Dynamical constraints on using precise spike timing to compute in recurrent cortical networks

Arunava Banerjee¹, Peggy Seriès², and Alexandre Pouget³

¹Computer and Information Science and Engineering, University of Florida, Gainesville, Florida, USA 32611, ²Gatsby Computational Neuroscience Unit, University College London, London, UK WC1N 3AR, and ³Brain and Cognitive Science Department, University of Rochester, Rochester, New York, USA 14627.

Acknowledgment: We would like to thank Peter Latham for helpful discussions. A.B. and A.P. were supported by an ONR grant (N00014-00-1-0642). P.S. was supported by a grant from the James McDonnell foundation.

Several recent models have proposed the use of precise timing of spikes for cortical computation. Such models rely on growing experimental evidence that neurons in the thalamus as well as many primary sensory cortical areas respond to stimuli with remarkable temporal precision. Models of computation based on spike timing, where the output of the network is a function not only of the input but also of an independently initializable internal state of the network, must however satisfy a critical constraint—the dynamics of the network should not be sensitive to initial conditions. We have previously developed an abstract dynamical system for networks of spiking neurons that has allowed us to identify the criterion for the stationary dynamics of a network to be sensitive to initial conditions. Guided by this criterion, we analyzed the dynamics of several recurrent cortical architectures, including one from the orientation selectivity literature. Based on the results, we conclude that under conditions of sustained, Poisson-like, weakly correlated, low to moderate levels of internal activity as found in the cortex, it is unlikely that recurrent cortical networks can robustly generate precise spike trajectories, i.e., spatio-temporal patterns of spikes precise to the millisecond time scale.

Key words: sensitive dependence, dynamical systems, spiking neurons

Introduction

In recent years, there has been growing experimental evidence that neurons in the thalamus as well as many primary sensory cortical areas respond to stimuli with remarkable temporal precision. For example, Reinagel & Reid (2000) have shown that neurons in the cat lateral geniculate nucleus (LGN) respond to randomly modulated visual stimuli in a highly reproducible manner with individual spikes occurring with a precision of better than 1 msec (see also Panzeri et al., 2001, for coding of sensory stimuli in the rat somatosensory cortex). Since these experiments strictly relate external stimuli to the response of neurons, they are however, by design, confined to the analysis of computation that occurs in a functionally

feedforward¹ manner, i.e., where the output of the system under consideration is mostly a function of the present and past input. Stated formally, the model underlying such systems is assumed to be $O(t)=\mathfrak{F}(I(t))$, where O(t) and I(t) represent, respectively, the time varying output and input of the system and $\mathfrak{F}(\cdot)$ represents the functional that maps the input to the output. It is now well recognized that for behaviourally relevant inputs, $\mathfrak{F}(\cdot)$ can be highly precise (Rieke et al., 1997). Simulation results within this framework, such as that in (Diesmann et al., 1999), have also shown that for particular classes of inputs and outputs it is possible to maintain spike timing in feedforward networks. In particular, it was demonstrated that a synchronous volley comprising approximately 100 spikes can be propagated robustly across several layers of a feedforward network of integrate-and-fire neurons in the presence of approximately 2 Hz of random background activity and to a small amount of jitter in the spike timings.

In contrast, experiments that assess working memory, such as delayed response and delayed match-to-sample tasks, have hitherto shown that the spike trains generated by the constituent neurons of the system, under identical stimulus presentations, are highly variable (Compte et al., 2003), although there is evidence of occasional coincidental spikes between neurons. In themselves, these results do not preclude the possibility of the cortical structures in question being capable of robustly generating precise spike trajectories, since the output in these cases is a function not only of the input but also of the internal state of the network. With little or no experimental control over the internal states of such systems, it is hardly surprising that variable spike trains are generated under identical input presentations. The significant impact of the rich internal dynamics of these systems is also only recently coming to light, with evidence showing that neurons as far as the prefrontal cortex can on average phase lock to the theta oscillations generated by the hippocampus (Siapas et al., 2005). In essence, the form of computation explored by these experiments is

_

¹ We label such systems functionally feedforward to highlight the fact that they are not constrained to be architecturally feedforward. Such systems are distinguished by the property that their output is determined solely by the input, with the internal state of the network not playing an independent role in the mapping.

of a more general nature where the output of the system is a function not only of the input but also of an independently initializable internal state of the network².

Such systems can not be modelled under the framework $O(t) = \mathcal{F}(I(t))$ as the following simple example demonstrates. Consider a recurrent network that is capable of sustained reverberatory activity. The system, under the trivial input drive of I(t)=0 for t=0 to $-\infty$, can either be in the quiescent state or in the sustained reverberatory state. Since the output of the system is not uniquely determined by the input, the system can not be modelled using $O(t)=\mathcal{F}(I(t))$. Instead, it has to be modelled as $O(t)=\mathcal{F}(I(t),s(t_0))$, where $s(t_0)$ denotes the internal state of the system at any particular time t_0 . Whether such systems in the cortex can perform computation using the precise timing of spikes is an issue that remains to be resolved. It is this question that we address in this article. To elaborate, we investigate whether input driven recurrent cortical networks can generate spike trajectories precise to the millisecond time scale. We must emphasize here that this is quite distinct from the issue of patterns of spiking at coarser (on the order of 25+ msec) time scales that are induced primarily by network oscillations, for which there is considerable experimental evidence (Nadasdy et al., 1999; Hahnloser et al. 2002).

Experimental determination in the case of the dynamics at the finer time scale is substantially more challenging because it is practically impossible to ascertain and/or control the internal state of cortical networks in vivo. Without knowledge of the internal state of a network, the variability in the output resulting from the variability in the internal state of the network can be confounded with noise.

With recent proposals for neural computation that rely on the timing of spikes to represent and manipulate information (Maass, 1997; Maass et al., 2003), the question of whether or not recurrent cortical networks can reliably compute using the precise timing of spikes has taken on added importance. If information is to be coded in the precise timing of

² In the case of digital computers, this would correspond to the distinction between combinatorial circuits where the output is a function solely of the input, and sequential circuits / finite-state machines where the output is a function both of the input and the internal state of the circuit.

spikes of the internal neurons of a network, the timing of such spikes must be robust to perturbations, i.e., small perturbations in the spike timing of the internal neurons induced by thermal or other noise must not result in successively larger fluctuations in the timing of subsequent spikes generated in the network. In the parlance of dynamical systems theory, the dynamics of cortical networks should not be sensitive to initial conditions.

The issue of the stability of spike timing in recurrent networks of excitatory and inhibitory neurons has been explored by other researchers. Hansel & Sompolinsky (1996), for example, have suggested that spike timing in such networks might be unstable. However their conclusions were based only on simulation results from individual networks. In contrast, van Vreeswijk & Sompolinsky (1998) have derived formal results consistent with the above conclusion, but their analysis was based on a highly simplified model of the cortical circuit, namely, a network of randomly connected binary stochastic units receiving uncorrelated external inputs.

