February 25, 2006

Introduction to Computational Neuroscience

Xiao-Jing Wang

Volen Center for Complex Systems, Brandeis University,
Waltham, MA 02254, USA

Correspondence: Xiao-Jing Wang
Volen Center for Complex Systems, MS 013, Brandeis University
415 South Street, Waltham, MA 02254-9110, USA
Tel (781) 736 3147 - Fax (781) 736 2915
Email: xjwang@brandeis.edu

Introduction

Adaptive computation by single synapse and neuron

- Short-term synaptic plasticity
- Neuronal adaptation
- Decorrelation

Coupled neurons and brain rhythms

- Cellular pacemakers
- Synchronization mechanisms

Recurrent neural networks

- Working memory and persistent activity
- Perceptual decision making
- Time integration

Concluding remarks

Introduction

In Computational Neuroscience, also known as Theoretical Neuroscience, one studies how the brain works using theories, models and analysis (Sejnowski et al 1988, Koch and Segev 1998). Because of its emphasis on quantitative methods, this discipline is highly cross-disciplinary, and has flourished in recently years with the participation of physicists, applied mathematicians and engineers, working together with neuroscientists. Neurophysiology is one of the most quantitative branches of Biology, with a long tradition of precise measurements and mathematical analysis. Today, as experimental data are accumulating at a staggering pace, we know more and more details about the "building blocks" of the brain (genes, ion channels, neurons and synapses). This allows us to start to put pieces together, and truly elucidate brain functions in terms of the underlying synaptic circuitry and neurophysiology. In this endeavor, theory and models provide a powerful tool to synthesize our knowledge, test hypotheses and uncover fundamental principles.

There are three types of modeling approaches. First, *descriptive models* are designed to quantitatively characterize experimental data. Signal processing algorithms and stochastic process models for neuronal spike trains belong to this category, so are linear filter models of sensory neurons, or population coding and decoding algorithms (Rieke et al. 1999). Second, *computational theories* aim at explaining brain processes at the functionally level. For instance, Horace Barlow proposed decorrelation, a computation that renders neural coding of sensory information more efficient by reducing redundancy in stimulus inputs, as a general principle for understanding multiple aspects of adaptation in early sensory systems. Statistical Bayesian inference theory argues that neural coding and processing of sensory stimuli depends on the organism's prior knowledge about the environment, hence can be optimized based on stimulus' prior probability distribution (Rao et al 2002). Yet another example is reinforcement learning theory, initially developed in Computer Science, which is now applied to the study of reward-dependent decision behavior. Third, *biophysically-realistic models* are constructed based on the principle that nervous membrane and synapses can be described by equivalent elec-

trical circuits. Well known examples include Hodgkin and Huxley's ionic current model of action potential, and Rall's cable model of neuronal dendritic tree (Koch 1999). With the advance of cellular neurobiology in recent years, one can now build models firmly grounded in the known "hardware" of the brain, not just for single neurons and synapses, but also large-scale networks. Such networks are described using concepts and tools from statistical physics and nonlinear dynamical systems theory. It is sometimes said that "top-down" theories are concerned with uncovering computational principles, whereas "bottom-up" realistic models deal with biological implementations. This dichotomy is a misconception, an indication of the still large existing gap between computational theories and neurobiology. The situation, however, is changing rapidly, as we are entering a new era when computational theories and biologically-constrained models are becoming integral parts of a unified framework.

The physicist Leo Kadanoff is fond of saying: "The World is like an onion". What he means is that we progress in understanding the world by peeling through layers, one at a time, going from observed phenomena at one level to the underlying mechanism at a deeper level. In Neuroscience, there are many levels of investigation, from molecules, neurons, networks, brain systems, to psychology. This chapter will be mainly concerned with neuronal microcircuitry, which is ideally suited for bridging the gap between functions at the network level and underlying biophysical substrates at the cellular level. This is also the level of investigation optimal for close two-way interactions between models and rigorous experimentation. However, the examples discussed below reflect but one particular point of view. To gain a broader perspective about the field, the reader is recommended to consult the textbook by Dayan and Abbott (2001), and recently published collections of articles (Feng 2004, Chow et al 2005).

Structurally, a neuronal circuit is a "network graph", i.e. a collection of "nodes" connected by "links". For a neuronal circuit, nodes are nervous cells and links are made of synaptic connections. Because neurons and synapses can be described as electrical elements, a network is equivalent to a complex electrical circuit, and mathematically described by coupled nonlinear differential equations (dynamical

systems). To understand how such a circuit works, we need to have two kinds of information: the inner working of neurons and synapses, circuit architecture and statistical properties of network connectivity. I will begin by talking about single nodes and isolated links, using examples to show that single neurons and synapses are capable of sophisticated computations. Then, I will discuss how to study networks of interconnected neurons, and for that purpose synchronous oscillations represent an example *par excellence*. Finally, I will cover models of strongly recurrent networks. The discussion on single synapses and cells focuses on sensory adaptation, which can be conceptualized as calculating the time derivative of an external input. By contrast, the discussion on recurrent networks deals with cognitive functions such as working memory and decision making, which can be conceptualized in terms of time integration of inputs. In the brain, it is probably impossible to attribute a particular function completely to either single synapse/neuron dynamics or to a purely network phenomenon. Network behavior is a result of an interplay between cellular processes and synaptic network dynamics.

Adaptive computation by single synapse and neuron

In order for external stimuli to be coded efficiently in the brain, it is desirable that sensory neurons adapt to the natural environment in which the organism lives. A simple but important form of such adaptation is reduction of redundancy, or correlations in the input across space and time. When an input stays relatively constant, our senses become less responsive to it over time. This is accomplished by sensory neurons that become "fatigued" and show a decreasing time course of spiking activity. Neurons at different stages of processing in the brain may be specialized in adaptation to distinct features of inputs. For example, in the mammalian visual system, retina is responsible for adaptation to light luminance, whereas adaptation to contrast (relative luminance of an object with respect to its background) and stimulus pattern take place mostly in the primary visual cortex.

Short-term synaptic plasticity

Adaptation of a neuron's activity can result from either short-term plasticity of

afferent synapses, or intrinsic membrane properties in the cell. Short-term synaptic plasticity refers to the phenomenon that synaptic transmission between two neurons is not static, but can either decrease (Fig. 1A left) or increase (Fig. 1A right) depending on the history of the presynaptic neuron's firing activity. Consider a presynaptic terminal with a certain number $N \leq N_0$ of readily releasable vesicles of neurotransmitters. Upon arrival of an action potential, a vesicle of neurotransmitters is released with certain probability P_{rel}, hence in average $P_{rel}N$ vesicles are released, and N is reduced to $N - P_{rel}N = f_D N$ ($f_D = 1 - P_{rel}$). Between two stimuli, the pool of docked vesicles is gradually refilled, with a recovery time constant τ_D. Hence, the postsynaptic potential response (proportional to $P_{rel}N$) is initially $P_{rel}N_0$ but can exhibit short-term depression (Fig. 1A left) as N decreases with time, typically on a timescale of hundreds of milliseconds. On the other hand, P_{rel} depends on the dynamical state of the transmitter release machinery, and can increase over time in response to a stimulus train, leading to short-term facilitation (Fig. 1A right). Interestingly, experimentalists and theorists recently showed that synapses endowed with short-term depression are sensitive to changes in the input but not its steady state. This can be seen using a simple phenomenological model (Dayan and Abbott 2001). Let $D = N/N_0$, then after each release in average D is reduced to $f_D D$ with $f_D \leq 1$, and between release events D recovers towards one with a time constant τ_D. For a Poisson spike train at rate r, the dynamical equation for D can be approximately written as

$$\frac{D}{dt} = -(1 - f_D)rD + \frac{1 - D}{\tau_D}$$

The steady state of this equation is given by $dD/dt = 0$, which yields

$$D_{ss} = \frac{1}{1 + (1 - f_D)r\tau_D}$$

Therefore, for the input rate r higher than $1/(1-f_D)\tau_D$, $Dss \simeq 1/((1-f_D)r\tau_D)$, inversely proportional to r. For instance, if $f_D = 0.4$ and $\tau_D = 500$ ms, then $1/((1 - f_D) = 1/0.3 \simeq 3$ Hz. The transmission rate (release per unit time) is $P_{rel}rD$, its steady state becomes independent of input rate r when the latter is

above $1/(1-f_D)\tau_D$. These model predictions have been confirmed by experimental observations on synapses between neocortical neurons (Abbott et al. 1997, Tsodyks and Markram 1997). Functionally, this means that the impact on the postsynaptic cell is not sensitive to the steady state level of presynaptic activity. On the other hand, when there is a sudden change in the input, the synapse can detect it by a transient response before the slow depression has the time to react. This is illustrated in Fig. 1B, where the transmission rate, $P_{rel}rD$, is plotted in response to a series of steps in the presynaptic firing rate. Note first that regardless of the presynaptic firing rate 25, 100, 10, 40 Hz, the steady state $P_{rel}rD$ is quite similar, as predicted before. Second, when there is a sudden jump in the presynaptic rate, because the depression process takes time to react, the synapse can still signal input change by a rigorous but transient response. When r is decreased or increased to $r + \Delta r$, D briefly remains $\sim 1/r$ (where r is the rate prior to the jump), thus the response jump is $\sim \Delta r/r$, reflecting the change in the input relative to the baseline, rather than the absolute amount of input change. This feature is apparent in Fig. 1B, where the transient response is similar for the transitions from 25 Hz to 100 Hz and from 10 Hz to 40 Hz. In both cases, $\Delta r/r = 3$. Moreover, the time constant governing the decay process toward a steady state is given by $\tau_D/(1+(1-f_D)r\tau_D)$, which is shorter at a higher input rate r. This explains why the transient response is longer lasting for the 10 to 40 Hz transition than for the 25 to 100 Hz transition. Therefore, short-term depression provides a synaptic mechanism for adaptation, i.e. for postsynaptic neurons to be responsive to changes in the input but not sensitive to constant input level.

