Cortical circuit-based lossless neural integrator for perceptual decision-making

Jung H. Lee[1], Joji Tsunada[2], Sujith Vijayan[3], and Yale E. Cohen[2,4,5]

Correspondence and Lead Contact: Jung H. Lee at jungl@alleninstitute.org

[1] Allen Institute for Brain Science, Seattle, WA 98109

[2] Departments of Otorhinolaryngology, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA 19104

[3] School of Neuroscience, Virginia Tech, Blacksburg VA, 24061

[4] Department of Neuroscience, University of Pennsylvania, Philadelphia PA, 19104

[5] Department of Bioengineering, University of Pennsylvania, Philadelphia, PA 19104

Summary

The intrinsic uncertainty of sensory information (i.e., evidence) does not necessarily deter an observer from making a reliable decision. It is believed that uncertainty is reduced by integrating (accumulating) incoming sensory evidence. Traditionally, recurrent network models have been thought to underlie the integration of evidence necessary for decision-making, but they have yet to fully explain humans' ability to hold accumulated evidence during temporal gaps in sensory inputs. In our study, using a computational model, we show that cortical circuits can readily explain humans' ability to make robust decisions during temporal gaps. Specifically, our model predicts that the interplay among depressing synapses and the two major interneuron types allow the cortex to switch flexibly between 'retention' and 'integration' modes. The evidence can be read out in two different modes mirroring 'stepping' and 'ramping' activity observed empirically, suggesting that these two activity patterns emerge from the same mechanism.

Keywords: perceptual decision-making; neural integrator; computational model of cortex; inhibitory neuron types; depressing synapses; ramping and stepping activity

Introduction

One of the fundamental operations of the brain is to transform representations of external sensory stimuli into a categorical judgment, despite the inherent uncertainty of the incoming sensory evidence. For instance, we can determine the direction of the wind, even though its instantaneous direction continuously fluctuates. It is widely thought that this moment-by-moment uncertainty is minimized by temporally integrating (accumulating) this incoming sensory evidence (Gold and Shadlen, 2007; Goldman et al., 2009; Roitman and Shadlen, 2002). Potential neural correlates of this neural accumulation of sensory evidence have been identified in a variety of brain areas, including the lateral intraparietal cortex (LIP), prefrontal cortex (Kim and Shadlen, 1999), and frontal eye fields (Ding and Gold, 2012). In particular, spiking activity in these brain areas appears to 'ramp up' (accumulate) prior to a perceptual decision. Further, the rate of this accumulation (and, thus, the time to the perceptual decision) is negatively correlated with the ambiguity of the sensory evidence: as the evidence becomes less ambiguous (e.g., the instantaneous fluctuations in wind direction decrease), the rate of the ramping increases (Gold and Shadlen, 2007).

Such neural integration has been modeled in two very different ways, which rely on different coding strategies and mechanisms of integration (Goldman et al., 2009). In the first type of model, rate-code neural integrators (NI) integrate sensory evidence and represent accumulated evidence as monotonically increasing ('ramping') spiking activity. In this rate-code model (Gold and Shadlen, 2007; Roitman and Shadlen, 2002; Wang, 2012), the firing rates of individual neurons increase over time in response to continuous inputs. In an alternative model, location-code NIs store accumulated evidence as the location of a highly localized elevated spiking activity. In such a location-code NI (Skaggs et al., 1995; Song and Wang, 2005), the location of

highly active neurons, referred to as ‘bump’, travels through a network over time. That is, the location of bump activity corresponds to the total amount of accumulated evidence.

Because ramping activity has been found in several studies of perceptual decision-making (Gold and Shadlen, 2007; Goldman et al., 2009), it is generally believed that a rate-code NI is the more natural circuit candidate for neural integration of sensory information. However, recent behavioral studies have questioned whether a rate-code NI could, in fact, accurately capture the dynamics of perceptual decision-making. For example, a temporal gap between stimulus presentations has little impact on the accuracy of an observer’s behavioral choices (Kiani et al., 2013; Liu et al., 2015), indicating that accumulated evidence can be maintained during this temporal gap. Yet, during this gap, the firing rates of neurons in a rate-code NI are likely to deviate from the desired values if the network is perturbed even slightly. This deviation can occur because the balance between a rate-code NI’s feedback (recurrent) inputs and its leaky currents has to be precise in order to maintain the desired values (Cain et al., 2013; Kiani et al., 2013). In addition, a rate-code NI does not fully account for empirically observed ‘stepping’ spiking activity, which may underlie the dynamics of decision-making (Latimer et al., 2015).