In this article, we investigate through simulations and formal analysis whether or not spike trajectories emanating from input driven recurrent cortical networks, with spatiotemporal characteristics similar to those found in the cortices of awake behaving animals (i.e., weakly correlated, Poisson-like, with spike rates in the 0.1-50 Hz range), can be generated robustly to within the precision set by the noise intrinsic to the neurons in the system (widely estimated to be in the 1-5 msec range (Mainen & Sejnowski, 1995; Novak et al., 1997)). We have previously developed an abstract dynamical system for recurrent networks of spiking neurons that supports a general model of the spiking neuron and is amenable to formal analysis (Banerjee, 2001a). The dynamical system has allowed us to identify the formal criterion for the stationary dynamics of a system of spiking neurons to be sensitive to initial conditions (Banerjee, 2001b; Banerjee & Pouget, 2003; Banerjee, 2006). Guided by the criterion, we analyzed the dynamics of several recurrent cortical architectures whose internal activity (as compared to the input) contributed to determining the output of the network to varying degrees. The networks ranged from ones that could sustain internal spike activity even after the input had ceased, to one from the orientation selectivity literature where the network eventually returned to quiescence after cessation of input. The results indicate that under conditions of sustained, Poisson-like, weakly

correlated, low to moderate levels of internal activity as found in the cortex, the dynamics of the networks ought to be almost surely sensitive to initial conditions.

Materials and Methods

Model of the neuron and network. The abstract dynamical system for a network of spiking neurons is described in detail in (Banerjee, 2001a). Briefly, the dynamical system is formulated based on the assumption that the internal state of a network can be specified by enumerating the temporal positions of all spikes generated in the network over a bounded past. This assumption is valid if the neurons are finite precision devices with fading memory. To illustrate, in Figure 1, the present state of the system is specified by the positions of all spikes (solid lines) in the shaded region at t=0, and the state at a future time T is specified by their positions in the shaded region at t=T. Each internal neuron i (as opposed to an input neuron) is assigned a membrane potential function $P_i(\cdot)$ that maps the spike configuration given by the present state to the instantaneous potential at the soma of neuron i. Internal neuron i generates a new spike whenever $P_i(\cdot)$ reaches the threshold from below. The particular instantiation of the $P_i(\cdot)$'s determines both the electrophysiological properties of the neurons as well as their connectivity in the network.

Sensitive dependence on initial conditions. Whether or not a spike trajectory is sensitive to initial conditions can be tested as follows. Consider the network in Figure 1 initialized at the state described by the shaded region at t = 0. We let the dynamics of the network unfold for a time T and record the positions of all spikes (solid lines). We then reset the network to its initial state, perturb the initial set of internal spikes (dotted lines), and let the dynamics of the network unfold for the same time T. The initial set of perturbations will propagate from spike to spike, affecting the timing of subsequent spikes. Stated informally, the dynamics of the network is sensitive if the successive perturbations tend to grow larger with each new spike generated in the network, and is insensitive otherwise.

Formally, let column vectors $\Delta \vec{x}_0$ and $\Delta \vec{x}_T$ denote, respectively, infinitesimal perturbations on the spikes of internal neurons (as opposed to input neurons) at t=0 and t=T. The dimensionality of $\Delta \vec{x}_0$ and $\Delta \vec{x}_T$ is then the number of internal spikes in the

respective state descriptions. Since we are concerned with sensitive dependence with respect to the internal state of the network, spikes emitted by the input neurons are not perturbed, and therefore we pad $\Delta \vec{x}_0$ and $\Delta \vec{x}_T$ with as many zeroes as there are input spikes in the respective states. The dimensionality of $\Delta \vec{x}_0$ and $\Delta \vec{x}_T$ is now the number of total (internal and input) spikes in the respective state descriptions. Let A_T denote the matrix such that $\Delta \vec{x}_T = A_T \Delta \vec{x}_0$. In the current framework, both $\Delta \vec{x}_0$ and $\Delta \vec{x}_T$ have components that are parallel to the trajectory, i.e., components that correspond to simple translation in time. To identify whether or not the spike trajectory is sensitive, these components have to be discarded. To better comprehend this subtlety, consider a perturbation $\Delta \vec{x}_T$ that lies parallel to the trajectory. This would correspond to the spikes in the perturbed trajectory being advanced in time, with the relative positions of all spikes in the perturbed trajectory being identical to those in the original trajectory. In dynamical systems theory, the original trajectory would not be considered sensitive, however large the magnitude of $\Delta \vec{x}_T$ might be. This issue is highlighted in the following extreme example. Consider a system all of whose spikes, internal as well as input, have been advanced in time by 1 msec. The ensuing dynamics of the system, assuming that all future input spikes are also advanced in time by 1 msec, would simply be a 1 msec advanced version of the dynamics of the original system. Clearly, this scenario has little relevance to the question of sensitive dependence. It was shown in (Banerjee, 2001b) that the discarding of the noted components can be achieved by pre and post multiplications with particular projection matrices. Let B and C be, respectively, these matrices that discard the noted component from the final and initial perturbations. $B * A_T * C$ then maps an initial perturbation that lies orthogonal to the spike trajectory to a corresponding final perturbation that also lies orthogonal to the spike trajectory. The spike trajectory is sensitive to initial conditions if $\lim_{T\to\infty} \|B*A_T*C\| = \infty$. If instead, $\lim_{T\to\infty} \|B * A_T * C\| = 0$, the trajectory is insensitive to initial conditions.

Simulation of single (spike response) model pyramidal neuron. A model pyramidal neuron was constructed using the Spike Response Model (Gerstner & van Hemmen, 1992), with 10,000 synapses, 85% of which were chosen to be excitatory and the rest inhibitory.

The threshold of the neuron was set at 15 mV above resting potential. The membrane potential function $P_i(\cdot)$ for the neuron was modeled as the sum of excitatory and inhibitory postsynaptic potentials (PSP) triggered by the arrival of spikes at synapses, and after-hyperpolarization potentials (AHP) triggered by the spikes generated by the neuron. PSPs were modeled using the function

$$P(t) = \frac{Q}{d\sqrt{t}} e^{-\beta d^2/t} e^{-t/\tau},\tag{1}$$

(see MacGregor and Lewis, 1977) to fit the physiological data reported in (Mason et al., 1991). Here Q denotes the connection strength of the synapse, d denotes the distance (in dimensionless units) of the synapse from the soma, and β and τ control the rate of rise and fall of the PSP. For EPSPs β was set at 1.0 and for IPSPs at 1.1. τ was set at 20 msec for all PSPs. d was sampled uniformly from the range [1.0,2.5]. Finally, Q for EPSPs and IPSPs were sampled uniformly from different ranges so as to satisfy the following final profiles. For EPSPs, the peak amplitudes ranged between 0.045 and 1.2 mV with the median around 0.15 mV, 10-90 rise times ranged from 0.75 to 3.35 msec and widths at half amplitude ranged from 8.15 to 18.5 msec. For IPSPs, the peak amplitudes were on average twice as large and the 10-90 rise times and widths at half amplitude were slightly larger. The AHP was modelled as the sum of a rectangular pulse that lasted 1msec (that introduced an absolute refractory period) and $-30*e^{-t/30}$ (that introduced a relative refractory period). Slow EPSPs/IPSPs/AHPs were not modelled since their corresponding α_i 's would, by definition, be small (see Eq. (2) in Results for the precise definition of α_i . Briefly, α_i is the normalized slope of the EPSP/IPSP/AHP generated by spike i, computed at the instant of the generation of the output spike). Finally, the length of the bounded past described in the previous section was set at 200 msec.