Neuronal adaptation

An alternative mechanism is membrane dynamics of single neurons. Contrast adaptation by visual neurons offers an example of this scenario. Our visual system can detect an object only when its contrast (difference in luminance relative to the surrounding background) is above a threshold. This threshold is dynamically adapted to the natural environment. For instance, after being exposed to a high contrast stimulus, the visual system becomes less sensitive to the same stimulus

at low contrast. Physiological investigations revealed that presentation of a high contrast visual stimulus results in a prolonged hyperpolarization of neurons in the primary visual cortex, and this reduced excitability correlates with the persistent decrease in sensitivity following adaptation. The long (seconds) hyperpolarization appears to be mediated to a large extent by a Na^+-activated K^+ current (I_{KNa}) in visual cortical neurons. What happens is that during a high contrast stimulation, prolonged firing of action potentials leads to a gradual Na^+ influx into the cell through the opening of sodium channels. Intracellular sodium accumulation eventually activates a potassium conductance, which gives rise to hyperpolarization of neuronal membrane and a reduced responsiveness of the cell. After the offset of high contrast stimulus, sodium ions are slowly extruded from the cell, probably through a Na^+/K^+ ionic pump. This process underlies recovery of neuronal responsiveness with a time constant of about ten seconds.

Computational approach has been used to test this mechanism which has supporting evidence but not conclusive experimental proof. A single neuron model was constructed that includes a fast sodium current (I_{Na}) and a potassium delayed rectifier (I_K) for action potential generation, an I_{KNa}, and intracellular Na^+ dynamics (influx through I_{Na} and outflux through a Na^+/K^+ pump) (Wang et al. 2003). Fig. 2A shows a computer simulation of this model. The input is a sinusoidal current (at 2 Hz) mimicking a periodic visual stimulus such as a moving grating pattern. Input amplitude represents stimulus contrast. Initially, the model neuron responds to a low amplitude input at a constant rate. During a high amplitude input, the firing rate transiently increases, then decays to a lower steady state level according to a time course in parallel with that of intracellular Na^+ accumulation, hence activation of I_{KNa}. After the stimulus returns from the high to the original low amplitude, the membrane potential is hyperpolarized thus the model neuron is incapable of spiking response for seconds. Slowly, however, intracellular Na^+ decays back to baseline, and the model neuron resumes the same activity level as before. Thus, the model captures essential observations from neurons in the primary visual cortex.

Decorrelation

It has long been assumed that neuronal adaptation provides a mechanism for a decorrelation operation that reduces input redundancy in time. If I_{KNa} subserves adaptation, is it capable of input decorrelation as well? Specifically, light intensity signals of optical flow, observed with a narrow visual field, exhibit correlations at all timescales. This is quantified by a scale-free power spectrum, with the power growing like $1/f$ at decreasing frequencies f (red curve in Fig. 2B), characterizing correlations at increasingly longer timescales. By contrast, a completely uncorrelated signal has a constant power spectrum, i.e. the same power at all frequencies (hence the name "white noise"). How would a cell endowed with adaptation respond to an 1/f-type input? This question can be investigated in model simulations, using stochastic inputs with strong time correlations. When an $1/f$ stochastic input is used as injected current to drive the model neuron, long-term correlations of the spiking activity are suppressed, so that the power spectrum of the output becomes essentially flat ("whitened") at frequencies below 1 Hz (blue curve, Left panel in Fig. 2B). Decorrelation takes place due to the negative feedback mediated by I_{KNa}: a higher firing activity leads to a larger I_{KNa}, hence a stronger negative signal to subtractively counterbalance the external input drive. Decorrelation takes place on a timescale of a second or longer, hence whitening of the power spectrum below 1 Hz. Therefore, a single neuron mechanism can subserve decorrelation function. This model prediction was directly tested in an *in vitro* preparation of ferret cortical slices. When a real neuron of primary visual cortex receives a stochastic $1/f$-type input, the output of the neuron exhibits the same kind of decorrelation as the model (Right panel, Fig. 2B). This experiment demonstrates that decorrelation can definitely occur in single cortical cells. Whether the same mechanism is responsible for adaptation at the psychological level remains an open question. Nonetheless, it is clear that single neuron dynamics is capable of complex computations likely important to behavior. Furthermore, this example illustrates a two-way process of interaction between experimentation and modeling: a model is built on experimental findings and gives rise to a prediction that is later tested and confirmed by new laboratory experiments.

Coupled neurons and brain rhythms

I have discussed how to model a single synapse or a single neuron. Now I turn to studies of neurons coupled by synapses. This is a vast topic, since there are virtually unlimited varieties of coordinated firing patterns in large-scale neural networks. To be specific, I will focus on synchronous neural oscillations, a type of neural population dynamics that has been intensively studied by modelers and experimentalists. Rhythmicity is obviously essential for certain brain functions, like locomotion produced by central pattern generators (Marder and Calabrese 1996). Diverse types of synchronous rhythms have also been observed in the mammalian cortex, such as spontaneous spindle wave during quiet sleep or gamma ($\sim$ 40 Hz) oscillation induced by external stimulation. Another prominent example is theta ($\sim$ 8 Hz) rhythm observed in hippocampus during exploratory movement and spatial navigation. "Place cells" in hippocampus code an animal's current (or immediate future) spatial location. Spikes of a place cell are locked to a particular phase of the theta cycle that systematically changes as the animal moves across the place field of that cell, indicating that precise timing of spike discharges relative to a coherent oscillation plays a role in information coding and processing.

Typically, a coherent neural network rhythm is generated within a relatively localized brain circuit, which nevertheless is composed of many hundreds or thousands of neurons. Therefore, studies of the neural mechanisms underlying brain rhythms provide an excellent venue to investigate how circuit dynamics emerges from the interplay between synaptic and intrinsic cellular properties. There are two general questions in terms of mechanisms of a coherent oscillation: First, what determines its oscillation frequency; are there pacemaker neurons, or is rhythmicity largely a network phenomenon? Second, what are the synaptic mechanisms for network synchronization? I will discuss these two issues in turn.

Cellular Pacemakers

Single neurons in the CNS are endowed with a large repertoire of voltage- and calcium-gated ion channels, distributed across the dendritic and somatic membrane,

which can give rise to complex neuronal dynamics. In general, oscillation occurs in a single cell, when a strong fast positive feedback (generating the rising phase of membrane voltage) interacts with a slower negative feedback (producing the decay phase of the cycle). Positive feedback within a cell can be provided by activation of voltage-gated inward Na^+ or/and Ca^{2+} currents, whereas negative feedback is mediated by either inactivation of inward currents or activation of outward K^+ currents. A special combination of such membrane processes can endow a neuron with the pacemaker property of exhibiting robust oscillations in a well defined frequency range. Fig. 3 shows three such examples.

Spindle oscillations during quiet sleep originate in the thalamus. It was discovered by H. Jahnsen and R. Llinás that thalamocortical projection cells show two modes of firing patterns: upon depolarization they discharge single spikes tonically; whereas upon hyperpolarization they fire bursts of spikes, possibly in a rhythmic fashion (Fig. 3A, upper panel). During quiet sleep, thalamic cells are in the burst mode and entrain the spindle oscillations in the entire thalamocortical system. The bursts of spikes are produced by a low-threshold voltage-gated (T-type) Ca^{2+} ion channel I_T, which de-inactivates during hyperpolarization; and a hyperpolarization-activated cation channel I_h. Such bursting mechanism is demonstrated by a single thalamic neuron model in Fig. 3A (lower panel). Intuitively, rhythmic bursting can be generated as follows (Fig. 3A, upper panel) A hyperpolarizing input slowly activates I_h and de-inactivates I_T. The buildup of the I_T eventually leads to a depolarization wave triggering a *rebound burst* of rapid (250-500 Hz) spikes. The burst is terminated by the inactivation of the same I_T at depolarized voltage, and the oscillatory cycle can start over again. The period of oscillation ($\sim$ 100 ms) is determined by the inactivation time constant of I_T and the activation time constant of I_h, during hyperpolarization. Waking process is associated with a switch of thalamic cells from the burst to tonic firing mode, due to an increase in the neuromodulatory (cholinergic, noradrenergic, and other) inputs.

Gamma ($\sim$ 40 Hz) rhythm is commonly observed in waking and behaving states. In the neocortex, intrinsic γ oscillations have been reported in a subclass of neurons, called 'chattering cells' (Fig. 3B, upper panel). These cells display repetitive burst

firing in the γ frequency range, with intraburst spike rates of 300-500 Hz. A compartmental model suggests that the fast rhythmic bursting in chattering neurons is generated by a Ca^{2+}-independent ionic mechanism (Fig. 3B, lower panel). Instead, it relies on voltage-gated Na^+ currents in the dendrite. In this scenario, perisomatic action potentials propagate back to the dendritic sites, where a Na^+-dependent slow depolarization is produced, which in turn triggers more spikes in the soma. This somato-dendritic "ping-pong" interplay underlies a burst of spikes, which is terminated by the activation of a K^+ current. The de-activation of the K^+ current during hyperpolarization leads to the recovery and to the start of a new burst. Experimental evidence in support for such a Na^+-dependent mechanism has recently been reported from cortical slice studies. It still remains to be demonstrated that chattering neurons indeed serve as pacemakers for γ oscillations observed in the neocortex *in vivo*.