In contrast, a location-code NI can maintain stable states in the absence of external inputs (Song and Wang, 2005). Further, neurophysiological studies have identified sequential activation similar to this propagation of bump activity (Beggs and Plenz, 2004; Harvey et al., 2012; Ikegaya et al., 2004; Rajan et al., 2016; Xu et al., 2012). Therefore, we constructed a computational model of cortical circuits to test the hypothesis that a location-code NI is a viable model for perceptual decision-making.

We found, as in previous location-code NIs, that our model sustains bump activity at a specific location in the absence of sensory input, whereas sensory input causes activity to propagate

through the network. Our model is unique in that it is based on depressing synapses and the interplay between two inhibitory neuron types that are commonly found in the cortex (Beierlein et al., 2003; Rudy et al., 2011). We also note that the sensory evidence, stored as the location of bump activity, can be read out in two different modes, depending on the connections to downstream readout neurons. When the connectivity is dense, readout neurons predominantly show classic ramping activity (i.e. linearly increasing activity over time) as the sensory evidence is accumulated into a decision variable (Gold and Shadlen, 2007; Goldman et al., 2009; Mazurek et al., 2003). In contrast, when the connectivity is sparse, readout neurons predominantly exhibit stepping activity (Latimer et al., 2015). This observation predicts that either ramping or stepping mode can emerge depending on the connectivity between the integrator and downstream readout neurons. This dual-readout mode may, in part, reconcile the degree to which components of decision-making are encoded as ramping- or stepping-like spiking activity.

Results

The goal of this study was to propose an alternative location-code NI and to test its potential links with perceptual decision-making. We first performed a bifurcation analysis to quantify the stability of a rate-code NI during a temporal gap in the sensory evidence. That is, we tested the ability of a rate-code NI to act as a lossless (perfect) integrator. Next, we propose an alternative location-code NI, which is different from earlier location rate-code NIs (Skaggs et al., 1995; Song and Wang, 2005; Stringer et al., 2002): our model relied on depressing excitatory synapses and the interplay between two types of inhibitory neurons that are commonly found in the cortex (Beierlein et al., 2003; Markram et al., 2004; Pfeffer et al., 2013; Rudy et al., 2011). Finally, we discuss how our proposed integrator can map onto two different modes of spiking activity that have been empirically observed during decision-making: the classic 'ramping' activity (Roitman and Shadlen, 2002) and newly identified 'stepping' activity (Latimer et al., 2015).

Recurrent networks are unstable without sensory evidence

In principle, rate-code NIs can be perfect lossless integrators when the recurrent inputs are precisely equivalent to the leak currents (Goldman et al., 2009). However, the dynamics of these rate-code NIs can become unstable when external sensory inputs are removed (Kiani et al., 2013). For instance, the amount of accumulated sensory evidence can decrease, due to leak current, during temporal gaps in the incoming sensory evidence. However, the accuracy of the perceptual decision does not degrade during such temporal gaps (Kiani et al., 2013; Liu et al., 2015). This discrepancy suggests that rate-code NIs may not fully account for sensory-evidence integration.

To quantify the degree to which rate-code NI can leak during temporal gaps in incoming sensory evidence, we developed a firing-rate model that describes a single recurrent network (Equation 1) similar to the one described previously (Goldman et al., 2009):

$$\tau_m \frac{dF_e}{dt} = -F_e + F_{max} \frac{1}{[1+e^{-\beta(rF_e+E-\theta)}]} \quad (1)$$

, where F_e and r are the firing rate and recurrent connections, respectively; F_{max} is the maximum firing rate; θ is the spiking threshold; E is the external input; and β represents the strength of stochastic inputs (Ermentrout and David, 2010). The first term in the right-hand side of Equation 1 represents the leak current; this type of firing rate equation corresponds to the subthreshold dynamics of leaky integrate-and-fire neurons (Miller and Fumarola, 2012). The selected default parameters are F_{max}=20, β=1, θ=0.5, r=1 and E=0 unless stated otherwise. We modeled the gain (transfer function; i.e., the number of spikes that a neuron can generate in response to afferent synaptic activity) with a logistic function because it best describes the response of neurons receiving stochastic inputs (Ermentrout and David, 2010); the firing rate of this neuron is not zero even when the synaptic inputs are smaller than the spike threshold.

