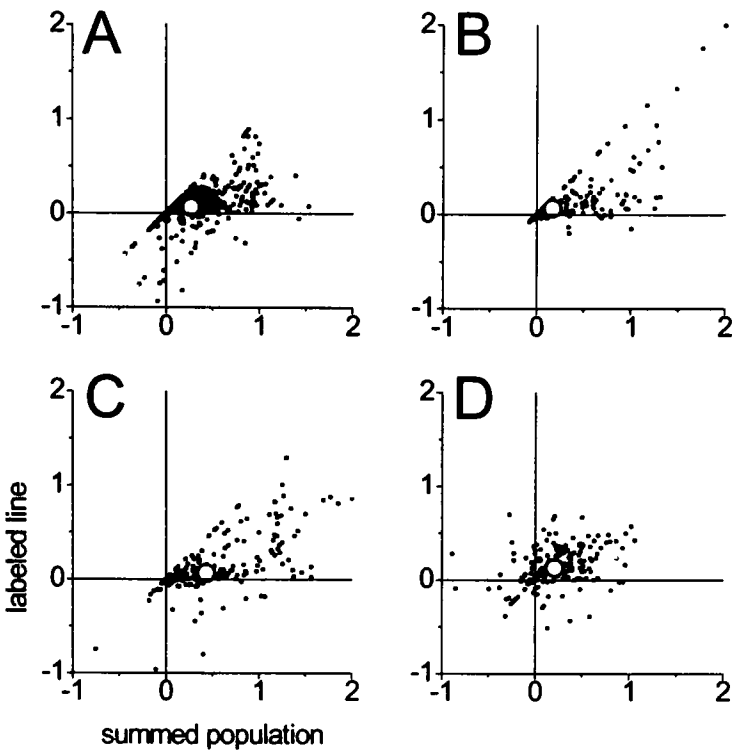

Figure 7.13. *Summary of redundancy indices across all recording sites, neuron combinations of different sizes, and stimulus types.* Redundancy indices for the summed-population code are plotted on the horizontal axis, and for the labeled-line code on the vertical axis. Median values are given by the large, open circles. **A**: formal (includes data from responses to m-sequences, drifting gratings, and stationary gratings that vary in contrast, spatial phase, and along multiple stimulus dimensions). **B**: contrast-specific. **C**: spatiotemporal pattern-specific. **D**: confounded. Data in **B**, **C**, and **D** come from responses to m-sequences, drifting gratings, and stationary gratings at multiple contrasts.

Discussion

The major result in this chapter is that nearby neurons in the primary visual cortex convey nearly independent stimulus-related information. Information is significantly more independent if both the times of spikes and their neurons of origin are considered, rather than the spike times alone (median redundancy index for formal information: 0.062 vs. 0.27, respectively). Thus, any downstream neuron in the cortex that averages incoming spike trains across a set of input neurons, which would occur if inputs were simply combined linearly, is bound to sacrifice a substantial fraction of the available information. Of course, utilization of this extra information is costly in terms of simplicity of processing and energy consumption, because averaging or summing requires fewer computational resources than keeping track of disparate inputs.

It is important to note that both of the across-neuron codes considered here—the labeled-line code, which considers the neuron of origin, and the summed-population code, which ignores it—utilize information encoded into spike times with relatively high precision (typically 2 to 15 msec). This does not reflect an assumption that precise spike times are important for conveying stimulus-related information, although this has been directly shown to be the case in V1 under a variety of stimulus conditions, both in published work (Richmond and Optican, 1990; Victor and Purpura, 1996; Mechler et al., 1998b; Victor and Purpura, 1998) and in other chapters of this thesis. Rather, the direct method of information calculation searches for the bin size that yields the highest information rates. Information rates calculated with wider bins tend to be significantly lower, and when bins of 30 or 60 msec are used, the amount of lost information tends to be even larger than the amount of information that is lost by averaging responses across neurons in a population (not shown).

Formal and attribute-specific information rates. As described in Chapter 6, the direct method can be modified to parse the formal information rates into attribute-specific components. When single time bins are used to calculate information rates, attribute-specific information rates reflect relative changes in spike-count distributions. For example, when responses are collected to stimuli that differ in contrast, the contrast-specific information rate is larger than the spatiotemporal pattern-specific information rate if the effect of contrast on the spike-count distribution is greater than the effect of spatiotemporal pattern.

The results indicate that attribute-specific information, like formal information, can be conveyed nearly independently by V1 neurons that are located in close proximity to one another. For stimuli that vary in contrast and spatiotemporal pattern, contrast-specific information is conveyed independently by both the labeled-line and summed-population codes, so that little contrast-specific information is lost when responses are averaged across neurons. Moreover, the rate of contrast-specific information transmission is independent of stimulus type for combinations of neurons, just as it is for individual neurons (Chapter 6). This suggests that information about contrast does not need to be

decoded in subtle ways; instead, such information is readily extracted from the summed responses of populations of V1 cells.

On the other hand, spatiotemporal pattern-specific information rates are quite sensitive to the code, and information rates are substantially lower when responses are averaged across neurons, especially those that have dissimilar receptive fields such as the pair in Fig. 7.1A. Information rates are also sensitive to stimulus type, which is not surprising since the stimuli used here are characterized by different spatiotemporal patterns but the same, fixed set of contrasts. Although an attempt was made to use the direct method to calculate spatial phase-, spatial frequency-, and orientation-specific information rates from stationary-grating responses (Figs. 7.10 and 7.12), those estimates likely suffer more from the fact that insufficient data are available to estimate the information in long time windows, which would take into account correlations within the response.

Finally, it is not the case that confounded information disappears when the responses of multiple neurons are analyzed together, at least for groups of up to eight nearby neurons. Rather, confounded information rates, like formal and attribute-specific information rates, grow in proportion to the number of neurons in the pool. Moreover, the confounded information rate proves to be independent in the responses of nearby neurons. Thus, it seems that V1 neurons need not encode stimulus-related information along the same cardinal axes that are used by experimenters to describe elementary visual stimuli (i.e., contrast, orientation, spatial frequency, spatial phase). Confounded information may be used by the visual system in the performance of certain functions that require simultaneous knowledge of multiple stimulus attributes and the interactions between them. Alternatively, it may be simply a corollary of the way that the visual system works, without any specific function.