Theta (7-10 Hz) rhythm in hippocampus and surrounding limbic structures is believed to depend on the input pathway from the medial septum, where pacemaker-like neural discharges have been observed. There are two major cell types in the septum, which are thought to play distinct roles in the theta rhythm generation: cholinergic cells modulate slowly the excitability of hipoccampal neurons, whereas GABAergic cells play a role of pacemakers. Recent physiological studies have revealed that non-cholinergic (putative GABAergic) neurons in the medial septum display robust intrinsic oscillations in the theta frequency range, where clusters of spikes are inter-nested in time with episodes of *subthreshold membrane oscillations* (Fig. 3C, upper panel). Interestingly, similar membrane oscillations have been observed in single principal neurons of the rat olfactory bulb, which displays prominent γ and θ rhythms. A conductance-based model (Fig. 3C, lower panel) suggests that such intrinsic rhythmicity can be generated by a low-threshold, slowly-inactivating K^+ current I_{KS}. When the cell fires, the I_{KS} slowly de-inactivates during spike afterhyperpolarization, and a sufficiently large I_{KS} terminates the spiking episode. When the cell does not fire spikes, I_{KS} slowly decreases due to inactivation, until the cell is sufficiently recovered and can start to fire again. The subthreshold oscillations are produced by the interplay between a Na^+ current and the low-threshold

activation of I_{KS}. In this scenario, the periodicity of the theta rhythm is largely controlled by a single current (I_{KS}) in septal GABAergic cells. This hypothetical mechanism has not yet been tested experimentally.

Several general comments can be made. First, a single neuron can display different dynamical (e.g. single-spiking and bursting) modes, which depend on the membrane potential level and are under neuromodulatory control. Second, there are at least two general classes of ionic mechanisms for rhythmogenesis, one depends on an interplay between Na^+ and K^+ currents, the other on Ca^{2+} currents. Pacemaker neurons for the γ and θ rhythms of the waking brain seem to rely on Na^+ and K^+ currents, whereas pacemaker neurons for the spindle and delta sleep rhythms depend on Ca^{2+} currents. Third, and finally, subthreshold oscillations and repetitive bursting may have different implications for synchronization of coupled neurons. Subthreshold oscillations can serve as a signal carrier for phase-locking and resonance, by virtue of the cell's sensitivity to small but precisely timed inputs (at the peak of the membrane oscillation cycle). On the other hand, bursts may provide a reliable signal for the rhythmicity to be transmitted across unreliable and facilitating synapses between neurons (Fig. 1A, right panel).

Synchronization mechanisms

A neural circuit, be it thalamic, neocortical or hippocampal, consists of two major cell types: excitatory principal neurons and inhibitory interneurons. It follows that three types of synchronization mechanisms by chemical synapses are conceivable: recurrent excitation between principal neurons, mutual inhibition between interneurons, and feedback inhibition through the excitatory-inhibitory loop.

Recurrent excitation model. Recurrent excitatory connections have been historically the first synchronization mechanism under detailed experimental and computational analyses. This was motivated by the observation that blockade of synaptic inhibition in a cortical network led to extremely synchronous neural firing patterns, resembling epileptic discharges. Intuitively, mutual excitation is expected to synchronize coupled neurons: if cells firing earlier in time excite the others and advance their firing times, a network can thus be brought to fire in phase. However, model

simulations of biophysically realistic coupled neurons have shown that synaptic excitation often delays rather than advances the firing time in the postsynaptic cell, whether mutual excitation can synchronize depends on the intrinsic membrane properties of the constituent neurons (Hansel et al. 1995, van Vreeswijk et al. 1994). In general, synchronization of normal brain rhythms is not realized by excitation alone; it depends critically on synaptic inhibition.

Interneuronal network model. Computational studies have revealed that, surprisingly, reciprocal synaptic inhibition is capable of synchronizing certain rhythmic activities in an interneuronal network (Wang and Rinzel 1992). This finding was derived from an appreciation of the importance of quantitative synaptic current kinetics. If mutual inhibition is instantaneous, coupled neurons obviously would supress each other's firing hence cannot be synchronized to zero phase. However, as it turns out, if the inhibitory synaptic current has a rise time, and decay slowly compared to the intrinsic neuronal time constant, then neurons can recover from synaptic inhibition together and fire action potentials synchronously. One feature of this mechanism is that the decay time of the inhibitory synaptic current is comparable with the oscillation period. For example, $GABA_B$ receptor mediated inhibition with a time constant of 100-200 ms could in principle synchronize slow oscillations at a few Hz. $GABA_A$ receptor mediated inhibition with a time constant of about 10 ms is too fast for synchronizing an oscillation at a few Hz, but is sufficiently slow for a 40 Hz oscillation (with a period of about 25 ms) (Fig. 4). Indeed, a physiological study using *in vitro* slices provided evidence that $GABA_A$ receptor mediated inhibition in a hippocampal interneuronal network, without the involvement of pyramidal neurons, could give rise to coherent 40 Hz oscillations (Whittington et al. 1995).

Feedback inhibition model. A competing network mechanism for coherent γ oscillations is based on feedback between excitatory and inhibitory neural populations. W. J. Freeman first proposed this scenario to explain 40 Hz oscillations observed in the olfacory bulb and cortex. Such a scenario does not depend on single neuron's clock-like behavior, but emerges as a large-population phenomenon. Like any strongly nonlinear dynamical system, the interplay between a fast positive feedback (mediated by the AMPA receptors) and a slower negative feedback (mediated by the

$GABA_A$ receptors) readily gives rise to oscillations. This scenario has been demonstrated in hippocampal slices, where spontaneously occurring 40 Hz oscillations have been shown to depend both on the excitatory and inhibitory synaptic transmissions. This experiment can be reproduced robustly in a network of pyramidal cells and interneurons, even in a randomly connected network model with in the presence of a large amount of synaptic noise (Fig. 5). As can be seen in Fig. 5, while the neural population as a whole oscillates in the 40 Hz frequency range, individual neurons fire more randomly and intermittently in time (at about 10 Hz for interneurons and only 2 Hz for principal cells). This is similar to neural firing activities during γ rhythms of the intact brain.

Therefore, coherent brain oscillations primarily rely on synaptic inhibition, within interneuronal networks or/and through feedback between excitatory and inhibitory cells. The conceptual shift from excitation to inhibition, somewhat contrary to intuition, is a direct result of biophysically-realistic modeling work. Moreover, two theoretical frameworks have emerged. One approach considers synchronization in terms of coupled oscillators, with individual neurons firing regularly like a clock. This view is applicable to pacemaker-driven rhythms. By contrast, a different approach describe networks in which single neurons are driven by noise and exhibit highly irregular spiking discharges, and coherent oscillation emerges as a collective population dynamical phenomenon.

Synchronous rhythms represent only one form of network dynamics. Other possible activity patterns are synfire chains, propagating waves, asynchronous stochastic firing, etc. Now I am going to discuss computations that can be performed by recurrent network dynamics.

Recurrent neural networks

Neural processing can proceed in a feedforward fashion. For instance, in the mammalian visual system, light signals propagate from retina to lateral geniculate thalamic nucleus, primary visual cortex, and on to higher visual areas in a series. On the other hand, within a local network neurons interact with each other through recurrent connections that are characterized by feedback loops (from neuron A to

B, then back to A either directly or via other neurons). From anatomical studies we know that neocortical microcircuits are endowed with massive amount of synaptic recurrency. However, it is notoriously difficult to dissect contributions from feedforward and feedback mechanisms to a network's function. For example, in layer 4 of primary visual cortex, the input layer for thalamaic afferents, a neuron is selective for orientation of a light bar, whereas its presynaptic thalamic cells are not. How orientation selectivity emerges has been intensely and fruitfully studied collaboratively by experimentalists and theorists. A typical layer 4 neuron has thousands of synapses, even though less than 1% of them originating from the thalamus, orientation selectivity can still arise from a feedforward mechanism, a feedback mechanism, or a combination of the two (Sompolinsky and Shapley 1997).

In contrast to early sensory systems, cognitive processes are not "enslaved" to the instantaneous information flow from the external world. A central cognitive function is working memory, the brain's ability to actively hold and manipulate information internally, in the absence of external input. The stored information can be a sensory stimulus which guides a prospective action, a delayed perceptual decision or behavioral response. It can also be an item retrieved from long-term memory, for instance when the memory of a face is activated and used in visual search of a friend in a crowd. Neurophysiological studies with behaving monkeys have established that working memory is stored in the form of self-sustained persistent neuronal activity. Mnemonic persistent activity is especially common and robust in the prefrontal cortex, which is critically important to working memory (Goldman-Rakic 1995), but it has also been observed in other brain areas like posterior parietal, inferotemporal and premotor cortices. Persistent activity is clearly not driven directly by external stimulus, thus cannot be explained by feedforward processing.