Whenever the neuron generated a new spike, the α_i 's for all contributing spikes were recorded and $\sum \alpha_i^2$ was computed. The sensitivity indicator, i.e., the mean $\langle \sum \alpha_i^2 \rangle$, was then computed over the set of all spike generations. In order to produce conservative estimates of the mean, samples with value above 10^4 (which was found to be about 0.1% of

the data) were discarded. The datasets ranged in size from 3000 to 15,000 spike generations.

Simulation of recurrent systems of neurons. A system comprising of 1000 internal neurons (modelling a cortical column, with 80% percent of the neurons set to be excitatory and the rest inhibitory) and 800 excitatory input neurons (modelling the input into the column) was constructed. Each internal neuron received 100 synapses from other (internal and input) neurons in the system. Two kinds of input drives were modelled; input neurons were set to generate random independent Poisson spike trains either at a fixed rate of 5 Hz, or at a rate (common across all input neurons) that modulated sinusoidally between 2.5 Hz and 7.5 Hz at a rate of 5 Hz. $P_i(\cdot)$'s for internal neurons were modelled using Eq. (1) as in the previous case, with parameters set to fit physiological data (Mason et al., 1991). Through exploratory simulations, a fixed set of synaptic strengths for excitatory and inhibitory synapses were found such that under the fixed rate input drive, the internal neurons spiked at a mean rate of 5 Hz (for excitatory) and 15 Hz (for inhibitory). These and all other parameters were then held constant over the entire system (under both input drive conditions) leaving the network connectivity and axonal delays as the only free parameters. After the generation of a spike, an absolute refractory period of 1 msec was introduced. There was no voltage reset. However, each spike triggered an AHP with a decay constant of 30 msec that led to a relative refractory period. Some of the networks thus generated were found to be capable of sustaining internal activity in the absence of input, albeit at a higher spike rate. Simulations were performed in 0.1 msec time steps. Whenever an internal neuron generated a new spike, the α_i 's for all contributing spikes were recorded and $\sum \alpha_i^2$ was computed. The sensitivity indicator, i.e., the mean $\langle \sum \alpha_i^2 \rangle$, was then computed over the set of all such spike generations. To compare this local indicator to the global rate of convergence/divergence, additional experiments were performed where the spike trajectory (for identical input spike trains) resulting from an internal state where exactly one randomly chosen spike was perturbed by 1 msec, was recorded. The divergence between the original and the perturbed spike trajectory was then computed using the spike time metric $D^{\text{spike}}[q]$ (Victor & Purpura, 1996). $D^{\text{spike}}[q]$ measures the distance between two spike trajectories as the minimal 'cost' of transforming one train into the other via a sequence of elementary steps: inserts, deletions and shifts of spikes. Each elementary step is associated with a fixed cost. For eliminating or inserting a spike, the cost is 1. To shift a spike, the cost is equal to $q\Delta t$, where Δt is the extent of the shift. We used a high value of q, making the metric sensitive to small spike time differences. However, we found that the results were similar for a wide range of values for q. The temporal course of the metric $D^{\rm spike}[q]$ between the original and perturbed spike trains was analyzed. Intuitive understanding of the value taken by $D^{\rm spike}[q]$ can be obtained by noting that for high values of q, the distance between two independent spike trains is primarily controlled by the rate of firing of the two trains. Consider for example two independent spike trains with a rate of 10 Hz. This rate corresponds to an average of 1 spike in any window of 100 msec. If the spike trains are sufficiently different, the minimal cost of transforming one spike train into the other is obtained by deleting each spike (at a cost of 1) and replacing it (at an additional cost of 1). The average $D^{\rm spike}[q]$ per window of 100 msec is thus 2, or more generally 0.2 times the average firing rate.

Simulation of orientation model. The orientation model that was simulated is similar to the model of Somers et al. (1995). The details of our particular implementation can be found in (Seriès et al., 2004). Briefly, the model represented three stages, retina, LGN and V1. The retinal stage spanning a 4° by 4° monocular patch of the central visual field, corresponded to 21×21 grid of ON and OFF ganglion cells, modelled by Difference-of-Gaussian filters. The output of each filter was used to drive the LGN cells which generated Poisson spike trains. The output of the LGN cells was pooled using Gabor receptive fields and acted as input to V1. The V1 stage, representing one hypercolumn of layer IV, was composed of 1260 simple cells, modelled as conductance-based integrate-and-fire neurons, 80% (i.e., 1008 cells) of which were regular-spiking excitatory and the rest (i.e., 252 cells) fast-spiking inhibitory neurons. These cells were coupled through lateral projections. Excitatory cells formed short-range connections, while inhibitory cells could target cells with a broad range of preferred orientations. Specifically, the probability of a connection between two cells was a Gaussian function of the difference between their preferred orientations. The standard deviation for the Gaussian was 7.5° for excitatory projections

and 60° for inhibitory projections. The sampling was done in a manner such that each cell (excitatory as well as inhibitory) received input from 40 excitatory V1 cells and 30 inhibitory V1 cells. The values of the synaptic conductances, and all other parameters can be found in (Seriès et al., 2004, Suppl Info). Exploratory simulations showed that the network eventually became quiescent with the cessation of input. Oriented flashed bar stimuli were presented to the retinal cells, and the responses of the cortical cells were recorded. The dynamics of the network during a normal simulation was then compared to a simulation in which one spike of one excitatory cortical cell was perturbed by 1 msec (the LGN spike trains being identical). To quantify the divergence of the spike trajectory, only the spike time metric $D^{\rm spike}[q]$ was used since the local perturbation matrix or the sensitivity indicator, i.e., $\langle \sum \alpha_i^2 \rangle$, could not be recovered from the dynamics of the conductance-based model of the neuron.