Redundancy indices and the compounded gains of using a labeled-line code. In this chapter, the redundancy index is used as a normalized measure of the degree to which responses of nearby neurons convey independent information. As discussed above, it is found that redundancy indices are significantly lower (more independence) when responses of different neurons are kept separate than when they are averaged together. Low

redundancy indices can have multiple causes. In some cases (for example, Fig. 7.1A), each neuron appears to convey information about different aspects of the stimulus. In others (for example, Fig. 7.1B), high levels of variability may render the simultaneous signals carried by each neuron independent even though their receptive fields are similar, so that the information rate approaches the separate sum even for the summed-population code. This reflects the fundamental property that the independence or redundancy of information transmission takes into account both signal and noise correlations (Panzeri et al., 1999), and that even neurons with similar stimulus selectivities can convey largely independent information (Abbott and Dayan, 1998).

Since information is a logarithmic measure of the number of stimuli that can be distinguished, linear growth in information rates and size-invariance of redundancy indices—typical findings in the neuronal responses presented here—imply that the number of stimuli that can be discriminated increases exponentially with the number of neurons in the pool. In this way, relatively small differences in information rates between the summed-population and labeled-line codes for neuron pairs can yield large differences in the number of stimuli that can be discriminated even by modest-sized populations. Median information rates, which are heavily influenced by sample size and which can be quite variable from site to site (but less variable within a site), do not accurately reflect the compounded benefits of the labeled-line code in comparison with the summed-population code.

Sources of systematic error. Foremost among the potential sources of systematic error in this study is the fact that, by the method presented, information rates can be reliably obtained only for single, short time bins, which means that longer-term correlations in the responses are effectively ignored. This limitation is likely to be more significant for responses to some stimuli, such as stationary gratings that vary in spatial phase, than to others, such as m-sequences that vary in contrast. The inability to measure information in multiple time bins most strongly affects the confounded information rate, as discussed in Results. Although analysis of a few single-neuron data sets for which sufficient data are available (see Chapter 6) suggests that confounded information does not disappear when

multiple time bins are used, it is certainly possible that the confounded information rate in the responses of groups of nearby neurons is actually lower than what is reported here.

A second source of systematic error is in the spike sorting (see Chapter 2). When the direct method is used to estimate information rates, errors consisting of the assignment of individual spikes to the wrong neuron can only reduce the information transmitted through the labeled-line code and cannot affect the information transmitted through the summed-population code. Thus, the analysis method is conservative with respect to errors in spike sorting, so that if spikes have been misidentified, information transmission across nearby neurons under the labeled-line code is likely to be even more independent than what is reported here.

Relationship to earlier work. The results in this chapter, which come from the responses of V1 neurons in macaque monkeys, are consistent with data from recordings of nearby simple cells in the primary visual cortices of kittens and cats (DeAngelis et al., 1999). In that study, which does not explicitly focus on the independence or redundancy of information encoding, such neurons are found to share many but not all response properties. Unfortunately, because of the limitations imposed by the use of the direct method with single time bins, it is difficult to assess accurately whether some stimulus attributes, such as spatial phase, are more independently encoded, as seems to be the case in kittens and cats (DeAngelis et al., 1999), whereas other attributes, such as orientation, are more redundantly encoded. Although summed-population redundancy indices are lower for contrast than for spatiotemporal pattern, consistent with the notion that some stimulus attributes are more independently encoded than others, the labeled-line redundancy indices are approximately equal. Other methods of information calculation, such as the metric-space approach (Victor and Purpura, 1996), are better-suited to examining responses with prominent long-lasting correlations in time.

When information is calculated by more traditional methods, there is always a fixed upper bound given by the logarithm of the number of stimulus categories. As the responses of more neurons are combined, information rates must eventually saturate as they approach this upper bound. In earlier studies (Abbott et al., 1996; Rolls et al., 1997),

saturation effects were seen in part because of the limited stimulus set, in part because attribute-specific information (rather than information rates) was calculated, and in part because responses of sequentially (rather than simultaneously) recorded neurons were used. In the results presented here, there is little evidence of information-rate saturation. To some degree, this is because the variability of neuronal responses themselves, rather than algorithms that assign responses to a fixed number of stimulus categories, is used in the calculation of information rates. Moreover, the entropy of (and hence the upper bound on the information transmitted about) certain stimulus attributes—particularly spatiotemporal pattern—can be quite high. Even in the relatively few cases in which saturation of the median information rates is suggested (such as the responses of simple cells to drifting gratings and of all neurons to stationary gratings that vary in spatial phase), the result may be due more to the undersampling of large clusters of neurons than to true information-rate saturation.

The results in this chapter are generally consistent with earlier estimates of the redundancy of information transmission across pairs of nearby neurons in V1 (Gawne et al., 1996a). In that study, approximately 20% of the information conveyed by supragranular complex cells in response to stationary visual patterns was found to be redundant. This value sits between the median redundancy indices for the labeled-line code and the summed-population code reported in this chapter, despite differences in the type of neuron studied (supragranular complex cells vs. simple and complex cells distributed across all V1 layers in our study), stimulus type (stationary only vs. a variety of spatiotemporal patterns), and method of analysis (principal components and a neural-network method of information calculation vs. spike counts in short time bins and the direct method).

Conclusion and speculation. The present results indicate that the responses of nearby neurons are nearly independent. In sensory cortex, neurons with similar stimulus preferences are located in close proximity to one another, organized in a columnar fashion (Mountcastle, 1957; Hubel and Wiesel, 1962), making it likely that information conveyed by neurons that are more distant is even more independent than information conveyed by nearby neurons. This is true despite the fact that correlations in the background activities of these neurons can extend across millimeters of cortex (Arieli et al., 1996;

Lampl et al., 1999), since such background correlations are strongest for nearby neurons, and since independence of information transmission is affected by both signal and noise correlations (Panzeri and Schultz, 2000). Moreover, the results in this chapter do not bear a fixed relationship to the tight, millisecond-precision response synchronization that has sometimes been observed across large distances in cortex (Singer, 1999a); such synchronization can potentially lead to information transmission that is redundant, independent, or even synergistic.