Moreover, the persistence time (up to 10 seconds) of sustained firing activity during working memory is orders of magnitude longer than the biophysical time constants (tens of milliseconds) of fast electrical signals in neurons and synapses. For this reason, persistent activity is believed to be generated by recurrent network "reverberation" a concept going back to the work in 1930's of R Lorente De Nó. The characteristic horizontal connections found in the superficial layers II-III of

the dorsolateral prefrontal cortex may provide the anatomical substrate for such a recurrent circuit (Goldman-Rakic 1995). This idea is made precise in theoretical work where persistent activity is described as "dynamical attractors" (Amit 1995, Wang 2001). The mathematical term "attractor" simply means any self-sustained and stable state of a dynamical system, such as a neural network. According to this picture, in a working memory system, the spontaneous state and stimulus-selective memory states are assumed to represent multiple and co-existing attractors, such that a memory state can be switched on and off by transient inputs.

Fig. 6 illustrates the biophysics of an attractor network. In an object working memory model, subpopulations of neurons are selective to different object stimuli. When the strength of excitatory connections between neurons within each subpopulation is larger than a threshold level, persistent activity appears as an all-or-none phenomenon (Fig. 6A-B). Below the critical threshold, only the spontaneous state exists. Above the threshold, the spontaneous activity state is still dynamically stable to small perturbations, because at low firing rates excitation is effectively counteracted by feedback inhibition (Fig. 6C). However, if a stimulus generates a transient high activity in a neural subpopulation, now recurrent reverberation is sufficiently powerful to drive this group of cells to "escape" from the spontaneous state. A higher firing activity leads to an even larger recurrent synaptic excitation, which becomes sufficient to sustain a persistent active state after the stimulus is withdrawn. The firing rate is stabilized by negative feedback (Fig. 6C). As a result, a stable attractor of persistent activity with an elevated firing rate is realized, that coexists with the stable spontaneous state (Fig. 6C). Biophysical mechanisms that control the firing rates in a working memory network remain to be identified. Among possible contributors are outward ion currents in single cells, feedback inhibition, short-term synaptic depression, and saturation of the synaptic drive at high frequencies.

Biophysically-realistic models shed insights into the circuit properties required for the generation of stimulus-selective persistent activity. In particular, it was found that a network with strong recurrent loops is prone to instability, if excitation (positive feedback) is fast compared to negative feedback, as is expected for a nonlinear dynamical system in general. This is the case when excitation is mediated

by the AMPA receptors, which is typically about 2-3 times faster than inhibition mediated by $GABA_A$ receptors (time constant 5-10 ms). As we have seen earlier, the interplay between AMPA and $GABA_A$ receptors in a excitatory-inhibitory loop naturally gives rise to fast network oscillations. In a working memory model, the large amount of recurrent connections, needed for the generation persistent activity, often lead to excessive oscillations that are detrimental to network stability. Working memory function can be rendered robust and stable, if excitatory reverberation is slow, i.e. contributed by the NMDA receptors (time constant 50-100 ms) at recurrent synapses. Thus, the model predicts a critical role of NMDA receptors in working memory (Wang 2001).

On the other hand, it is essential that exuberant excitation be tightly balanced by inhibition (Brunel and Wang 2001). Synaptic inhibition plays a major role in stimulus selectivity of persistent activity. The model predicts quantitative features of GABAergic inhibitory cells, which have been supported by measurements of putative inhibitory neurons from behaving monkeys, and by the observation that $GABA_A$ receptor antagonists resulted in the loss of spatial tuning of prefrontal neurons during a delayed oculomotor task. Further, a key aspect of memory maintenance is resistance against distractors: while behaviorally relevant information is actively held in mind, irrelevant sensory stimuli should be denied entrance to the working memory system. We found that synaptic inhibition enables our model network to resist distracting stimuli during working memory, and that the network's ability to ignore distractors can be enhanced by dopamine modulation of recurrent excitation and inhibition (Brunel and Wang 2001).

Perceptual decision making

Cortical areas that are engaged in working memory – like the prefrontal and parietal cortices – also play a critical role in other cognitive functions such as decision making, selective attention, behavioral control. This suggests that microcircuit organization in these areas is endowed with the necessary properties to subserve both *internal representation* of information and *dynamical computations* of cognitive types. As it turns out, models originally developed for working memory can

account for decision making processes as well (Wang 2002). An example is shown in Fig. 7 from model simulations of a visual motion discrimination experiment (Shadlen and Newsome 2001, Schall 2001). In this two-alternative forced choice task, monkeys are trained to make a judgment about the direction of motion (say, left or right) in a stochastic random dot display, and to report the perceived direction with a saccadic eye movement. A percentage of dots (called motion strength) move coherently in the same direction, so the task can be made easy or difficult by varying the motion strength (close to 100% or 0%) from trial to trial. While a monkey is performing the task, single-unit recordings revealed that neurons in the posterior parietal cortex and prefrontal cortex exhibit firing activity correlated with the animal's perceptual choice (Shadlen and Newsome 2001). For example, in a trial the motion strength is low (say 3.2%), if the stimulus direction is left whereas the monkey's choice is right, cells selective for right display a higher activity than those selective for left. This experiment can be simulated using the same model designed for working memory (Brunel and Wang 2001). The only difference between a working memory simulation and a decision simulation is that, while for a delayed response task only one stimulus is presented, for a perceptual discrimination task conflicting sensory inputs are fed into competing neural subpopulations in the circuit. This is schematically depicted in Fig. 7A, where the relative difference in the inputs $(I_A - I_B)/(I_A + I_B)$ mimicks the motion strength in the visual motion discrimination experiment. Fig. 7B shows a simulation with zero coherence ($I_A = I_B$). At the stimulus onset, the firing rates of the two competing neural populations r_A and r_B initially ramp up together for hundreds of milliseconds, before diverging from each other when one (r_A in this case) increases while the other (r_B) declines. The perceptual choice is decided based on which of the two neural populations wins the competition. Therefore, consistent with the physiological observations from the monkey experiment (Shadlen and Newsome 2001), decision process proceeds in two steps. Sensory data are first integrated over time in a graded fashion, which in the model is instantiated by the NMDA receptor dependent slow reverberation. This is followed by winner-take-all competition produced by synaptic inhibition, leading to a categorical (binary) choice).

Note that in the case of zero coherence, the input is identical to both neural populations, hence which of the two wins is determined by noise, due to irregular spiking dynamics of neurons in the decision circuit. From one trial to the next, the network's choice is probabilistic at chance level (Fig. 7C). When the motion strength is not zero, $I_A > I_B$, neuronal ramping activity is faster, with a slope increasing with the motion strength (Wang 2002). In other words, the network integrates sensory information at a higher accumulation rate, and the decision is reached more rapidly, when the evidence about a stimulus is stronger. The network's probability of correct choice as a function of the motion strength (Fig. 7D) is similar to the animal's behavioral performance (Shadlen and Newsome 2001). Hence, the model is able to account for both neurophysiological and psychophysical observations of the monkey experiment. Once again, the essential network requirement is slow reverberation (for time integration) balanced by synaptic inhibition (for winner-take-all competition). In the model slow excitation depends on the synaptic NMDA receptors. Conceivably, though, slow voltage- or/and calcium-gated inward ion channels in single neurons could also contribute to such reverberatory circuit dynamics.

Time integration

Qualitatively speaking, working memory requires neurons to convert a transient input pulse into a sustained persistent activity, like a time integral of the stimulus. Similarly, in perceptual decisions, approximate linear ramping activity, at a rate proportional to input strength, can also be conceptualized as time integration. It is worth noting, however, a genuine integrator implies that, after a transient input is turned off, the activity is persistent at a level that varies with the input amplitude in a graded manner. This is not the case in Fig. 7, where after the stimulus offset the neural activity is binary (representing one of the two categorical choices), independent of the input motion strength. Indeed, this is what has been observed in posterior parietal neurons, and is the kind of neural signals required by the behavioral task (Shadlen and Newsome 2001). Using other task paradigms and in a number of neural systems, graded (or "parametric" persistent activity has been observed, and found to store short-term memory of an analog quantity like spatial

location of a stimulus (Goldman-Rakic 1995), frequency of a somatosensory vibrational input (Romo et al 1999), allocentric head direction of an animal (Taube and Bassett 2003), and gaze position during eye fixation (Seung et al. 2000).

There are two known coding schemes by neural integrators. In a "rate code", persistent firing rate of each neuron varies linearly with the encoded feature. As a result, if rates of different neurons are plotted against each other, they fall on a straight line in the "firing-rate space". This observation led to the theoretical concept of "line attractors" (Seung et al 2000). By contrast, according to a "location code" neurons exhibit a bell-shaped Gaussian tuning curve of an encoded feature, therefore distinct neural groups are engaged in storing different values of the analog feature. In other words, the stimulus feature is specified by "how much active are neurons" in a rate code, and by "which neurons are active" in a location code. Mathematically, a perfect integrator is described by $dX/dt = I(t)$, where $I(t)$ is the input, X is either neural firing activity level (in a rate code) or the peak locus of network activity (in a location code).