Next, we tested the stability of this network by conducting a bifurcation analysis with the XPPAUT analysis platform (Ermentrout, 2007). A bifurcation analysis identifies the steady-state solutions, in which a system can stay indefinitely until perturbed. Moreover, this analysis clarifies whether the steady-state solutions are stable in response to the perturbations of bifurcation parameters (which, in our analysis, is the strength of the recurrent connections r and the external inputs E; see Fig. 1A). In Figs. 1B and C, the stable and unstable steady-state solutions are shown in red and black, respectively. As seen in these figures, this recurrent network (Equation 1) has only two stable steady states (attractor), in which neurons in the

network either fire at their maximum rate (F_{max}) or become quiescent. This means that a small perturbation in the strength of the recurrent connections or changes in the external inputs (e.g., a temporal gap in the incoming sensory information) could lead to a loss in the network of the temporally accumulated information (Kiani et al., 2013).

Cortical-circuit based location-code neural integrator

This finding opens the possibility that lossless integration may be mediated by a circuit other than a rate-code NI; in particular, by a location-code NI. The structure of this cortical integrator is inspired by the three broadly observed properties of cortical neurons. First, pyramidal (Pyr) neurons are topographically organized as a function of their sensory response profiles via spatial (Hubel and Wiesel, 1962, 1968) and functional (Ko et al., 2013) connections. Second, parvalbumin positive (PV) and somatostatin positive (SST) inhibitory interneurons are two of the most common types of inhibitory neuron (Rudy et al., 2011) in the neocortex; PV neurons have a fast-spiking pattern of activity, whereas SST neurons have a low-threshold spiking pattern. For our purposes, it is important to note that, although most inhibitory interneurons are broadly tuned to sensory inputs, the response profiles of SST neurons can be as sharply tuned as those of Pyr neurons (Ma et al., 2010). Third, SST neurons inhibit neighboring cortical neurons through lateral inhibition (Markram et al., 2004; Zhang et al., 2014). Below, we discuss the structure and properties of a discrete and continuous version of this integrator (Fig. 2) and how information from this integrator can be read out by downstream neurons.

Discrete location-code neural integrator

Our discrete integrator is based on the generic cortical-network structure (Ardid et al., 2007; Compte et al., 2000) but it includes two major inhibitory neuron types, PV and SST neurons (Beierlein et al., 2003; Rudy et al., 2011). In our model, Pyr neurons drive SST neurons locally

within a population but drive PV neurons globally (i.e., across different neuronal populations). As a consequence of this connectivity, SST neurons can have more sharply tuned response profiles than PV neurons (Ma et al., 2010).

Specifically, we constructed 17 neuronal populations, each consisting of 16 SST neurons and 400 Pyr neurons. We note that the size of our network model (i.e., 17 populations of Pyr and SST neurons) is rather arbitrary and is designed to demonstrate how lossless integration works at a reasonable time scale: if we were to incorporate additional populations, the network would integrate synaptic inputs over a longer period of time. Each pair of neighboring neuronal populations are connected via Pyr-Pyr connections. This connectivity is unidirectional to ensure that bump activity propagates in one direction only; such unidirectional connections have been employed in previous network models to specify the direction of signal propagation (Diesmann et al., 1999; Goldman, 2009). The recurrent connections between Pyr neurons form depressing synapses (see Table 1 and **Methods**). All of the other synapses are static including the Pyr-Pyr synapses that form across neighboring populations. Additionally, our model had two PV neuronal populations (PV_1 and PV_2) that interacted with each of these 17 neuronal populations. These two PV neuronal populations provide feedforward and feedback inhibition (Hu et al., 2014). Importantly, the SST neurons mediate specific lateral inhibition: that is, they can inhibit all of the Pyr neurons except those in their population (Markram et al., 2004; Zhang et al., 2014). Fig. 2A and B illustrate the model structure, and Tables 1 and 2 provide details of the neuron/synapse models and the connectivity patterns.