Results

Every time a spike is generated by a neuron in a network, it is due to the combined effects of the postsynaptic potentials (PSP) of all the spikes that have already arrived at the synapses of the neuron, and the after-hyperpolarization potentials (AHP) of the previous spikes generated by the neuron. As illustrated in Figure 2, perturbing any of these spikes affects the timing of the output spike (only presynaptic spikes are considered in the diagram, but the result applies just as well to the previous spikes generated by the neuron).

Let $\Delta x_{new} = \sum_{i=1}^{n} \alpha_i \Delta x_i$, where Δx_{new} is the perturbation in the output spike and Δx_i is the perturbation in the i^{th} contributing spike (i.e., a presynaptic spike of the neuron or a spike generated by the neuron). It was shown in (Banerjee, 2001b) that the weight α_i is proportional to the slope, ρ_i , of the PSP (or AHP) triggered by spike i (the partial derivative of the membrane potential with respect to spike i, in the general case) at the instant of the generation of the output spike. The weight is given by the normalized slope:

$$\alpha_i = \rho_i / \sum_{j=1}^n \rho_j , \qquad (2)$$

where j ranges over all the contributing spikes. To further elucidate, consider the membrane potential of a neuron that has just crossed threshold, causing the neuron to generate a spike. Had any of the contributing pre/postsynaptic spikes of this neuron been perturbed, the neuron would not have generated a spike at the noted time, and therefore, the membrane potential would not have been at threshold at that time. ρ_i corresponds to the ratio of this putative change in the membrane potential, to the perturbation in the timing of the i^{th} contributing spike, in the limit as the perturbation goes to zero. In the case of a spike-response model, this reduces to the slope of the PSP/AHP generated by the i^{th} contributing spike. α_i is computed by normalizing ρ_i across all contributing spikes.

For each internal spike generated in a network, one can in principle compute the scalar quantity $\sum \alpha_i^2$, the summation taken over all those spikes that contributed to the generation of that spike. It was demonstrated in (Banerjee, 2001b) that the mean of the $\sum \alpha_i^2$'s computed over the history of spike generations in the network, determines the sensitivity of the spike trajectory to initial conditions (see also Coombes & Bressloff, 1999). We have refined this approach for systems receiving stationary inputs (Banerjee & Pouget, 2003; Banerjee, 2006) and have found under a reasonable set of assumptions (described next) that if:

$$\left\langle \sum_{i=1}^{n} \alpha_{i}^{2} \right\rangle > \frac{(2 + O(1/m))}{\mu} - 1, \tag{3a}$$

then the spike trajectory is almost surely sensitive to initial conditions, i.e., $\lim_{T\to\infty} \|B*A_T*C\| = \infty$ with probability 1 (see Materials and Methods).

If instead,

$$\left\langle \sum_{i=1}^{n} \alpha_{i}^{2} \right\rangle < \frac{1}{\mu} - 1, \tag{3b}$$

then the spike trajectory is almost surely insensitive to initial conditions, i.e., $\lim_{T\to\infty} \|B*A_T*C\| = 0$ with probability 1. The angular brackets $\langle \cdot \rangle$ in (3a) and (3b) denote the mean computed over the entire spike trajectory, O(1/m) denotes a term on the order of (1/m) where m is the number of total (internal and input) spikes in the state, and μ denotes

the average ratio of the number of internal spikes to the number of total (internal and input) spikes in the state.

Under stationary conditions (i.e., when the stochastic process modelling the spike generation process can be described by a time invariant, fixed probability distribution) one can assume that at the generation of each internal spike, the ρ_i 's for internal and input contributing spikes arise from stationary distributions, and that the ratio of the number of internal to the total (internal plus input) spikes in any state remains close to a fixed quantity μ at all times. We assume in addition that at the generation of each spike, the ρ_i 's for both internal and input contributing spikes arise from the same fixed distribution independent of one another and of all ρ_i 's for spikes that were generated earlier in the system (an assumption that we have tested to hold well in our simulations). This assumption is quite unlike the more restrictive assumptions made by van Vreeswijk & Sompolinsky (1998), such as random connectivity and uncorrelated inputs, because it applies to the local gradients around spike trajectories rather than the trajectories themselves. The proof of (3) is based on a small modification to the one presented in (Banerjee, 2001b) and is reported in (Banerjee, 2006).

If we assume that input spikes, on average, account for as many as half the total number of spikes in state descriptions (i.e., μ =0.5, a liberal assumption by cortical network standards), then since the number of total spikes in any state description is very large, the constraint (3) reduces to $\langle \sum \alpha_i^2 \rangle > 3$ for spike trajectories to be almost surely sensitive to initial conditions.

An examination of $\sum \alpha_i^2$, taking Eq. (2) into consideration, reveals that its value rises as the size of the subset of ρ_i 's that are negative grows larger. Indeed, negative ρ_i 's reduce the value of $\sum_{j=1}^n \rho_j$, thus increasing $\sum \alpha_i^2$. A spike trajectory is therefore more likely to be sensitive when a substantial number of the EPSPs are on their falling phase (and IPSPs on their rising phase) at the instant of the generation of each spike. This is the case, for instance, in a network containing excitatory and inhibitory neurons firing unsynchronized, near Poisson spike trains—as observed in cortex. Conversely, if spikes are

generated soon after the arrival of synchronized bursts of spikes all of whose EPSPs are presumably on their rising phase, the spike trajectory is less likely to be sensitive.

Single (Spike Response) Model Pyramidal Neuron

One of the advantages of the criteria in (3) is that if we were to assume that the individual neurons in a network are operating under similar conditions, we could estimate the value of $\langle \sum \alpha_i^2 \rangle$ for the dynamics of the entire network by computing its value for a single neuron. We therefore simulated a single pyramidal neuron under various input scenarios to identify conditions under which the dynamics of a network would be almost surely sensitive or insensitive to initial conditions (See Methods).

Three experiments simulating various levels of uncorrelated input/output activity were conducted. In particular, excitatory Poisson inputs at 2, 20 and 40 Hz were balanced by inhibitory Poisson inputs at 6.3, 63 and 124 Hz to generate output rates of approximately 2, 20 and 40 Hz, respectively. The output in all three cases was Poisson-like (CV=0.77, 0.74, and 0.89, respectively). The mean $\langle \sum \alpha_i^2 \rangle$ for the three experiments was 4.37, 5.66, and 9.52, respectively.

Next, two sets of experiments simulating the arrival of regularly spaced synchronized bursts of spikes were conducted. In the first set the random background activity was set at 2 Hz and in the second, at 20 Hz. The synchronized bursts comprised of spike volleys that arrived every 50 msec. Four experiments were conducted within each set: volleys were composed of either 100 or 200 spikes (producing jolts of around 10 and 20 mV respectively) that were either fully synchronized or were dispersed over a Gaussian distribution with σ =1 msec. The mean $\langle \sum \alpha_i^2 \rangle$ for the experiments was as follows. At 2 Hz background activity, it was 0.49 (200 spikes/volley, synchronized), 0.60 (200 spikes/volley, dispersed), 2.46 (100 spikes/volley, synchronized), and 2.16 (100 spikes/volley, dispersed). We can therefore conclude that the corresponding spike trains are unlikely to be sensitive, which is similar to the finding of Diesmann et al. (1999) except that our result applies to recurrent activity. By contrast, at 20 Hz background activity, we found that the spike trains

are likely to be sensitive; $\langle \sum \alpha_i^2 \rangle$ was 4.39 (200 spikes/volley, synchronized), 8.32 (200 spikes/volley, dispersed), 6.77 (100 spikes/volley, synchronized), and 6.78 (100 spikes/volley, dispersed).