Although the experiments described in this chapter do not directly address mechanisms of information decoding in real cortical networks, the existence of a complex dendritic machinery for processing synaptic inputs (Segev and Rall, 1998) suggests that neurons may, in fact, take advantage of the extra information available in the labeled-line code. If so, the size of a functional cluster of neurons may be considerably smaller than previously estimated (Zohary et al., 1994; Shadlen and Newsome, 1998). If this is true, it becomes easier to understand how the visual system can solve the problem of discriminating among a large array of stimuli with only a limited number of neurons (Ghose and Maunsell, 1999).

Chapter 8

General Discussion

The research described in this thesis is designed to address fundamental issues in neuroscience: the mechanisms by which neurons in the cerebral cortex represent and encode information. The issues are not specific to visual neuroscience nor even to sensory neuroscience more broadly, although vision is chosen as a model because it is experimentally tractable and because an extensive body of knowledge about cortical mechanisms of visual processing, gleaned from forty years of research, can be brought to bear. The particular goals of the thesis are to determine whether visual information is represented in a rate code, a temporal code, or some mixture of the two, and to investigate whether individual neurons convey information redundant with or distinct from the information conveyed by their neighbors.

Where are we? The major results can be briefly summarized. The responses of visual neurons—in particular, neurons in the primary visual cortex or V1—are demonstrated (in Chapters 2 and 5) to have a statistical structure that can support coding schemes more complicated than a simple transformation of input into the rate of spiking activity. That such schemes exist is demonstrated directly for responses to transiently presented stimuli that vary in contrast (Chapter 4); in particular, one purely temporal aspect of the response—its latency—typically carries more contrast-related information than the response magnitude. Another example (presented in Chapter 5) involves response to flickering checkerboard patterns, in which spikes preceded by interspike intervals of different durations can convey distinct information. These results provide new support for the notion that information can be multiplexed into spike trains of single neurons. Such multiplexing can increase the efficiency of information transfer relative to a

scheme that assigns no more than one message—and a noisy one at that—to each neuron's response.

More general results about neural coding are obtained from the application of information theory to responses of single neurons and groups of nearby neurons. V1 neurons are shown to transmit between 1 and 20 bits/sec of information in response to a variety of visual stimuli (Chapter 6). Information rates are not necessarily higher when the stimulus itself has a complicated, dynamic time course, which means that the spatial attributes of a stimulus can play a major role in determining the time course of the response. Moreover, information about specific stimulus attributes—in particular, contrast and spatiotemporal pattern—can be confounded in the responses of V1 neurons. Such confounded information is also observed in the responses of groups of well-isolated nearby neurons recorded simultaneously on a single tetrode (Chapter 7). But the major result is that averaging together the responses of nearby neurons—commonly suggested as a mechanism by which the brain improves its signal-to-noise ratio—results in a substantial lowering of the rate of information transmission per neuron. Thus, in order to take full advantage of the information encoded into neuronal responses, the brain would have to keep track of the identity of the neuron that fires each action potential.

What does it mean? This thesis is part of a longstanding project of sensory neurophysiologists to characterize the ways in which sensory stimuli are transformed into neuronal responses. The work departs from the norm in terms of the careful attention it pays to defining the relevant aspects of those responses. As such, it avoids the implicit assumption that the firing rate averaged across neurons with similar stimulus selectivities conveys all the important information. The results reveal that an ultimate understanding of the processes by which visual information is represented and encoded in the brain will need to take into account more than just the average firing rate.

The results also reveal aspects of the functional organization of vision, beyond issues of encoding and decoding. It has long been known that neurons with similar stimulus selectivities are located in close proximity to one another, in what is known as a "cortical column" (Mountcastle, 1957; Hubel and Wiesel, 1962). The data described here do

not challenge this hypothesis, but they do provide insight into the segregation of roles within such columns. In particular, the results indicate that the response to a visual stimulus is not randomly distributed across neurons, but rather that each neuron must be considered a distinct functional unit within the network. This is not to say that each neuron, on its own, provides all the necessary information about a visual stimulus, nor does it say that neurons necessarily have particular roles that remain fixed over long periods of time. Rather, since a neuron can encode information about multiple stimulus attributes, it would not be surprising if each neuron simultaneously participates in a number of small networks to represent different aspects of visual stimuli. It is difficult to see how this could be accomplished if a neuron's response were simply averaged together with the responses of other neurons.

What is to be done? It is beyond the scope of this thesis to provide a demonstration that the cerebral cortex actually uses the information encoded into the times of individual action potentials distributed across multiple neurons. There is certainly evidence, referenced at various points in the thesis, that the cortex possesses neural hardware (for example, depressing and facilitating synapses, and coincidence detectors) that could support this. Experiments that would directly resolve the issue—often called the "decoding problem"—have not yet been performed. Such experiments would involve testing whether manipulations that selectively change some aspects of the neural response (for example, the detailed timing of action potentials but not the total number of spikes) can have measurable effects on behavior. There are two primary reasons why these experiments have not yet been performed: first, they are extremely difficult to carry out, in large measure because the cortex—with its billions of neurons and trillions of synapses—is so complicated; and second, a great deal more focus can be applied to the decoding problem once the mechanisms of encoding are known. Hopefully, the results in this thesis advance this knowledge.

In a simpler preparation such as the locust olfactory system, it has been possible, through painstaking experiments, to describe the behavioral consequences of manipulating different aspects of neuronal responses (Stopfer et al., 1997). Elsewhere, the neural code of a specific visual neuron in the brain of the blowfly may be on the verge of being

understood (Ruyter van Steveninck et al., 2000). However, even though it has been demonstrated that the responses of small groups of neurons in a motion-sensitive region of the cortex can be directly correlated with an experimental animal's behavior (Salzman et al., 1992), it is not yet possible to characterize the breadth of real-time inputs to those neurons or to determine the particular aspects of the responses that most influence the animal's behavior. This may begin to change with the advent of new technologies by which the responses of many well-isolated neurons can be monitored simultaneously. Forerunners of these technologies include tetrode recording (see Chapter 2) and *in vivo* optical imaging of individual neurons (Helmchen et al., 1999). In a similar vein, it is possible that through the application of some of the computational techniques developed and described in this thesis to simpler *ex vivo* preparations, such as cortical slices in which small networks of neurons can be identified and examined, new principles will emerge to provide insight into the vastly more complicated realm of *in vivo* dynamics.