Each neuron received background inputs that were simulated as Poisson spike trains (Table 1). Afferent (i.e., ‘thalamic’) sensory inputs to Pyr and PV neurons were also modeled as Poisson spike trains. These sensory inputs modeled two (200-ms) bursts of sensory information (evidence)

that are separated by a 500-ms temporal gap in which there is not any sensory evidence. In this study, we assumed that transient (onset) sensory input (see **Methods**) can provide extra inputs to the first neuronal population (i.e., 'population 1') so that population 1 could always generate bump activity at the onset of stimulus. Due to the nonspecific feedback inhibition of PV_1 neurons, other populations can be kept quiescent at stimulus onset; see below. With this assumption, we exclusively introduced transient inputs to Pyr neurons in population 1, whereas sustained sensory inputs projected to all of Pyr and PV neurons during the entire duration of stimulation periods.

To establish a baseline of activity, we had 100 ms of spontaneous activity prior to the onset of the sensory input. Fig. 3A shows the activity of each of the Pyr neurons in each of the 17 neuronal populations during stimulus presentation. During each of the two stimulus periods (time=100-300 ms and 800-1000 ms, marked by red arrows), Pyr neurons were sequentially activated. However, and more importantly, Pyr neurons within the same population continues to fire even during the temporal gap (time=300-800 ms). That is, this integrator had two functional modes: integration and retention.

What is the basis for these two functional modes? First, the integration was based on the sequential activation of different Pyr neuronal populations (Fig. 3A). During the two stimulus periods (100-300 ms and 800-1000 ms, marked by red arrows), when one of Pyr neuronal populations was activated, the other Pyr neuronal populations were quiescent (Fig. 3A). This exclusive activation of a Pyr neuronal population depended on feedback inhibition, which was mediated by PV_1 neurons which nonspecifically projected to Pyr neurons. Because PV_1 neurons were more active during the stimulus period than during the temporal gap (Fig. 3B), their feedback inhibition was stronger during the stimulus period than during the temporal gap. Further, because the PV_1 neurons fired strongly and synchronously during the stimulus period

(supplemental Fig. 1A), the inhibition from these synchronously firing neurons can effectively hyperpolarize target neurons (Börgers et al., 2008; Fries, 2005). Finally, the depressing synapses opened a window of opportunity for Pyr neurons in a neighboring population to respond to lateral excitation by lowering the feedback inhibition. As seen in supplemental Fig. 1B, the peak population activity in Pyr_1 decreased over time due to depressing recurrent synapses, making PV_1 activity weaker over time. When bump activity jumped to population 2 from population 1 a time=~210 ms, PV_1 activity stayed at a low rate. Similarly, Pyr_2 activity became weaker, on average, over time; see the black line in supplemental Fig.1B. In this way, Pyr neuronal populations were sequentially activated during the stimulus period.

Second, the network mode switched into a retention mode when we removed the sensory input. In this retention mode, SST neurons became active (Fig. 3C). Unlike the global inhibition mediated by PV_1 neurons, SST neurons selectively inhibited Pyr neurons in other populations but did not inhibit Pyr neurons within the same population. Due to this connectivity pattern, when SST neurons in a population fired, Pyr neurons in the same population (as the SST neurons) received the least amount of inhibition. If the inhibition induced by active SST neurons was strong enough, the next population could not respond to the excitatory inputs and remained quiescent. In this way, the same population remained active during the temporal gap.