Finally, two sets of experiments simulating the arrival of randomly spaced synchronized bursts of spikes were conducted. In the first set the random background activity was set at 2 Hz and in the second, at 20 Hz. The synchronized bursts comprised of a sequence of spike volleys that arrived randomly at a rate of 20 Hz. Two experiments were conducted within each set: volleys were composed of either 100 or 200 synchronized spikes. The mean $\langle \sum \alpha_i^2 \rangle$ for the experiments was as follows. At 2 Hz background activity, it was 4.30 (200 spikes/volley) and 4.64 (100 spikes/volley). At 20 Hz background activity, it was 5.24 (200 spikes/volley) and 6.28 (100 spikes/volley). Therefore, when the synchronized bursts arrive randomly, the spike trains are likely to be sensitive regardless of the level of background activity.

Recurrent Systems of Neurons

The results from the simulations of a single pyramidal neuron indicated that networks composed of a majority of excitatory neurons that operated at low to moderate levels of sustained, Poisson-like, weakly correlated activity, as found in the cortex, would likely be sensitive to initial conditions. In order to determine whether this conclusion applied to the particular spike trajectories generated by recurrent networks of neurons, we simulated several instantiations of model cortical networks and applied the preceding analysis to the recorded spike trajectories.

Each network was composed of 1000 neurons driven by Poisson input spike trains at fixed or sinusoidally modulating rates, from 800 excitatory neurons (see Methods). While the input was asynchronous in the fixed rate case, it was bursty in the modulating rate case. Several connectivity patterns were considered. On the one extreme was an ordered two layer ring network with input neurons forming the lower layer and internal neurons (with the inhibitory neurons placed evenly among the excitatory neurons) forming the upper layer. Each internal neuron received inputs from a sector of internal and input

neurons that was centered on that neuron. As a result, any two neighboring internal neurons shared 96 of their 100 inputs (with the same axonal delay of 0.8 msec). This led the system to generate spikes in bursts, irrespective of whether the input drive was at a fixed rate or at a sinusoidally modulating rate (as is apparent from the streaks/coincidental spikes in the spike rasters in Figure 4a. and 4b.). On the other extreme was a network where each internal neuron received inputs from 100 randomly chosen neurons from the entire population of internal and input neurons with random axonal delays of 0.5-1.1 msec. Here the system generated spikes asynchronously when the input drive was at a fixed rate, and in bursts when the input drive was at a sinusoidally modulating rate (as is apparent from the presence and absence of streaks/coincidental spikes in the spike rasters in Figure 3b. and 3a., respectively). Several other networks where neighboring internal neurons shared an intermediate percentage of their inputs were also simulated.

In each case, two different sets of experiments were performed. In the first set, a randomly chosen spike in the state description was perturbed by 1 msec and the subsequent spike trajectory was recorded. The global divergence between the original and the perturbed spike trajectory was then computed using the metric $D^{\text{spike}}[q]$ for intervals of 100 msec at every 10 msec, averaged over several trials. In each case we found that the systems were sensitive and the spike trajectories desynchronized within 100 to 200 msec, eventually reaching a distance $D^{\text{spike}}[q]$ similar to the one that would be observed between completely independent spike trains with the same statistics (as is apparent from the blue and red spikes in Figures 3a., b. and 4a., b., and the levelling off of the $D^{\text{spike}}[q]$ graphs in Figures 3c. and 4c.).

In the second set of experiments, the statistic $\left\langle \sum \alpha_i^2 \right\rangle$ was computed from data recorded from the sequence of births of internal spikes in the trajectory. The value of the statistic was found to be consistent with the formal result. The value of $\left\langle \sum \alpha_i^2 \right\rangle$ was 47.7 for the fixed rate input and 25.9 for the modulating rate input for the ring network, and 11.3 for the fixed rate input and 5.6 for the modulating rate input for the random network. $\left\langle \sum \alpha_i^2 \right\rangle$ for all the other networks were also greater than 3.

Orientation Model

The statistic $\langle \sum \alpha_i^2 \rangle$ can be difficult to compute for neuronal models such as the conductance based integrate-and-fire neuron. To test whether the results of the previous experiments generalized to such models, we perturbed the dynamics of a model of orientation selectivity in V1, consisting of a recurrent network of conductance based neurons (see Methods). The cortical neurons in this model received Poisson spike trains from the lateral geniculate nucleus and in turn generated spike trains with near Poisson statistics.

Model cortical cells exhibited sharp orientation selectivity despite receiving weakly tuned thalamocortical inputs. We compared the dynamics of the network during normal simulations to simulations in which one spike of an excitatory cortical cell was perturbed by 1 msec. Just as in the previous experiments, all our simulations revealed that perturbing a single spike by 1 msec desynchronized the spike trajectory within 100 to 200 msec (as is apparent from the blue and red spikes in Figure 5a. and 5b.).

Discussion

In (Banerjee, 2001b; Banerjee & Pouget, 2003; Banerjee, 2006) we derived formal results that showed that whether or not the dynamics of a recurrent network of spiking neurons depends sensitively on its internal state is intimately tied to the value of $\langle \sum \alpha_i^2 \rangle$ computed over the set of all internal spikes generated in the network. In order to identify scenarios in which the dynamics of a system would (or would not) be sensitive to initial conditions, we computed the value of $\langle \sum \alpha_i^2 \rangle$ for the dynamics of a single model neuron operating under such conditions. The underlying assumption was that the results from the single neuron experiments would be indicative of the value of $\langle \sum \alpha_i^2 \rangle$ for a system with all of its neurons operating under like conditions.

In these simulations, we considered among other cases, the neuron firing at near Poisson statistics in response to independent (i.e., unsynchronized) Poisson input spike trains from excitatory and inhibitory neurons. For all the input rates tested, $\langle \sum \alpha_i^2 \rangle$ was above 3, indicating that the corresponding dynamics of the system would almost surely be sensitive. The only case where $\langle \sum \alpha_i^2 \rangle$ was found to be less than 3 was when the neuron received synchronized bursts of 100 or more spikes at regular intervals of 50 msec in the presence of 2 Hz of random background activity. Apart from the fact that our result concerns recurrent activity, this scenario is much like that simulated in (Diesmann et al., 1999). However, when the neuron received these bursts of spikes randomly (as opposed to regularly) with a mean interval of 50 msec between bursts, $\langle \sum \alpha_i^2 \rangle$ was found to be larger than 3. Finally, for higher values of background activity (20 Hz), $\langle \sum \alpha_i^2 \rangle$ was found to be substantially larger than 3 in all scenarios.