For me, therefore, it is a very exciting time to be involved in systems neurophysiology. Detailed knowledge about the functional organization of the mammalian brain, obtained assiduously over the course of four or five decades, is beginning to be transformed into knowledge about the dynamics of information encoding and decoding. I look forward to the completion of this transformation.

References

Abbott LF, Dayan P (1998) The effect of correlated variability on the accuracy of a population code. Neural Comput 11:91-101.

Abbott LF, Rolls ET, Tovee MJ (1996) Representational capacity of face coding in monkeys. Cereb Cortex 6:498-505.

Abeles M (1982) Role of the cortical neuron: integrator or coincidence detector? Isr J Med Sci 18:83-92.

Abeles M, Goldstein MH (1977) Multispike train analysis. Proc IEEE 65:762-773.

Ahmed B, Allison JD, Douglas RJ, Martin KA (1997) An intracellular study of the contrast-dependence of neuronal activity in cat visual cortex. Cereb Cortex 7:559-570.

Albrecht DG (1995) Visual cortex neurons in monkey and cat: effect of contrast on the spatial and temporal phase transfer functions. Vis Neurosci 12:1191-1210.

Albrecht DG, Geisler WS (1991) Motion selectivity and the contrast-response function of simple cells in the visual cortex. Vis Neurosci 7:531-546.

Albrecht DG, Hamilton DB (1982) Striate cortex of monkey and cat: contrast response function. J Neurophysiol 48:217-237.

Alonso JM, Usrey WM, Reid RC (1996) Precisely correlated firing in cells of the lateral geniculate nucleus. Nature 383:815-819.

Arieli A, Sterkin A, Grinvald A, Aertsen A (1996) Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. Science 273:1868-1871.

Bair W, Koch C (1996) Temporal precision of spike trains in extrastriate cortex of the behaving macaque monkey. Neural Comput 8:1185-1202.

Bendat JS, Piersol AG (1971) Random data: Analysis and measurement procedures. New York: Wiley-Interscience.

Berry MJ, 2nd, Meister M (1998) Refractoriness and neural precision. J Neurosci 18:2200-2211.

Berry MJ, Warland DK, Meister M (1997) The structure and precision of retinal spike trains. Proc Natl Acad Sci U S A 94:5411-5416.

Bialek W, Rieke F, Ruyter van Steveninck RRd, Warland D (1991) Reading a neural code. Science 252:1854-1857.

Bland M (1995) An introduction to medical statistics, 2nd Edition. Oxford: Oxford University Press.

Bolz J, Rosner G, Wassle H (1982) Response latency of brisk-sustained (X) and brisk-transient (Y) cells in the cat retina. J Physiol (Lond) 328:171-190.

Bonds AB (1991) Temporal dynamics of contrast gain in single cells of the cat striate cortex. Vis Neurosci 6:239-255.

Borst A, Théunissen FE (1999) Information theory and neural coding. Nat Neurosci 2:947-957.

Brenner N, Strong SP, Koberle R, Bialek W, Ruyter van Steveninck RRd (2000) Synergy in a neural code. Neural Comput 12:1531-1552.

Bura as GT, Albright TD (1999) Gauging sensory representations in the brain. Trends Neurosci 22:303-309.

Bura as GT, Zador AM, DeWeese MR, Albright TD (1998) Efficient discrimination of temporal patterns by motion-sensitive neurons in primate visual cortex. Neuron 20:959-969.

Carandini M, Barlow HB, O'Keefe LP, Poirson AB, Movshon JA (1997a) Adaptation to contingencies in macaque primary visual cortex. Philos Trans R Soc Lond B Biol Sci 352:1149-1154.

Carandini M, Heeger DJ, Movshon JA (1997b) Linearity and normalization in simple cells of the macaque primary visual cortex. J Neurosci 17:8621-8644.

Carandini M, Mechler F, Leonard CS, Movshon JA (1996) Spike train encoding by regular-spiking cells of the visual cortex. J Neurophysiol 76:3425-3441.

Carlton AG (1969) On the bias of information estimates. Psychol Bull 71:108-109.

Cattaneo A, Maffei L, Morrone C (1981) Patterns in the discharge of simple and complex visual cortical cells. Proc R Soc Lond B Biol Sci 212:279-297.

Chance FS, Nelson SB, Abbott LF (1998) Synaptic depression and the temporal response characteristics of V1 cells. J Neurosci 18:4785-4799.

Cleland BG, Enroth-Cugell C (1970) Quantitative aspects of gain and latency in the cat retina. J Physiol (Lond) 206:73-91.

Connors BW, Gutnick MJ (1990) Intrinsic firing patterns of diverse neocortical neurons. Trends Neurosci 13:99-104.

Cover TM, Thomas JA (1991) Elements of Information Theory. New York: Wiley.

Dan Y, Alonso JM, Usrey WM, Reid RC (1998) Coding of visual information by precisely correlated spikes in the lateral geniculate nucleus. Nat Neurosci 1:501-507.

Dean AF (1981) The relationship between response amplitude and contrast for cat striate cortical neurones. J Physiol (Lond) 318:413-427.

Dean AF, Tolhurst DJ (1986) Factors influencing the temporal phase of response to bar and grating stimuli for simple cells in the cat striate cortex. Exp Brain Res 62:143-151.

DeAngelis GC, Ghose GM, Ohzawa I, Freeman RD (1999) Functional micro-organization of primary visual cortex: receptive field analysis of nearby neurons. J Neurosci 19:4046-4064.

DeAngelis GC, Ohzawa I, Freeman RD (1993) Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. II. Linearity of temporal and spatial summation. J Neurophysiol 69:1118-1135.

DeBusk BC, DeBruyn EJ, Snider RK, Kabara JF, Bonds AB (1997) Stimulus-dependent modulation of spike burst length in cat striate cortical cells. J Neurophysiol 78:199-213.

deCharms RC (1998) Information coding in the cortex by independent or coordinated populations. Proc Natl Acad Sci U S A 95:15166-15168.

DeWeese MR, Meister M (1999) How to measure the information gained from one symbol. Network 10:325-340.