In contrast to their activity during temporal gaps in the sensory input, when there was afferent sensory input (i.e., time=100-300 ms), SST neurons were quiescent (Fig. 3C). They were quiescent because of the strong inhibition from the PV_2-neuronal population that provided feedforward inhibition (Fig. 3B); see also Table 2. However, when the sensory inputs were removed at 300 ms, this feedforward inhibition was eliminated, allowing SST neurons to respond to the excitatory input from Pyr neurons within its population. As a consequence, those SST

neurons that belonged to the population of active Pyr neurons at 300 ms became active and then, in turn, inhibited the Pyr neurons in other neuronal populations. In other words, Pyr neuronal populations, which were inactive at 300 ms, remained inactive. This network is virtually equivalent to a stable recurrent network and the same neuronal population remained continuously active even in the absence of stimulus input (Fig. 3).

What happens to this network when afferent sensory input is returned to the network? When sensory input was delivered at time=800 ms, population activity resumed propagating as SST neurons were, once again, inactivated by the feedforward inhibition from the PV_1 and PV_2 neurons. We note that population 1 was <u>not reactivated</u> even though it received transient afferent inputs (Fig. 3A). This can be explained by the feedback inhibition mediated by PV_1 neurons. Because population 5 was active, population 1 received strong feedback inhibition induced by population 5 and did not fire in response to transient stimulus inputs. To confirm the role of feedback inhibition, we removed the inhibition from PV_1 to Pyr_1 neurons during the temporal gap and found that transient afferent inputs <u>reactivated</u> population 1 (Supplemental Fig. 2). That is, nonspecific feedback inhibition was necessary for an accurate encoding of the accumulated sensory information when that afferent information included temporal gaps. We instantiated 4 equivalent networks and ran simulations on each of these networks. All four simulations produced behaviors (Supplemental Fig. 3) consistent with that shown in Fig. 3A.

Finally, bump activity did not propagate (i.e., it stayed at neuronal population 1) when we replaced the recurrent depressing synapses with static synapses (Fig. 3D). This finding suggests that depressing excitatory synapses were fundamental to the propagation of bump activity through the network. See supplemental Fig. 4 for more examples with static synapses.

To better understand how bump activity propagated during stimulus presentation and was stable during the delay period, we created a reduced version of this discrete neural integrator, which allowed us to analyze and quantify the dynamics of the network. This reduced network contained two neuronal populations and generated action potentials using the following firing-rate models (Equation 2).

$$\tau_m \frac{dF_1}{dt} = -F_1 + F_{max} \frac{1}{\left[1+e^{-\beta(rF_1 - r_m F_2 + E_1 - \theta)}\right]}$$

$$\tau_m \frac{dF_2}{dt} = -F_2 + F_{max} \frac{1}{\left[1+e^{-\beta(rF_2 - r_m F_1 + E_2 - \theta)}\right]} \quad (2)$$

The two populations had their own recurrent connections (r) and interacted with each other via lateral connections (r_m); see Fig. 4A. This mutual inhibition models the lateral inhibition mediated by SST and PV neurons in the computational model. We assumed that population 1 fired at the maximum rate, and population 2 was quiescent; that is, in the initial condition, population 1 retained bump activity.

With this reduced network, we studied the response of this network to perturbations in (1) the recurrent connections within a population (r), (2) the external inputs (E_1, E_2) to populations 1 and 2, or (3) lateral interactions (r_m) between the two populations. Three main findings emerged from this analysis and can be seen in Fig. 4. First, as the recurrent-connection stregnth (r) increased, the network remained stable (Fig. 4B). Second, the network remained stable as we increased E_1 (i.e., the external input to population 1; the red lines in Fig. 4C) but became unstable (i.e., population 1 lost its bump activity) when E_1 was reduced (black lines in Fig. 4C). On the other hand, as shown in Fig. 4D, the network became unstable when E_2 (i.e., the external input to population 2) increased, but it became stable when E_2 was decreased. In other words, the noise introduced into quiescent populations needed to be regulated for reliable retention of

information. Third, the lateral interactions (r_m) strongly impacted the stability of the network. When r_m was postive but small (i.e., weak mutual inhibition), the network became unstable (the black lines in Fig. 4E). In contrast, when r_m was postive and big (i.e., strong mutual inhibition), population 1 reliably retained bump activity (Fig. 4E), and population 2 remained quiescent (Fig. 4F). That is, as long as population 1 retained bump activity intially, the mutual inhibition helped population 1 keep the bump activity. When the two populations excited each other (i.e., negative r_m), neurons in both populations fired at the maximum rate (Figs. 4E and F). In this case, bump activity was not confined to population 1, indicating that a read-out of bump activity based on location was not an accurate reflection of the accumulated evidence.