To confirm whether these findings applied to entire spike trajectories, we simulated the dynamics of several recurrent systems under a variety of input conditions. The spike trajectories generated by the recurrent systems of spiking neurons of both the spikeresponse type as well as the conductance based integrate-and-fire type were also found to be sensitive to the internal state of the network in all scenarios. This included cases where the dynamics of the network was either bursty or asynchronous, and when the input drive was either at a fixed rate or at a sinusoidally modulating rate. In addition to computing $\langle \sum \alpha_i^2 \rangle$, we also confirmed that the dynamics was in fact sensitive by perturbing a randomly chosen internal spike and recording the subsequent spike trajectory under identical input drive. The perturbed trajectories of the systems driven by sinusoidally varying inputs were particularly instructive in that they demonstrated that a system could generate synchronized bursts of spikes and at the same time be sensitive. Not only were the bursts themselves perturbed in time, but also the bursts were generated by different subsets of neurons (as can be seen in the spike rasters in Figures 3b. and 4b.). These results indicate that it is unlikely that recurrent cortical networks of integrate-and-fire, bounded memory neurons can robustly generate precise spike trajectories under conditions prevalent in the cortex, i.e., where the constituent neurons generate sustained, weakly correlated, Poissonlike spike trains at low to moderate rates. We must emphasize that this conclusion is quite independent of the observation that neurons do on average phase lock to oscillations either in the input drive (as seen in Figure 3b) or in the intrinsic dynamics of the network (as seen in Figure 4a and b).

As mentioned earlier, the interplay between rising and falling PSPs at the births of spikes plays a role in our findings. Indeed, approximate balance between the slopes of the EPSPs/IPSPs at the birth of spikes, makes the sum of the slopes, $\sum_{j=1}^{n} \rho_{j}$, small on average. This makes the α_{i} 's large, which in turn leads to sensitivity. This presents an interesting and testable experimental prediction. If we are correct about the sensitivity of the dynamics of recurrent cortical networks, the distribution of the slopes of the membrane potentials of cortical neurons near the firing threshold should be skewed towards small values, i.e., most threshold crossings should take place with small slopes. Under such conditions, small jitters in the timing of the presynaptic spikes would result in a large change in the timing of the postsynaptic spike. This prediction could be tested with patch, or intracellular recordings, in awake animals.

One potential concern with our result is our assumption that cortical neurons fire with near Poisson statistics, which is not strictly true in the cortex. For instance, the Fano factor (the ratio of the variance to the mean of the spike counts) which is equal to one for a Poisson process, varies between 0.3 and 1.8 in the cortex (Tolhurst et al., 1982; Gershon et al., 1998; Gur & Snodderly, 2005). The assumption of near-Poisson variability, however, is not critical to our results. As we have just discussed, the critical factor is $\left\langle \sum \alpha_i^2 \right\rangle$ which depends on the slope of the contributing PSPs/AHPs at threshold crossings. If the input spike trains are Poisson-like and weakly correlated, then the slope at threshold is small as indicated by our single neuron simulations. But this can also be true when spike statistics deviate from Poisson. In fact, in our network simulations using both the spike response and the conductance-based integrate-and-fire neurons, the Fano factor varied from 0.4 to 0.8, and yet the networks were found to be sensitive to initial conditions.

It must be noted that whether spike trajectories are robust to perturbations depends not only on the details of the network architecture and the statistics of the spike trains, but also on the range of firing rates. For instance, if neurons in a network fire at approximately

500Hz, the average spike timing is likely to be precise to a 2 msec time scale, because of the refractory period. In fact, we have found in other simulation experiments (not reported here) that the value of $\left\langle \sum \alpha_i^2 \right\rangle$ drops sharply as the spike rates of the neurons in a network rise beyond 200 Hz. This was found to be due to the positive α_i 's contributed by the AHPs of the efferent spikes of the neurons. In this article, we have focused instead on the question of whether entire spike trajectories—with spike timing precise to within a 1 to 5 msec range—can be generated in recurrent networks in the cortex of awake behaving animals. Accordingly, we considered situations in which the internal neurons fire in the 0.1-50 Hz range with near-Poisson statistics, and are weakly correlated, because these are the most common type of responses reported in the cortex of awake behaving animals (Tolhurst et al. 1982; Gershon et al., 1998; Gur & Snodderly, 2005). We also focused on recurrent networks where the output is a function not only of the input but also the internal state of the network, as opposed to feedforward networks, because lateral connections are ubiquitous throughout the cortex (Braitenberg & Schuz, 1991). Our results demonstrated that even those spike trajectories that may not be strictly categorized as Poisson-like were sensitive to initial conditions, when the neurons spiked at low to moderate rates.

Naturally, our results do not imply that neural circuits in general can not utilize precise timing of spikes to represent and manipulate information, and there are documented examples of such usage in functionally feedforward systems (Carr & Konishi, 1990; Kawasaki et al., 1988). Since our analysis is based on perturbations of the internals state of a network, systems where the output is solely a function of the input do not fall within the purview of our result. Several researchers have also reported spike patterns at a coarser time scale in the cortex, i.e., on the order of 25+ msec (Nadasdy et al., 1999; Hahnloser et al. 2002). Such reports do not contradict our result due to the coarseness of their time scale (the average spike in a 40 Hz Poisson spike train has 25+ msec precision).

Finally, it must be noted that sensitive dependence in the cortex does not preclude all kinds of temporal codes. On the contrary, it makes the question even more tantalizing by ruling out the most basic form of temporal code where information is represented in the *precise* (modulo the jitter introduced by thermal noise at each component neuron estimated

to be in the 1-5 msec range) spike trajectory generated by the *entire* network. In fact, the predominant impact of sensitive dependence is likely to be an increased complexity in the topology of the attractors (detectable only when all input into a network is withdrawn) in the dynamics of cortical networks, which would in turn cause the spatio-temporal signature of any information coded in its dynamics to be vastly more complex. Several researchers (Skarda & Freeman 1987; van Vreeswijk & Sompolinsky, 1998) have also noted other advantages of chaotic dynamics for computational systems. The precise implications that sensitive dependence has for the computational nature of systems of spiking neurons is however far from clear.

References

Banerjee, A. On the phase-space dynamics of systems of spiking neurons I: model and experiments. *Neural Computation* **13**, 161 (2001a).

Banerjee, A. On the phase-space dynamics of systems of spiking neurons II: formal analysis. *Neural Computation* **13**, 195 (2001b).