Di Lorenzo PM, Hecht GS (1993) Perceptual consequences of electrical stimulation in the gustatory system. Behav Neurosci 107:130-138.

DiCarlo JJ, Lane JW, Hsiao SS, Johnson KO (1996) Marking microelectrode penetrations with fluorescent dyes. J Neurosci Methods 64:75-81.

Efron B, Tibshirani RJ (1998) An Introduction to the Bootstrap. Boca Raton, FL: Chapman & Hall/CRC Press.

Enroth-Cugell C, Robson JG (1966) The contrast sensitivity of retinal ganglion cells of the cat. J Physiol (Lond) 187:517-561.

Fee MS, Mitra PP, Kleinfeld D (1996a) Automatic sorting of multiple unit neuronal signals in the presence of anisotropic and non-Gaussian variability [published erratum appears in J Neurosci Methods 1997 Feb;71(2):233]. J Neurosci Methods 69:175-188.

Fee MS, Mitra PP, Kleinfeld D (1996b) Variability of extracellular spike waveforms of cortical neurons. J Neurophysiol 76:3823-3833.

FitzHugh R (1957) The statistical detection of threshold signals in the retina. J Gen Physiol 40:925-948.

FitzHugh R (1958) A statistical analyzer for optic nerve messages. J Gen Physiol 41:675-692.

Funke K, Worgötter F (1997) On the significance of temporally structured activity in the dorsal lateral geniculate nucleus (LGN). Prog Neurobiol 53:67-119.

Gaska JP, Jacobson LD, Chen HW, Pollen DA (1994) Space-time spectra of complex cell filters in the macaque monkey: a comparison of results obtained with pseudowhite noise and grating stimuli. Vis Neurosci 11:805-821.

Gawne TJ, Kjaer TW, Hertz JA, Richmond BJ (1996a) Adjacent visual cortical complex cells share about 20% of their stimulus-related information. Cereb Cortex 6:482-489.

Gawne TJ, Kjaer TW, Richmond BJ (1996b) Latency: another potential code for feature binding in striate cortex. J Neurophysiol 76:1356-1360.

Gershon ED, Wiener MC, Latham PE, Richmond BJ (1998) Coding strategies in monkey V1 and inferior temporal cortices. J Neurophysiol 79:1135-1144.

Gerstein GL, Mandelbrot B (1964) Random walk models for the spike activity of a single neuron. Biophys J 4:41-68.

Gerstner W, Kreiter AK, Markram H, Herz AVM (1997) Neural codes: Firing rates and beyond. Proc Natl Acad Sci U S A 94:12740-12741.

Gestri G (1978) Dynamics of a model for the variability of the interspike intervals in a retinal neuron. Biol Cybern 31:97-98.

Ghose GM, Maunsell J (1999) Specialized representations in visual cortex: a role for binding? Neuron 24:79-85, 111-125.

Goldman MS, Nelson SB, Abbott LF (1999) Decorrelation of spike trains by synaptic depression. Neurocomputing 26-27:147-153.

Gray CM, König P, Engel AK, Singer W (1989) Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. Nature 338:334-337.

Gray CM, Maldonado PE, Wilson M, McNaughton B (1995) Tetrodes markedly improve the reliability and yield of multiple single-unit isolation from multi-unit recordings in cat striate cortex. J Neurosci Methods 63:43-54.

Gross C (1998) Brain, Vision, Memory. Cambridge, MA: MIT Press.

Gupta A, Wang Y, Markram H (2000) Organizing principles for a diversity of GABAergic interneurons and synapses in the neocortex. Science 287:273-278.

Harris KD, Henze DA, Csicsvari J, Hirase H, Buzsaki G (2000) Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements. J Neurophysiol 84:401-414.

Hartline HK (1938) The response of single optic nerve fibers of the vertebrate eye to illumination of the retina. Am J Physiol 121:400-415.

Hartline HK, Ratliff F (1958) Spatial summation of inhibitory influences in the eye of *Limulus*, and the mutual interaction of receptor units. J Gen Physiol 41:1049-1066.

Heeger DJ (1992) Normalization of cell responses in cat striate cortex. Vis Neurosci 9:181-197.

Heller J, Hertz JA, Kjaer TW, Richmond BJ (1995) Information flow and temporal coding in primate pattern vision. J Comput Neurosci 2:175-193.

Helmchen F, Svoboda K, Denk W, Tank DW (1999) In vivo dendritic calcium dynamics in deep-layer cortical pyramidal neurons. Nat Neurosci 2:989-996.

Hopfield JJ (1995) Pattern recognition computation using action potential timing for stimulus representation. Nature 376:33-36.

Hubel DH, Wiesel TN (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J Physiol (Lond) 160:106-154.

Hubel DH, Wiesel TN (1977) Ferrier lecture. Functional architecture of macaque monkey visual cortex. Proc R Soc Lond B Biol Sci 198:1-59.

Ikeda H, Wright MJ (1975) Spatial and temporal properties of 'sustained' and 'transient' neurones in area 17 of the cat's visual cortex. Exp Brain Res 22:363-383.

Jahnsen H, Llinás R (1984) Electrophysiological properties of guinea-pig thalamic neurones: an in vitro study. J Physiol (Lond) 349:205-226.

Kaplan E, Purpura K, Shapley RM (1987) Contrast affects the transmission of visual information through the mammalian lateral geniculate nucleus. J Physiol (Lond) 391:267-288.

Keener JP, Hoppensteadt FC, Rinzel J (1981) Integrate-and-fire models of nerve membrane response to oscillatory input. SIAM J Appl Math 41:503-517.

Knight BW (1972a) Dynamics of encoding in a population of neurons. J Gen Physiol 59:734-766.

Knight BW (1972b) The relationship between the firing rate of a single neuron and the level of activity in a population of neurons. Experimental evidence for resonant enhancement in the population response. J Gen Physiol 59:767-778.

Knight BW (2000) Dynamics of encoding in neuron populations: some general mathematical features. Neural Comput 12:473-518.

Knight BW, Brodie SE, Sirovich L (1979) Treatment of nerve impulse data for comparison with theory. Proc Natl Acad Sci U S A 76:6026-6029.

Lampl I, Reichova I, Ferster D (1999) Synchronous membrane potential fluctuations in neurons of the cat visual cortex. Neuron 22:361-374.