Fig. 8 shows an neural integrator model in which the encoded feature is specified by the peak location of the network activity profile. This model was proposed for head direction (HD) cells (Song and Wang 2005). HD neurons are part of the spatial navigation system and signal the animal's directional heading. When an animal turns its head, the angular velocity signal carried by vestibular inputs is integrated over time by HD cells into a positional signal, and the latter is sustained internally when the animal keeps the head direction fixed. HD cells are selective for angular head direction according to a Gaussian (bell-shaped) tuning curve. Interestingly, available evidence indicates that HD cells are generated in a neural circuit characterized by a paucity of local excitatory connections (Taube and Bassett 2003). Consistent with this observation, the model of Fig. 8 proposes a cross-inhibition mechanism, without recurrent excitation, that gives rise to direction-selective "hill of persistent activity" Moreover, the model surmises a shift mechanism by "rotation cells" consisting of two inhibitory cell populations (Fig. 8A). When the head direction is fixed, the inputs from the two inhibitory neural populations are balanced with each other, and the hill of activity of a third, excitatory, neural population is

maintained fixed (Fig. 8B, time epochs between input pulses). As the head turns, the angular velocity signal increases the firing of one inhibitory population, while decreases the firing of the other. The resulting asymmetric inhibitory inputs induce the activity pattern in the excitatory neural population to move, at a constant speed proportional to the input amplitude (angular velocity) (Fig. 8B, time epochs during input pulses). Computation performed by the model is quite close to an integral operation in the sense of Calculus. When the input intensity is doubled (second versus first input pulse), the hill of activity moves twice as fast. Moreover, the network can integrate both positive inputs (first and second pulses) and negative inputs (third pulse). In fact, with a third input half in amplitude but lasting twice as long as the second input, the hill of activity moves back to the position before the second pulse, as expected by perfect integration.

This model assumes certain network symmetry, whose biological basis remains unclear. Also, models of perfect line attractors require fine-tuning of network parameters (Seung 2000, Wang 2001). To see why this is the case, consider the following simple firing-rate equation

$$\frac{dr}{dt} = \frac{(-r + w_{rec} r)}{\tau} + I(t)$$

where r is a firing rate, τ is a typical biophysical (membrane or synaptic) time constant, and w_{rec} is the strength of recurrent connections. The effective time constant of the system by given by $\tau_{eff} = \tau_{syn}/(1 - w_{syn})$, which is longer than τ in the presence of w_{rec}. For instance, if $\tau = 100$ ms and $1 - w_{rec} = 0.05$, then $\tau_{eff} = 2$ sec. When $w_{rec} = 1$ (which requires fine-tuning of the parameter w_{rec}), $\tau_{eff} = \infty$, and the system becomes a perfect integrator. How neural integrators can be realized robustly by plausible biophysical mechanisms remain a topic of active current research. Also, as we have seen, integrators can rely either on synaptic excitation, or cross-inhibition, or intrinsic cellular mechanisms. Further progress depends on new experiments to put more constraints, and establish a firmer mechanistic basis, for computational modeling in this area.

In summary, it is worth contrasting neural integration with adaptation discussed at the beginning of this chapter: early sensory processing appears to require "time

derivative" type neural computation, whereas cognitive functions involve "time integral" type neural circuits.

Concluding remarks

Half century after Hodgkin and Huxley pioneered a quantitative description of the electrical activity of nervous membrane, Computational Neuroscience has grown into a mature field. As illustrated by various examples here, theory and models, in close interaction with experimentation, can be fruitfully applied to problems of different types (sensory processing versus cognitive functions) and at different levels (single synapse/cell versus networks). I have emphasize microcircuit modeling, at the interface between network behavior and underlying cellular mechanisms. Among open questions for future research is to delineate the contributions of detailed single neuron dynamics (like active dendritic properties) to network function. Another intriguing problem is to better quantitatively characterize the statistics of neural microcircuits as complex networks. A third direction is to study adaptive brain functions at many spatial and temporal scales. For instance, learning and memory formation span timescales from milliseconds to many years, and involve processes at the molecular (intracellular protein signaling network), neuronal (concerted firing pattern) and systems levels. Finally, biologically constrained modeling will need to be expanded to large neural networks encompassing multiple brain areas, for which a solid theoretical foundation is still lacking.

Acknowledgements. This work was supported by NIMH grant 062349.

Reading list

(1) Chow CC, Gutkin B, Hansel D, Meunier C and Dalibard J (2005) Methods and Models in Neurophysics. Amsterdam: Elsevier.

(2) Dayan P and Abbott LF (2001) Theoretical Neuroscience, Cambridge, MA: MIT Press.

(3) Feng J (Ed) (2004) Computational Neuroscience: a Comprehensive Approach. Boca Raton: Chapman and Hall.

(4) Koch C (1999) Biophysics of Computation: Information Processing in Single Neurons. New York: Oxford University Press.

(5) Koch C and Segev I (Eds) (1998) Methods in Neuronal Modeling: From Ions to Networks, 2nd edition. Cambridge, MA: MIT Press.

(6) Rao RPN, Olshausen BA and Lewicki MS (2002) Probabilistic Models of the Brain. Cambridge, MA: MIT Press.

(7) Rieke F, Warland D, de Ruyter van Steveninck R and Bialek W (1999) Spikes: Exploring the Neural Code. Cambridge, MA: MIT Press.

(8) Sejnowski TJ, Koch C, Churchland PS (1988) Computational neuroscience. Science 241: 1299-1306

Further references

(1) Abbott LF, Varela JA, Sen K, Nelson SB (1997) Synaptic depression and cortical gain control. *Science* 275: 220-224.

(2) Amit DJ (1995) The hebbian paradigm reintegrated: local reverberations as internal representation. *Behav. Brain Sci.* 18: 617-626.

(3) Brunel N and Wang X-J (2001) Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. *J Comput Neurosci* 11: 63-85.

(4) Goldman-Rakic P (1995) Cellular basis of working memory. *Neuron* 14: 477-485.

(5) Hansel D, Mato G and Meunier C (1995) Synchrony in excitatory neural networks. *Neural Comput* 7: 307-337.

(6) Romo R, Brody CD, Hernandez A and Lemus L (1999) Neuronal correlates of parametric working memory in the prefrontal cortex. *Nature* 399: 470-473.

(7) Marder E, Calabrese RL (1996) Principles of rhythmic motor pattern generation. *Physiol Rev* 76: 687-717.

(8) Schall JD (2001) Neural basis of deciding, choosing and acting. *Nature Neurosci* 2: 33-42.

(9) Seung HS, Lee DD, Reis BY and Tank D (2000) Stability of the memory of eye position in a recurrent network of conductance-based model neurons. *Neuron* 26: 259-271.

(10) Shadlen MN and Newsome WT (2001) Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J Neurophysiol* 86: 1916-1936.

(11) Sompolinsky H and Shapley R (1997) New perspectives on the mechanisms for orientation selectivity. *Curr Opin Neurobiol* 7: 514-522.

(12) Song P and Wang X-J (2005) Angular path integration by moving 'hill of activity': a spiking neuron model without recurrent excitation of the head-direction system. *J Neurosci* 25: 1002-1014.

(13) Taube JS and Bassett JP (2003) Persistent neural activity in head direction cells. *Cereb. Cortex* 13: 1162-1172.

(14) Tsodyks MV, Markram H (1997) The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc Natl Acad Sci USA* 94: 719-723.

(15) van Vreeswijk C, Abbott LF, Ermentrout GB (1994) When inhibition, not excitation synchronizes neural firing. *J Comput Neurosci* 1: 313-322.

(16) Wang X-J (2001) Synaptic reverberation underlying mnemonic persistent activity. *Trends in Neurosci* 24: 455-463.

(17) Wang X-J (2002) Probabilistic decision making by slow reverberation in neocortical circuits. *Neuron* 36: 955-968.

(18) Wang X-J and Rinzel J (1992) Alternating and synchronous rhythms in reciprocally inhibitory model neurons. *Neural Computat* 4: 84-97.

(19) Wang X-J, Liu YH, Sanchez-Vives MV and McCormick DA (2003) Adaptation and temporal decorrelation by single neurons in the primary visual cortex. *J Neurophysiol* 89: 2707-2725.

(20) Whittington MA, Traub RD and Jefferys JG (1995) Synchronized oscillations in interneuron networks driven by metabotropic glutamate receptor activation. *Nature* 373: 612-615.

Figure Captions

Figure 1: (A) Short-term plasticity of excitatory intracortical synapses. Left panel: depression of an excitatory synapse between two layer 5 pyramidal cells recorded in a slice of rat somatosensory cortex. Spikes were evoked by current injection into the presynaptic neuron and the postsynaptic potential was recorded with a second electrode (from H Markram and M Tsodyks *Nature* 382: 807-810 (1996)). Right panel: facilitation of an excitatory synapse from a pyramidal cell to an inhibitory interneuron in layer 2/3 of rat somatosensory cortex (from H Markram et al. *Proc. Natl. Acad. Sci. (USA)* 95: 5323-5328 (1998)). (B) The average rate of synaptic transmission for a synapse model with short-term depression when the presynaptic firing rate changes in a sequence of steps. The parameters of the model are $P_0 = 1$, $f_D = 0.4$ and $\tau_D = 500$ ms. This figure is adapted from Dayan and Abbott (2001).

Figure 2: Adaptation and decorrelation by a single neuron. (A) Response of the model neuron to a low-high-low sinusoidal current input. Top to bottom: Membrane potential, number of spikes per cycle, intracellular Na^+ concentration, and 2 Hz sinusoidal injected current that varies from low to high to low in amplitude. Following the injection of high-amplitude sinusoidal current there is a slow hyperpolarization lasting for about ten seconds (indicated by an arrow). Intracellular Na^+ accumulation mirrors the adaptation time course of the instantaneous firing rate during the high-contrast input; whereas Na^+ decay is correlated with the slow hyperpolarization and recovery during the second low-amplitude period. (B) Decorrelation of $1/f$-type stochastic input by the model (Left panel) and a visual cortical neuron recorded intracellularly in a ferret cortical slice (Right panel). See text for detailed discussion. This figure is adapted from Wang et al. (2003).