Banerjee, A., & Pouget, A. Dynamical constraints on computing with spike timing in the cortex. *Advances in Neural Information Processing Systems* **15** (2003).

Banerjee, A. On the sensitive dependence on initial conditions of the dynamics of networks of spiking neurons. *Journal of Computational Neuroscience* **20**, 321 (2006).

Braitenberg, V., & Schuz, A. Anatomy of the Cortex. Berlin: Spring-Verlag, (1991).

Carr, C. E., & Konishi, M. A circuit for detection of interaural time difference in the brain stem of the barn owl. *Journal of Neuroscience*, **10**, 3227 (1990).

Compte, A., Constantinidis, C., Tegner, J., Raghavachari, S., Chafee, M.V., Goldman-Rakic, P.S. & Wang, X-J. Temporally irregular mnemonic persistent activity in prefrontal neurons of monkeys during delayed response rask. *Journal of Neurophysiology* **90**, 3441 (2003).

Coombes, S., & Bressloff, P. C. Mode locking and Arnold tongues in integrate-and-fire neural oscillators. *Physical Review E* **60**, 2086 (1999).

Diesmann, M., Gewaltig, M. O., & Aertsen, A. Stable propagation of synchronous spiking in cortical neural networks. *Nature* **402**, 529 (1999).

Gershon, E.D., Wiener M.C., Latham, P.E. & Richmond, B.J. Coding strategies in monkey V1 and inferior temporal cortices. *Journal of Neurophysiology* **79**, 1135 (1998).

Gerstner, W., & van Hemmen, J. L. Associative memory in a network of spiking neurons. *Network* **3**, 139 (1992).

Gur, M., & Snodderly, D. M., High Response Reliability of Neurons in Primary Visual Cortex (V1) of Alert, Trained Monkeys. *Cerebral Cortex*, (2005).

Hahnloser, R. H. R., Kozhevnikov, A., & Fee, M. S. An ultrasparse code underlies the generation of neural sequences in songbirds. *Nature*, **419**, 65 (2002).

Hansel, D., & Sompolinsky, H. Chaos and synchrony in a model of a hypercolumn in visual cortex. *Journal of Computational Neuroscience* **3**, 7 (1996).

Kawasaki, M., Rose, G., & Heiligenberg, W. Temporal hyperacuity in single neurons of electric fish. *Nature* **336**, 173 (1988).

Maass, W., Natschläger T., & Markram, H. A model for real time computation in generic neural microcircuits. *Advances in Neural Information Processing Systems* **15** (2003).

Maass, W. Fast sigmoidal networks via spiking neurons. *Neural Computation* **9**, 279 (1997).

MacGregor, R. J., and Lewis, E. R. Neural Modeling, Plenum Press New York (1977).

Mainen, Z. F., & Sejnowski, T. Reliability of spike timing in neocortical neurons. *Science* **268**, 1503 (1995).

Mason, A., Nicoll, A., & Stratford, K. Synaptic transmission between individual pyramidal neurons of the rat visual cortex in vitro. *Journal of Neuroscience* **11**(1), 72 (1991).

Mehring, C., Hehl, U., Kubo, M., Diesmann, M., & Aertsen, A. Activity dynamics and propagation of synchronous spiking in locally connected random networks. *Biological Cybernetics* **88**, 395 (2003).

Novak, L. G., Sanches-Vives, M. V., & McCormick, D. A. Influence of low and high frequency inputs on spike timing in visual cortical neurons. *Cerebral Cortex* **7**, 487 (1997).

Nadasdy, Z., Hirase, H., Czurko, A., Csicsvari, J., & Buzsaki, G. Replay and time compression of recurring spike sequences in the hippocampus. *Journal of Neuroscience* **9**, 9497 (1999).

Osborne, A. R., & Provenzale, A. Finite correlation dimension for stochastic systems with power-law spectra, *Physica D* **35**, 357, (1989).

Panzeri, S., Petersen, R. S., Schultz, S. R., Lebedev, M., & Diamond, M. E. The role of spike timing in the coding of stimulus location in rat somatosensory cortex. *Neuron* **29**, 769 (2001).

Reinagel, P., & Reid, R. C. Temporal coding of visual information in the thalamus. *Journal of Neuroscience* **20**, 5392 (2000).

Rieke, F., Warland, D., de Ruyter van Steveninck, R. R., & Bialek, W., *Spikes: Exploring the Neural Code*, MIT Press, Cambridge, (1997).

Siapas, A. G., Lubenov, E. V., & Wilson, M. A. Prefrontal phase locking to hippocampal phase oscillations. *Neuron* **46**, 141 (2005).

Seriès, P., Latham, P., & Pouget, A. Statistical efficiency of orientation selectivity models. *Nature Neuroscience* **7**, 1129 (2004).

Skarda, C. A., & Freeman, W. J. How brains make chaos in order to make sense of the world *Behavioral and Brain Sciences*, **10**, 161 (1987).

Somers, D. C., Nelson, S. B., & Sur, M. An emergent model of orientation selectivity in cat visual cortex simple cells. *Journal of Neuroscience* **15**(8), 5448 (1995).

Tolhurst, D., Movshon, J., & Dean, A. The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Research* **23**, 775 (1982).

van Vreeswijk, C., & Sompolinsky, H. Chaotic balanced state in a model of cortical circuits. *Neural Computation* **10**, 1321 (1998).

Victor, J. D., & Purpura, K. Nature and precision of temporal coding in visual cortex: a metric-space analysis. *Journal of Neurophysiology*, **76**, 1310 (1996).

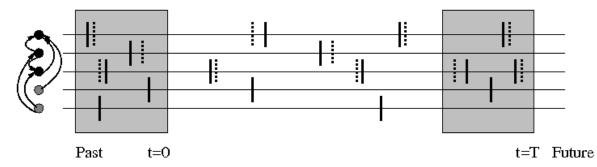


Figure 1: Schematic diagram of the dynamics of a system of neurons. Input neurons are colored gray and internal neurons black. Spikes are shown as solid lines and their perturbations as dotted lines. Note that the spikes generated by input neurons are not perturbed. Gray boxes demarcate a bounded past history starting at time t. The temporal position of all spikes in the boxes specify the state of the system at t=0 and t=T.

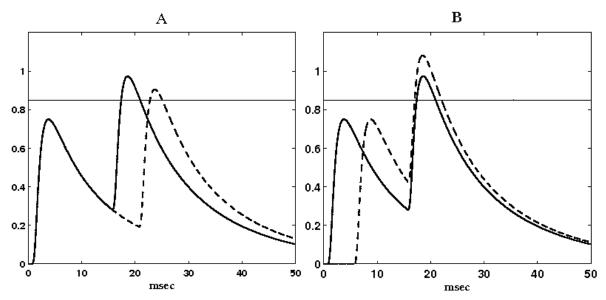


Figure 2: Schematic diagram of the membrane potential of a neuron that reaches threshold soon after the arrival of two spikes. The threshold is shown as a horizontal line and the EPSPs resulting from the spikes are shown in solid. In (a) (respectively, (b)), the result of perturbing the second (respectively, first) spike forward in time is shown as a dotted curve. Whereas in (a) the threshold is reached later, causing the output spike to be perturbed forward in time, in (b), it is reached earlier causing the output spike to be perturbed backward in time. The reason for this behavior lies in the sign of the slope of the respective PSPs at the instant of generation of the output spike; the second PSP is rising whereas the first is falling. The magnitude of the slope also matters. The larger the slope is, the greater is the perturbation on the output spike.