Legéndy CR, Salcman M (1985) Bursts and recurrences of bursts in the spike trains of spontaneously active striate cortex neurons. J Neurophysiol 53:926-939.

Lennie P (1981) The physiological basis of variations in visual latency. Vision Res 21:815-824.

Levick WR (1973) Variation in the response latency of cat retinal ganglion cells. Vision Res 13:837-853.

Levine MW (1991) The distribution of the intervals between neural impulses in the maintained discharges of retinal ganglion cells. Biol Cybern 65:459-467.

Lewicki MS (1994) Bayesian modeling and classification of neural signals. Neual Comput 6:1005-1030.

Lewicki MS (1998) A review of methods for spike sorting: the detection and classification of neural action potentials. Network 9:R53-78.

Lisman JE (1997) Bursts as a unit of neural information: making unreliable synapses reliable. Trends Neurosci 20:38-43.

Livingstone MS, Freeman DC, Hubel DH (1996) Visual responses in V1 of freely viewing monkeys. Cold Spring Harb Symp Quant Biol 61:27-37.

Maass W, Zador AM (1999) Dynamic stochastic synapses as computational units. Neural Comput 11:903-917.

MacKay DM, McCulloch WS (1952) The limiting information capacity of a neuronal link [published erratum appears in Bull Math Biophys 1953; 15:107]. Bull Math Biophys 14:127-135.

Maffei L, Fiorentini A (1973) The visual cortex as a spatial frequency analyser. Vision Res 13:1255-1267.

Mainen ZF, Sejnowski TJ (1995) Reliability of spike timing in neocortical neurons. Science 268:1503-1506.

Mandl G (1993) Coding for stimulus velocity by temporal patterning of spike discharges in visual cells of cat superior colliculus. Vision Res 33:1451-1475.

Margulis M, Tang CM (1998) Temporal integration can readily switch between sublinear and supralinear summation. J Neurophysiol 79:2809-2813.

Markram H, Gupta A, Uziel A, Wang Y, Tsodyks M (1998) Information processing with frequency-dependent synaptic connections. Neurobiol Learn Mem 70:101-112.

Maunsell JH, Gibson JR (1992) Visual response latencies in striate cortex of the macaque monkey. J Neurophysiol 68:1332-1344.

Maynard EM, Hatsopoulos NG, Ojakangas CL, Acuna BD, Sanes JN, Normann RA, Donoghue JP (1999) Neuronal interactions improve cortical population coding of movement direction. J Neurosci 19:8083-8093.

McClurkin JW, Optican LM, Richmond BJ, Gawne TJ (1991) Concurrent processing and complexity of temporally encoded neuronal messages in visual perception. Science 253:675-677.

McLaughlin D, Shapley R, Shelley M, Wielaard DJ (2000) A neuronal network model of macaque primary visual cortex (V1): Orientation selectivity and dynamics in the input layer 4Calpha. Proc Natl Acad Sci U S A 97:8087-8092.

Mechler F, Reich DS, Victor JD (1998a) Receptive field nonlinearities in macaque primary visual cortex (V1) are not selectively activated by edges. Soc Neurosci Abstr 24:1979 (#1789.1973).

Mechler F, Victor JD, Purpura KP, Shapley R (1998b) Robust temporal coding of contrast by V1 neurons for transient but not for steady-state stimuli. J Neurosci 18:6583-6598.

Meister M, Lagnado L, Baylor DA (1995) Concerted signaling by retinal ganglion cells. Science 270:1207-1210.

Mel BW (1994) Information processing in dendritic trees. Neural Comput 6:1031-1085.

Merrill EG, Ainsworth A (1972) Glass-coated platinum-plated tungsten microelectrodes. Med Biol Eng 10:662-672.

Milkman N, Schick G, Rossetto M, Ratliff F, Shapley R, Victor J (1980) A two-dimensional computer-controlled visual stimulator. Behav Res Meth Instrum 12:283-292.

Miller GA (1955) Note on the bias on information estimates. Information Theory in Psychology: Problems and Methods II-B:95-100.

Mountcastle VB (1957) Modality and topographic properties of single neurons of cat's somatic sensory cortex. J Neurophysiol 20:408-434.

Movshon JA, Thompson ID, Tolhurst DJ (1978a) Receptive field organization of complex cells in the cat's striate cortex. J Physiol (Lond) 283:79-99.

Movshon JA, Thompson ID, Tolhurst DJ (1978b) Spatial summation in the receptive fields of simple cells in the cat's striate cortex. J Physiol (Lond) 283:53-77.

Mukherjee P, Kaplan E (1995) Dynamics of neurons in the cat lateral geniculate nucleus: in vivo electrophysiology and computational modeling. J Neurophysiol 74:1222-1243.

Müller JR, Mehta AB, Krauskoph J, Lennie P (1999a) Why onset transients in responses of V1 neurons provide the most reliable information. Soc Neurosci Abstr 25:1549 (#1619.1511).

Müller JR, Metha AB, Krauskopf J, Lennie P (1999b) Rapid adaptation in visual cortex to the structure of images. Science 285:1405-1408.

Nowak LG, Sanchez-Vives MV, McCormick DA (1997) Influence of low and high frequency inputs on spike timing in visual cortical neurons. Cereb Cortex 7:487-501.

Ohzawa I, Sclar G, Freeman RD (1982) Contrast gain control in the cat visual cortex. Nature 298:266-268.

Optican LM, Richmond BJ (1987) Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. III. Information theoretic analysis. J Neurophysiol 57:162-178.

Panzeri S, Schultz SR (2000) A unified approach to the study of temporal, correlational and rate coding. Neural Comput in press.

Panzeri S, Schultz SR, Treves A, Rolls ET (1999) Correlations and the encoding of information in the nervous system. Proc R Soc Lond B Biol Sci 266:1001-1012.

Panzeri S, Treves A (1996) Analytical estimates of limited sampling biases in different information measures. Network 7:87-100.

Papoulis A (1991) Probability, random variables, and stochastic processes, 3rd Edition. New York: McGraw-Hill.

Peters A (1987) Number of neurons and synapses in primary visual cortex. In: Cerebral Cortex (Jones EG, Peters A, eds), pp 267-294. New York: Plenum.