Figure 3. Intrinsic membrane oscillations of single neurons. In each panel are shown experimental data and model simulations. (A) A thalamic relay cell displays two distinct spiking modes: tonic firing upon depolarization, and burst discharges upon hyperpolarization. (Upper trace adapted from DA McCormick and HC Pape *J. Physiol. Lond.* 431:319-342 (1990); lower trace adapted from X-J Wang *Neu-*

rosci 59:21-31 (1994)) (B) A chattering neuron from the cat visual cortex shows rhythmic bursting in the gamma frequency range. (Upper trace from CM Gray and DA McCormick *Science* 274:109-113 (1996); lower trace from X-J Wang *Neurosci* 89:347-362 (1999)) (C) A non-cholinergic (putative GABAergic) cell in the rat medial septum displays rhythmic alternations at theta frequency between clusters of spikes and epochs of subthreshold membrane potential oscillations. (Upper trace from M Serafin et al. *Neurosci* 75:671-675 (1996), lower trace: from X-J Wang, *J Neurophysiol* 87: 889-900 (2002)). The simulated oscillation is faster than the experimental data (see the different time scales), because the model simulation was done at body temperature (37^{o}C), whereas the *in vitro* trace was recorded at 32^{o}C. This figure is adapted from X-J Wang *Encyclopedia of Cognitive Science*, MacMillan Reference Ltd, pp. 272-280 (2003).

Figure 4. Interneuronal network model for synchronization of coupled oscillators. (A) An example of network synchronization in a fully connected regular network. Upper panel: the rastergram where each row of vertical bars represents spikes discharged by one of the neurons in the network. Lower panel: membrane potentials of two neurons. Neurons initially fire asynchronized, but quickly become perfectly synchronized by mutual inhibition. (B) In a random network, the network coherence is plotted versus the mean number of synaptic inputs per cell M_{syn} (The correction term ($\sim 1/N$) takes into account the finite size effect). Different curves correspond to different network size (N=100, 200, 500, 1000). There is a critical threshold for the connectedness above which network synchrony occurs. This threshold connectivity is independent of the network size. Figures are adapted from Wang and Buzsáki, *J Neurosci* 16: 6402-6413 (1996) and Buzsáki G, Geisler C, Hinze D and Wang X-J *Trends in Neurosci* 27: 186-193 (2003).

Figure 5. Feedback inhibition model for synchronous oscillations with irregularly firing neurons. Computer simulation of a model with two neural populations (pyramidal cells and interneurons) in a sparsely connected random network. The network shows a collective oscillation at 55 Hz (see population rates, and the power spectrum), whereas Single neurons fire spikes intermittently in time at low rates (2 Hz

for pyramidal cells, 10 Hz for interneurons; see rastergrams). The spike discharges are synchronized to zero-phase between the two populations (second trace on the left); whereas the inhibitory synaptic current shows a phase lag of about 2 ms compared to the excitatory synaptic current (third and bottom traces on the left). (N Brunel and X-J Wang *J Neurophysiol* 94: 4344-4361 (2003)).

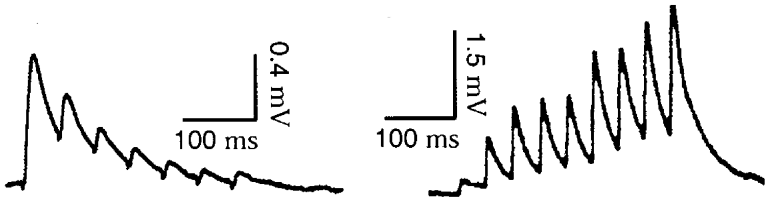
Figure 6: Attractor paradigm for persistent activity. (A) Delayed match-to-sample simulation of an object working memory model. Neural subpopulations (labeled 1 to 5) are selective to different stimuli. Average activity from each neural group is plotted below the rastergram. Inhibitory (I) cell population is shown in black. In the simulation, the shown stimulus triggers persistent activity in a pyramidal cell group (red) at about 30 Hz, but not in other pyramidal groups. Delay period activity is switched off by a transient excitatory input generating a brief surge of activity in all neurons. (B) Dependence of network activity on the strength of recurrent synaptic connectivity. Solid lines: spontaneous state and persistent memory state; dashed line: unstable states. There is a critical threshold of synaptic strength, above which persistent activity appears abruptly as an all-or-none phenomenon. The example in (A) corresponds to the parameter indicated by the star sign. (C) Schematic illustration of the biophysics underlying an attractor dynamics. An attractor is a neural firing state that is stable to perturbations: when a small input perturbs the network to a lower or higher activity level, there is a restoring force to bring the network back to the attractor state. In this case, the spontaneous state is stabilized from below by background inputs, and from above by feedback synaptic inhibition. A sufficiently powerful sensory stimulus can drive a cell assembly to escape from the spontaneous state, and after the stimulus is withdrawn the system settles in one of the active memory states at an elevated firing rate. The persistent activity state is stabilized from below by excitatory reverberation, and from above by various negative feedback rate control mechanisms. Finally, a behavioral response or reward signal can turn the network off and erase the memory. Adapted from Brunel and Wang (2001) and Wang (2001).

Figure 7: A simple model for two-alternative forced-choice tasks. (A) Model

scheme. There are two pyramidal cell groups, each of which is selective to one of the two directions (A=left, B=right) of random moving dots in a visual motion discrimination experiment. Within each pyramidal neural group there is strong recurrent excitatory connections which can sustain persistent activity triggered by a transient preferred stimulus. The two neural groups compete through feedback inhibition from interneurons. (B) A network simulation with zero coherence. Population firing rates r_A and r_B exhibit an initial slow ramping (time integration) followed by eventual divergence (categorical choice). (C) Decision dynamics shown in the two-dimensional plane where firing rates r_A and r_B are plotted against each other. Traces with different colors are from different trials, demonstrating probabilistic decision behavior. With $I_A = I_B$, the choice is at chance level across trials. (D) Decision performance (percentage of correct choices) as a function of the relative difference in the inputs. Stimulus is shown for a fixed duration of 1 second. Adapted from Wang (2002).

Figure 8: Time integration by a "hill of persistent activity". (A) Model scheme. An excitatory neural network, encoding a directional angle (0 to 360 degrees), receives inputs from two inhibitory neural populations that are balanced with each other when there is no input. A velocity signal increases input to one of the inhibitory population ($+I$) and decreases input to the other ($-I$), leading to a bias for the excitatory network. (B) Network firing pattern (Top) in response to a series of input signal steps (Bottom). Neurons in the network are aligned along the y-axis, labelled by their preferred directional angles. X-axis is time. Each dot is a spike. The network activity pattern has the form of a bell-shaped profile (Upper panel, right), its peak location encodes the directional angle (white line in the rastergram). In the absence of an external input, the directional information is maintained by persistent firing pattern. An input induces the hill of activity to move at a speed proportional to the stimulus amplitude, hence the network performs a time integral of the input. See text for more detailed discussion. Simulation by P. Song using the model published by Song and Wang (2005).

A


0.4 mV
100 ms
1.5 mV
100 ms

B

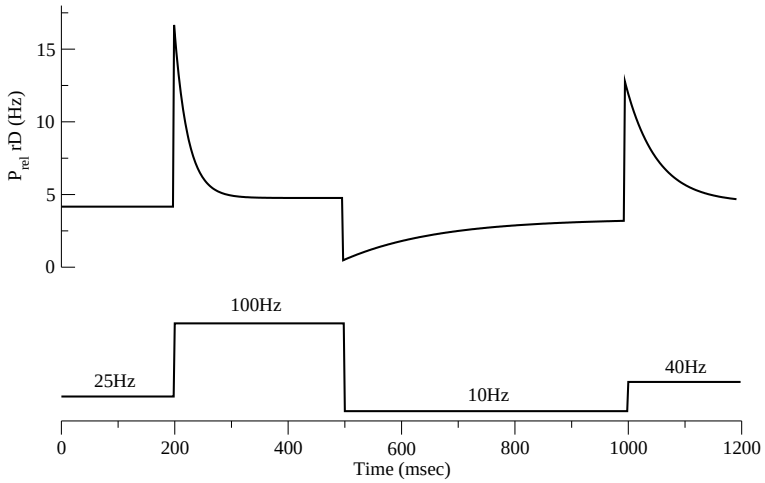

15
10
5
0
P rel rD (Hz)
100Hz
25Hz
10Hz
40Hz
0
200
400
600
800
1000
1200
Time (msec)