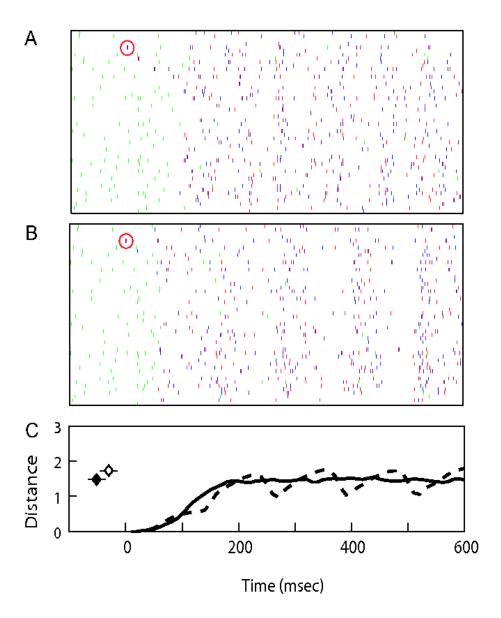


Figure 3: (a,b) Spike trains of 50 neighbouring neurons for 700 msec from the random network, for two types of input drives, constant (a) or sinusoidally modulated at 5 Hz (b). In both cases, the same system was initialized at identical states except for 1 spike (highlighted) generated at approximately 100 msec that was perturbed by 1 msec. Origin is set at the time of this perturbation. Simultaneous spikes (precise to within 1 msec) in the normal and perturbed simulations are depicted in green. The spikes whose timing were affected by more than 1 msec are shown in red (normal) and blue (perturbed). As is evident, the two spike trajectories diverge rapidly. (c) Mean divergence of the spike trains after the perturbation, for the entire population. The curves correspond to the distance $D^{\text{spike}}[q]$ between the normal and perturbed spike trains, for constant (full lines) and oscillatory (dashed lines) input, averaged over all neurons and 10 trials. At each time step t, $D^{\text{spike}}[q]$ was computed over [t-100, t] msec, with cost q set at 100 msec⁻¹. The black and white diamonds

indicate the baseline value of $D^{\text{spike}}[q]$ that would be obtained by comparing independent spike trains with the same response statistics, for constant and oscillatory input respectively. This value is computed by measuring, for each neuron, the distance between two spike trains recorded on independent trials (different seeds). Although $D^{\text{spike}}[q]$ depends on q for small q, the temporal course of the divergence was found to be similar for a large range of q's.

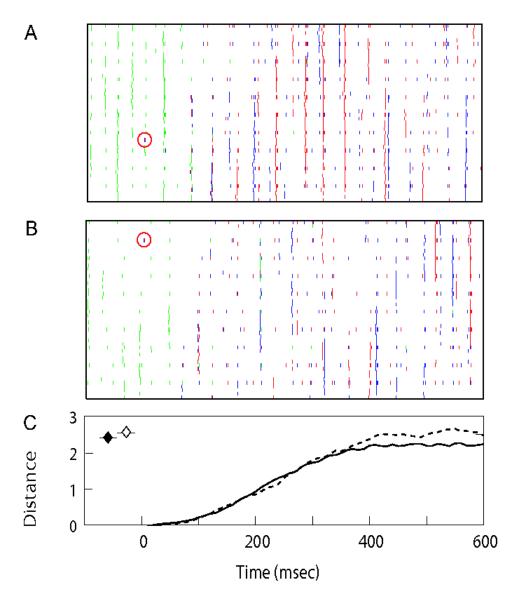


Figure 4: (a,b) Spike trains of 50 neighbouring neurons for 700 msec from the ring network for two types of input drives, constant (a) or sinusoidally modulated at 5 Hz (b). In both cases, the same system was initialized at identical states except for 1 spike (highlighted) generated at approximately 100 msec that was perturbed by 1 msec. Origin is set at the time of this perturbation. Simultaneous spikes (precise to within 1

msec) in the normal and perturbed simulations are depicted in green. The spikes whose timing were affected by more than 1 msec are shown in red (normal) and blue (perturbed). For these neurons located in the neighbourhood of the perturbed neuron, divergence is very fast. (c) Mean divergence of the spike trains after the perturbation, for the entire population. The curves correspond to the distance $D^{\text{spike}}[q]$ between the normal and perturbed spike trains, for constant (full lines) and oscillatory input (dashed lines), averaged over all neurons and 10 trials. The black and white diamonds indicate the baseline value of $D^{\text{spike}}[q]$ as in Figure 3.

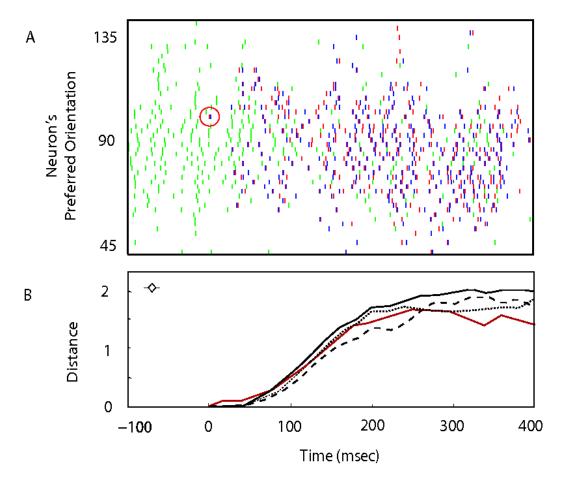


Figure 5: (a) Spike trains of 50 neurons with preferred orientation in [45°-135°] in response to a bar oriented at 90°, during a normal run and after perturbing 1 cortical spike (highlighted) by 1 msec, from a model of orientation selectivity in V1. Perturbation occurred around t=0 msec for a neuron with preferred orientation 80°. Simultaneous spikes (precise to within 1 msec) in the normal and perturbed simulations are depicted in green. The spikes whose timing were affected by more than 1 msec are shown in red (normal) and blue (perturbed). (b) Divergence of the spike trains after the perturbation, for the neuron receiving the initial perturbation (red) and other neurons picked at random (other curves). Each curve corresponds to the distance $D^{\text{spike}}[q]$ between the normal and perturbed spike trains, averaged over 50 trials.