Press WH, Teukolsky SA, Vetterling WT, Flannery BP (1992) Numerical Recipes in C: The Art of Scientific Computing, 2nd Edition. Cambridge, UK: Cambridge University Press.

Procaccia I, Thomae S, Tresser C (1987) First-return maps as a unified renormalization scheme for dynamical systems. Phys Rev A 35:1884-1900.

Purpura K, Chee-Orts MN, Optican LM (1993) Temporal encoding of texture properties in visual cortex of awake monkey. Soc Neurosci Abstr 19:#315.315.

Ramón y Cajal S (1989) Recollections of My Life. Cambridge, MA: MIT Press.

Rebrik SP, Wright BD, Emondi AA, Miller KD (1999) Cross channel correlations in tetrode recordings: implications for spike-sorting. Neurocomputing 26-27:1033-1038.

Regan D, Silver R, Murray TJ (1977) Visual acuity and contrast sensitivity in multiple sclerosis--hidden visual loss: an auxiliary diagnostic test. Brain 100:563-579.

Reich DS, Mechler F, Purpura KP, Victor JD (2000) Interspike intervals, receptive fields, and information encoding in primary visual cortex. J Neurosci 20:1964-1974.

Reich DS, Victor JD, Knight BW (1998) The power ratio and the interval map: spiking models and extracellular recordings. J Neurosci 18:10090-10104.

Reich DS, Victor JD, Knight BW, Ozaki T, Kaplan E (1997) Response variability and timing precision of neuronal spike trains in vivo. J Neurophysiol 77:2836-2841.

Reid RC, Alonso JM (1996) The processing and encoding of information in the visual cortex. Curr Opin Neurobiol 6:475-480.

Reid RC, Victor JD, Shapley RM (1992) Broadband temporal stimuli decrease the integration time of neurons in cat striate cortex. Vis Neurosci 9:39-45.

Reid RC, Victor JD, Shapley RM (1997) The use of m-sequences in the analysis of visual neurons: linear receptive field properties. Vis Neurosci 14:1015-1027.

Reinagel P, Godwin D, Sherman SM, Koch C (1999) Encoding of visual information by LGN bursts. J Neurophysiol 81:2558-2569.

Reinagel P, Reid RC (2000) Temporal coding of visual information in the thalamus. J Neurosci 20:5392-5400.

Reyes A, Lujan R, Rozov A, Burnashev N, Somogyi P, Sakmann B (1998) Target-cell-specific facilitation and depression in neocortical circuits. Nat Neurosci 1:279-285.

Richmond BJ, Optican LM (1987) Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. II. Quantification of response waveform. J Neurophysiol 57:147-161.

Richmond BJ, Optican LM (1990) Temporal encoding of two-dimensional patterns by single units in primate primary visual cortex. II. Information transmission. J Neurophysiol 64:370-380.

Riehle A, Grün S, Diesmann M, Aertsen A (1997) Spike synchronization and rate modulation differentially involved in motor cortical function. Science 278:1950-1953.

Rieke F, Warland D, Ruyter van Steveninck RRd, Bialek W (1997) Spikes: Exploring the Neural Code. Cambridge, MA: MIT Press.

Rolls ET, Treves A, Tovee MJ (1997) The representational capacity of the distributed encoding of information provided by populations of neurons in primate temporal visual cortex. Exp Brain Res 114:149-162.

Ruyter van Steveninck RRd, Bialek W (1988) Real-time performance of a movement-sensitive neuron in the blowfly visual system: coding and information transfer in short spike sequences. Proc R Soc Lond B Biol Sci 234:379-414.

Ruyter van Steveninck RRd, Borst A, Bialek W (2000) Real time encoding of motion: answerable questions and questionable answers from the fly's visual system. In: Motion vision in the real world: a survey of ecological and computational constraints (Zanker J, Zeil J, eds). New York: Springer.

Ruyter van Steveninck RRd, Lewen GD, Strong SP, Koberle R, Bialek W (1997) Reproducibility and variability in neural spike trains. Science 275:1805-1808.

Sahani M, Pezaris JS, Andersen RA (1998) On the separation of signals from neighboring cells in tetrode recordings. In: Advances in Neural Information Processing Systems (Jordan MI, Kearns MJ, Solla SA, eds). Cambridge, MA: MIT Press.

Salzman CD, Murasugi CM, Britten KH, Newsome WT (1992) Microstimulation in visual area MT: effects on direction discrimination performance. J Neurosci 12:2331-2355.

Schiff ND, Purpura KP, Victor JD (1999) Gating of local network signals appears as stimulus-dependent activity envelopes in striate cortex. J Neurophysiol 82:2182-2196.

Sclar G (1987) Expression of "retinal" contrast gain control by neurons of the cat's lateral geniculate nucleus. Exp Brain Res 66:589-596.

Sclar G, Lennie P, DePriest DD (1989) Contrast adaptation in striate cortex of macaque. Vision Res 29:747-755.

Segev I, Rall W (1998) Excitable dendrites and spines: earlier theoretical insights elucidate recent direct observations. Trends Neurosci 21:453-460.

Segundo JP, Moore GP, Stensaas LJ, Bullock TH (1963) Sensitivity of neurones in Aplysia to temporal patterns of arriving impulses. J Exp Biol 40:643-667.

Sestokas AK, Lehmkuhle S (1986) Visual response latency of X- and Y-cells in the dorsal lateral geniculate nucleus of the cat. Vision Res 26:1041-1054.

Shadlen MN, Newsome WT (1994) Noise, neural codes and cortical organization. Curr Opin Neurobiol 4:569-579.

Shadlen MN, Newsome WT (1998) The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. J Neurosci 18:3870-3896.

Shapley RM, Victor JD (1978) The effect of contrast on the transfer properties of cat retinal ganglion cells. J Physiol (Lond) 285:275-298.

Shapley RM, Victor JD (1981) How the contrast gain control modifies the frequency responses of cat retinal ganglion cells. J Physiol (Lond) 318:161-179.

Sherman SM (1996) Dual response modes in lateral geniculate neurons: mechanisms and functions. Vis Neurosci 13:205-213.

Singer W (1999a) Neuronal synchrony: a versatile code for the definition of relations? Neuron 24:49-65, 111-125.