Figure 1

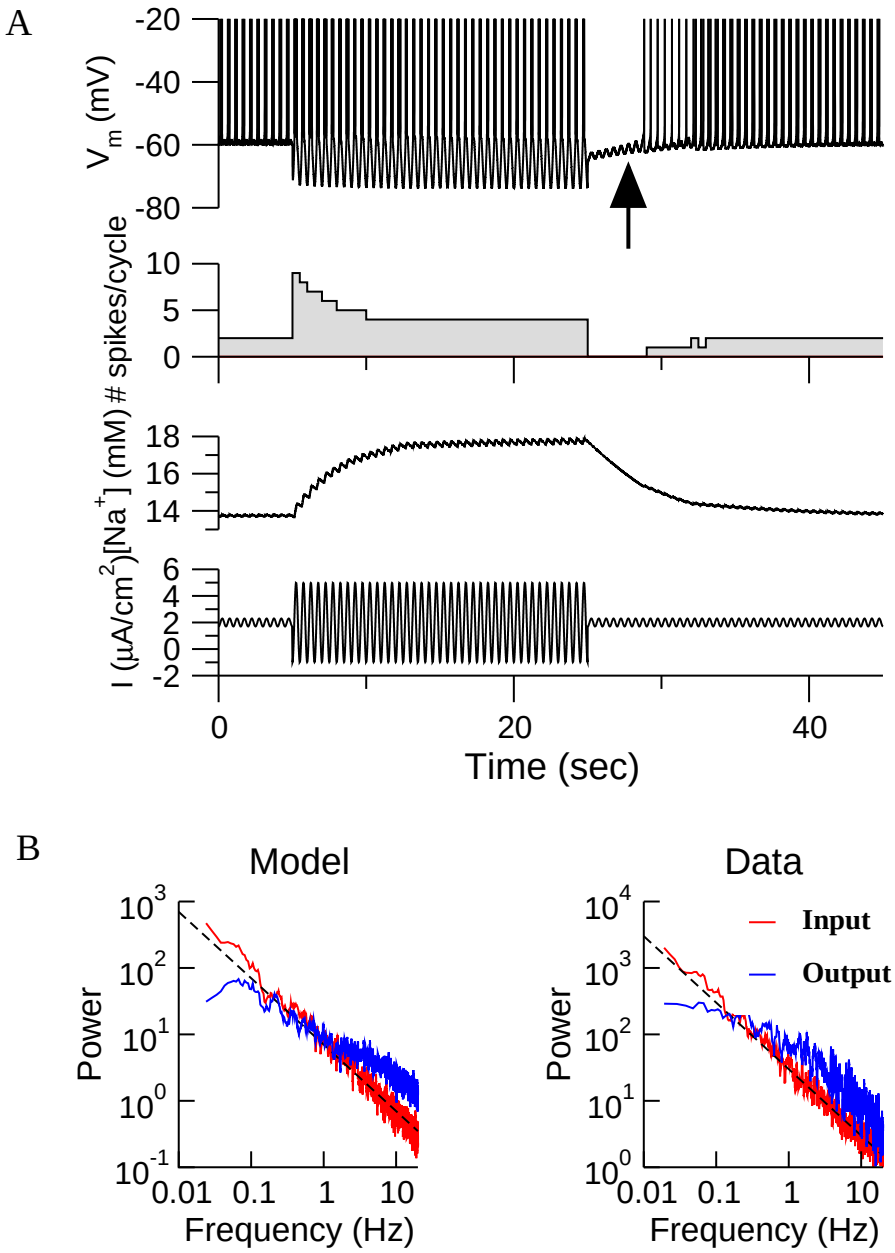

A
-20
-40
-60
-80
V_m (mV)
10
5
0
spikes/cycle
18
16
14
$[Na^+]$ (mM)
6
4
2
0
-2
I ($\mu A/cm^2$)
0
20
40
Time (sec)
B
Model
Data
Input
Output
Power
10^3
10^2
10^1
10^0
10^{-1}
10^4
0.01
0.1
1
10
Frequency (Hz)

Figure 2

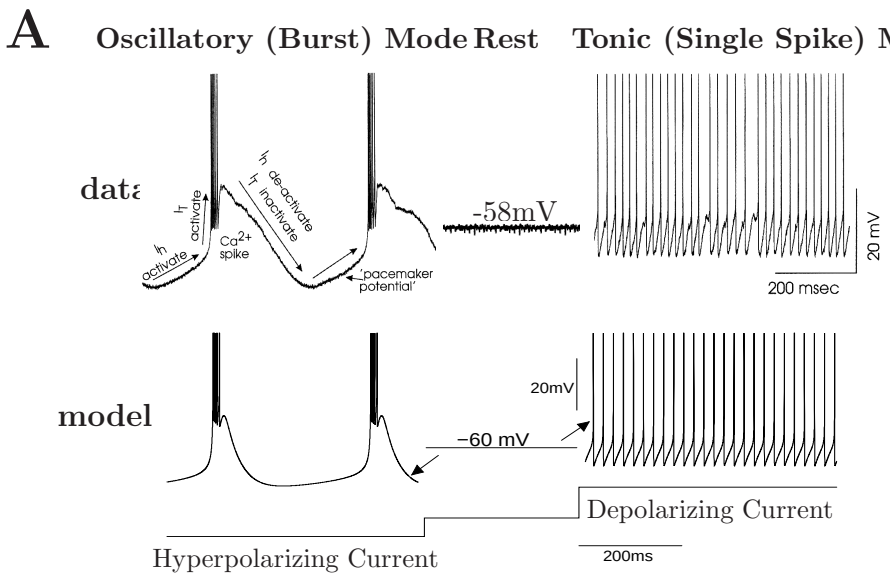
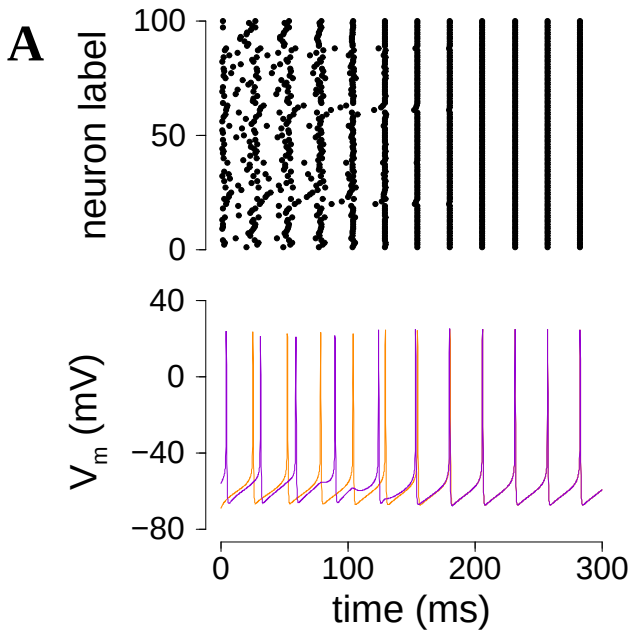

A
Oscillatory (Burst) Mode
Rest
Tonic (Single Spike) Mode
data
model
-58mV
20 mV
200 msec
20mV
−60 mV
Depolarizing Current
Hyperpolarizing Current
200ms
B
data
model
20ms
20mV
C
data
model
20 mV
500 ms
20mV
200ms

Figure 3


A
neuron label
100
50
0
40
0
−40
−80
V_m (mV)
0
100
200
300
time (ms)

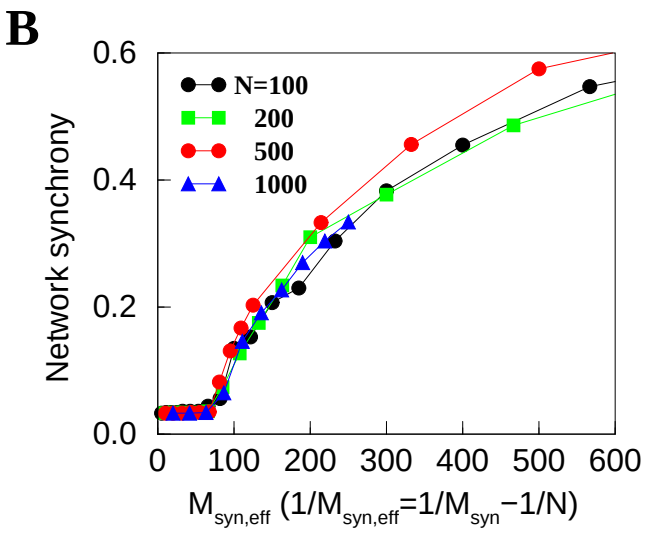

B
Network synchrony
0.6
0.4
0.2
0.0
N=100
200
500
1000
0
100
200
300
400
500
600
M_syn,eff (1/M_syn,eff=1/M_syn−1/N)

Figure 4

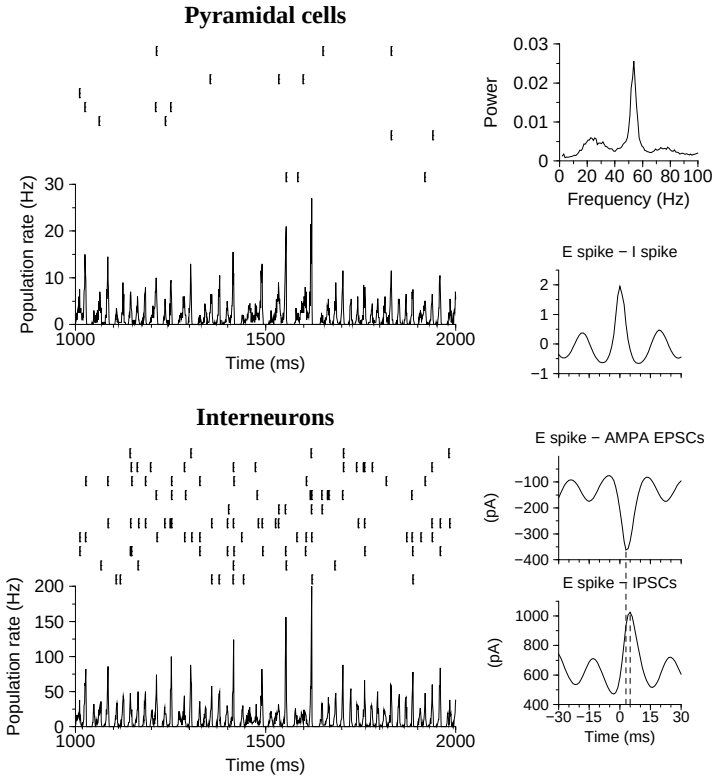

Pyramidal cells
Population rate (Hz)
Time (ms)
Power
Frequency (Hz)
E spike − I spike
Interneurons
E spike − AMPA EPSCs
(pA)
E spike − IPSCs
(pA)
Time (ms)

Figure 5

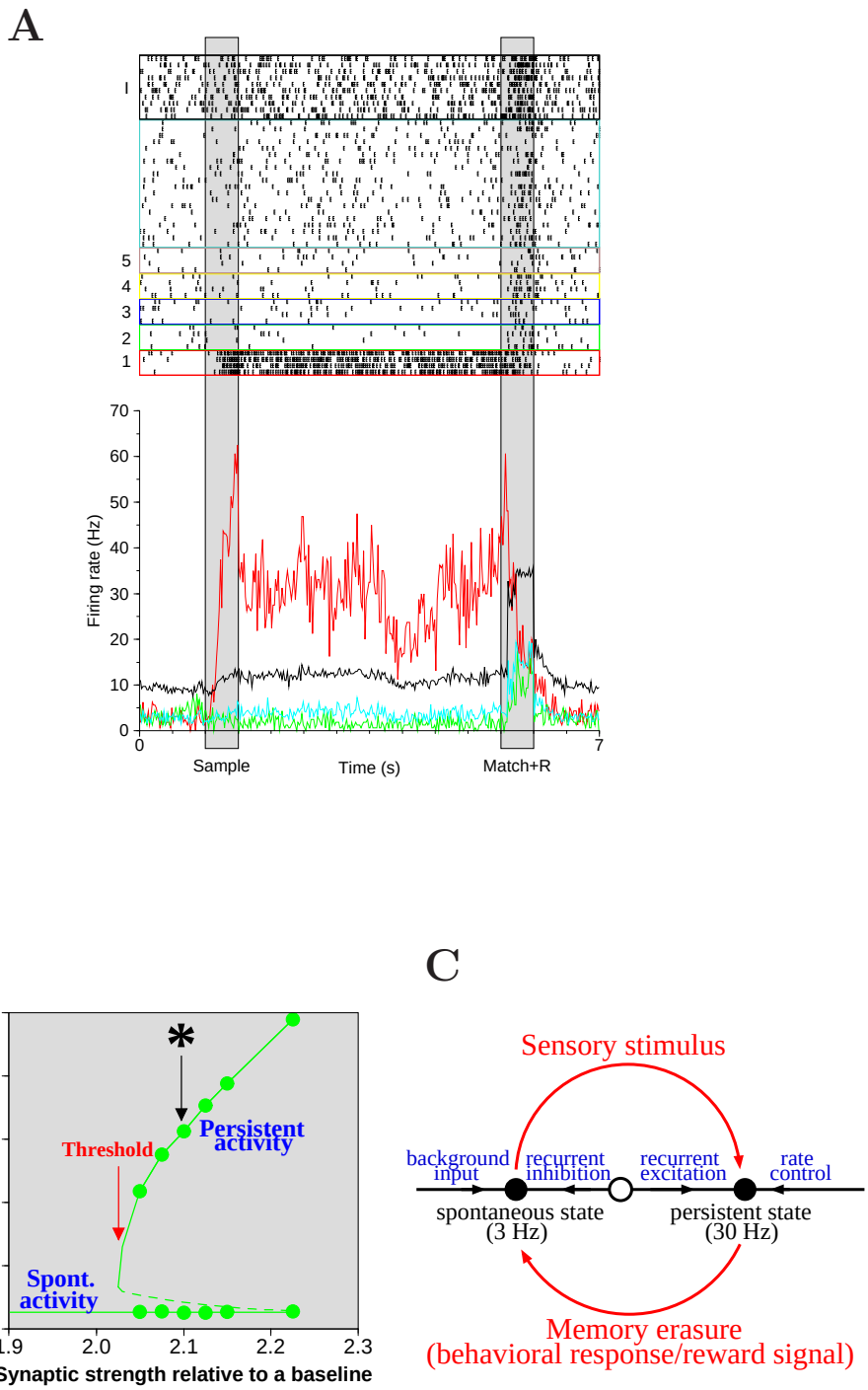

A
I
5
4
3
2
1
70
60
50
40
30
20
10
0
Firing rate (Hz)
0
7
Sample
Time (s)
Match+R
B
50
40
30
20
10
0
Firing rate (Hz)
1.9
2.0
2.1
2.2
2.3
Synaptic strength relative to a baseline
*
Threshold
Persistent activity
Spont. activity
C
Sensory stimulus
background input
recurrent inhibition
recurrent excitation
rate control
spontaneous state (3 Hz)
persistent state (30 Hz)
Memory erasure
(behavioral response/reward signal)

Figure 6

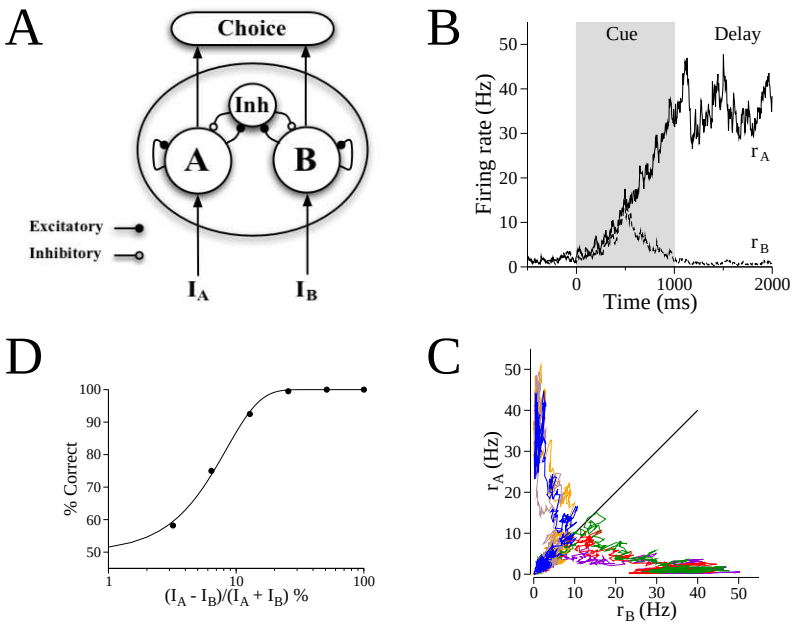
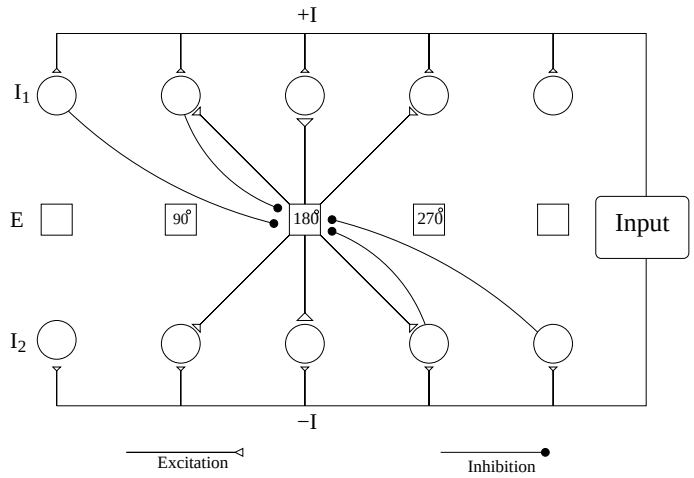

A
Choice
Inh
A
B
Excitatory
Inhibitory
I_A
I_B
B
Cue
Delay
Firing rate (Hz)
r_A
r_B
Time (ms)
D
% Correct
$(I_A - I_B)/(I_A + I_B)$ %
C
r_A (Hz)
r_B (Hz)

Figure 7

A


+I
I1
E
90°
180°
270°
Input
I2
−I
Excitation
Inhibition

B

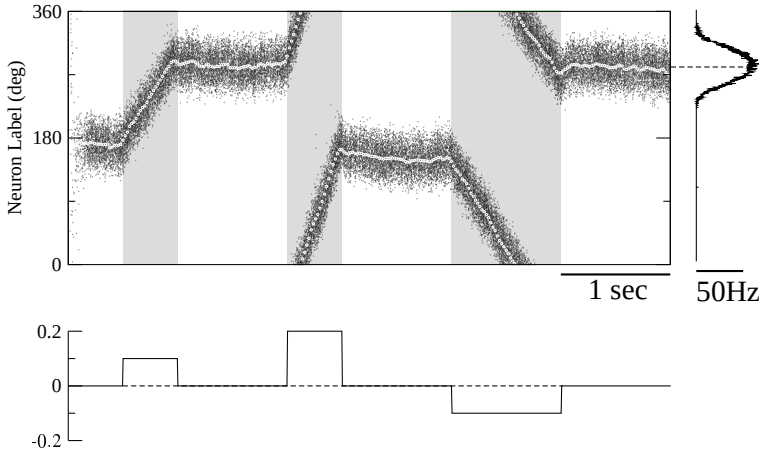

360
180
0
Neuron Label (deg)
1 sec
50Hz
0.2
0
-0.2

Figure 8