Singer W (1999b) Time as coding space? Curr Opin Neurobiol 9:189-194.

Singer W, Engel AK, Kreiter AK, Munk MHJ, Neuenschwander S, Roelfsma P (1997) Neuronal assemblies: necessity, signature and detectability. Trends in Cognitive Sciences 1:252-261.

Singer W, Gray CM (1995) Visual feature integration and the temporal correlation hypothesis. Annu Rev Neurosci 18:555-586.

Skottun BC, De Valois RL, Grosof DH, Movshon JA, Albrecht DG, Bonds AB (1991) Classifying simple and complex cells on the basis of response modulation. Vision Res 31:1079-1086.

Smith DR, Smith GK (1965) A statistical analysis of the continual activity of single cortical neurones in the cat unanaesthetized isolated forebrain. Biophys J 5:47-74.

Snider RK, Kabara JF, Roig BR, Bonds AB (1998) Burst firing and modulation of functional connectivity in cat striate cortex. J Neurophysiol 80:730-744.

Somers DC, Nelson SB, Sur M (1995) An emergent model of orientation selectivity in cat visual cortical simple cells. J Neurosci 15:5448-5465.

Stopfer M, Bhagavan S, Smith BH, Laurent G (1997) Impaired odour discrimination on desynchronization of odour-encoding neural assemblies. Nature 390:70-74.

Strogatz SH (1994) Nonlinear dynamics and chaos with applications to physics, biology, chemistry, and engineering. Reading, MA: Addison-Wesley.

Strong SP, Koberle R, Ruyter van Steveninck RRd, Bialek W (1998) Entropy and information in neural spike trains. Phys Rev Lett 80:197-200.

Sutter EE (1992) A deterministic approach to nonlinear systems analysis. In: Nonlinear Vision: Determination of Neural Receptive Fields, Function, and Networks (Pinter RB, Nabet B, eds), pp 171-220. Boca Raton: CRC Press.

Tarczy-Hornoch K, Martin KA, Stratford KJ, Jack JJ (1999) Intracortical excitation of spiny neurons in layer 4 of cat striate cortex in vitro. Cereb Cortex 9:833-843.

Théunissen F, Miller JP (1995) Temporal encoding in nervous systems: a rigorous definition. J Comput Neurosci 2:149-162.

Thomson AM (1997) Activity-dependent properties of synaptic transmission at two classes of connections made by rat neocortical pyramidal axons in vitro. J Physiol (Lond) 502:131-147.

Thomson AM (2000) Facilitation, augmentation and potentiation at central synapses. Trends Neurosci 23:305-312.

Tolhurst DJ (1989) The amount of information transmitted about contrast by neurones in the cat's visual cortex. Vis Neurosci 2:409-413.

Tolhurst DJ, Movshon JA (1975) Spatial and temporal contrast sensitivity of striate cortical neurones. Nature 257:674-675.

Tolhurst DJ, Movshon JA, Thompson ID (1981) The dependence of response amplitude and variance of cat visual cortical neurones on stimulus contrast. Exp Brain Res 41:414-419.

Treves A, Panzeri S (1995) The upward bias in measures of information derived from limited data samples. Neural Comput 7:399-407.

Troy JB, Robson JG (1992) Steady discharges of X and Y retinal ganglion cells of cat under photopic illuminance. Vis Neurosci 9:535-553.

Tzonev S, Rebrik S, Miller KD (1997) Response specificity of lateral geniculate nucleus neurons. Soc Neurosci Abstr 23:450 (#177.451).

Usrey WM, Reid RC (1999) Synchronous activity in the visual system. Annu Rev Physiol 61:435-456.

Vaadia E, Haalman I, Abeles M, Bergman H, Prut Y, Slovin H, Aertsen A (1995) Dynamics of neuronal interactions in monkey cortex in relation to behavioural events. Nature 373:515-518.

Van Sluyters RC, Oberdorfer MD (1994) Preparation and maintenance of higher mammals during neuroscience experiments. In:. Washington: National Institutes of Health.

Victor JD (1992) Nonlinear systems analysis in vision: overview of kernel methods. In: Nonlinear Vision: Determination of Neural Receptive Fields, Function, and Networks (Pinter RB, Nabet B, eds), pp 1-37. Boca Raton, FL: CRC Press.

Victor JD (2000) Asymptotic bias in information estimates and the exponential (Bell) polynomials. Neural Comput in press.

Victor JD, Purpura KP (1996) Nature and precision of temporal coding in visual cortex: A metric-space analysis. J Neurophysiol 76:1310-1326.

Victor JD, Purpura KP (1997) Metric-space analysis of spike trains: theory, algorithms and application. Network 8:127-164.

Victor JD, Purpura KP (1998) Spatial phase and the temporal structure of the response to gratings in V1. J Neurophysiol 80:554-571.

Victor JD, Reich DS, Mechler F, Aronov D (2000) Multi-neuronal signaling of spatial phase in V1. Soc Neurosci Abstr 26:2090 (#2779.2099).

Victor JD, Reich DS, Purpura KP, Mechler F (1998) Bursts and "reliable" spikes generated by neurons in V1: relationships and spatiotemporal origins. FASEB: Retinal Neurobiology and Visual Perception .

Wehr M, Pezaris JS, Sahani M (1999) Simultaneous paired intracellular and tetrode recordings for evaluating the performance of spike sorting. Neurocomputing 26-27:1061-1068.

Wiener MC, Oram MW, Richmond BJ (1999) Latency is a better temporal code than principal components. Soc Neurosci Abstr 25:1549 (#1619.1512).

Wiener MC, Richmond BJ (1999) Using response models to estimate channel capacity for neuronal classification of stationary visual stimuli using temporal coding. J Neurophysiol 82:2861-2875.

Yuste R, Majewska A, Cash SS, Denk W (1999) Mechanisms of calcium influx into hippocampal spines: heterogeneity among spines, coincidence detection by NMDA receptors, and optical quantal analysis. J Neurosci 19:1976-1987.

Zohary E, Shadlen MN, Newsome WT (1994) Correlated neuronal discharge rate and its implications for psychophysical performance [published erratum appears in Nature 1994 Sep 22; 371(6495):358]. Nature 370:140-143.