Overall, these simulations suggest that recurrent connections within a neuronal population and mutual inhibition are key components of the network and are responsible for the retention of bump activity during a temporal gap in the incoming sensory information. Conversely, the propagation of bump activity required a reduction in the strength of the recurrent connections within a population or the strength of the mutual inhibition. In our integrator model, we found such reductionships: during stimulus presentation, specific lateral inhibition was reduced due to the silencing of the SST neurons (Fig. 3C). The nonspecific feedback lateral inhibition of PV_1 neurons ensured that only a single Pyr cell population was active. Also, recurrent inputs within a population decreased over time due to depressing recurrent connections; this can account for our finding that bump activity did not propagate with static recurrent connections (Fig. 3D).

Continuous location-code neural integrator

Because this discrete network model has only a limited number of attractors (i.e., the number of populations that can hold accumulated sensory inputs), it necessarily lowers the precision of the stimulus integration. However, this discrete network can be generalized to have continuous

attractor states, as in linear attractor networks. Unlike the discrete integrator that consists of defined neuronal populations, in our continuous integrator, Pyr and SST neurons were distributed into circular lattices with uniquely assigned coordinates (Fig. 2C). For convenience, we refer to the direction from lower to higher coordinates as the clockwise direction and higher to lower as counterclockwise. Two Pyr neurons were connected with each other if the difference between their coordinates was ≤200. Because the connections are symmetrical, each Pyr neuron made excitatory synapses with 400 neighbors. These recurrent chain-like connections provided effective recurrent excitation to elevate Pyr activity enough to generate a highly localized 'bump' of elevated spiking activity. Further, because these connections between Pyr neurons are symmetrical in both directions, bump activity could propagate in both directions. To make bump activity propagate only in the clockwise direction, we added a second set of SST neurons (SST_1 and SST_2).

Except for the additional population of SST neurons (the SST_2 neurons), the continuous and discrete integrators have equivalent structures; the structure of the continuous integrator is summarized in Fig. 2C and Table 3. All of the Pyr and SST neurons formed non-specific connections with PV_1 neurons, and PV_2 neurons exclusively provided feedforward inhibition to Pyr and SST_1 neurons. The connections between Pyr neurons and SST neurons were formed based on their coordinates in the circular lattice. (1) Pyr neurons made one-to-one synaptic connections with SST_1 and SST_2 neurons, when they had the same coordinates; a connectional scheme that we refer to as 'topographical' hereafter. (2) A SST_1 neuron inhibited a Pyr neuron when the (absolute) difference between their coordinates was ≥200. (3) A SST_2 neuron inhibited a Pyr neuron when the coordinate of a Pyr neuron was lower than that of a SST_2 neuron and when the (absolute) difference was between 400 and 800. Because of this connectivity pattern,

the propagation of bump activity in the counter-clockwise direction was dampened, and only bump activity in the clockwise direction propagated through the network.

In our first analysis, we examined whether our continuous integrator could integrate sensory information. Transient sensory inputs activated only the first 400 Pyr neurons (i.e., those with the lowest arbitrary coordinates) during time=100-200 ms, and then sustained sensory input was introduced to all of the Pyr and PV neurons during time=100-1000 ms. As seen in Fig. 5A, spontaneous spiking activity in the Pyr neurons at time=50 ms was not strong enough to induce bump activity. However, the transient sensory inputs that targeted the first 400 Pyr neurons elevated the rate of spiking activity strongly enough to generate bump activity. Once generated, the feedback inhibition mediated by the PV_1 neurons was strong enough to prevent all of the other excitatory neurons from spiking during the presentation of this transient sensory input. That is, the onset of sensory information ensured that bump activity was generated at the designated place (the first 400 excitatory neurons) and that all other activity was quiescent.

After the offset of this transient input, bump activity propagated to other Pyr neurons in the clockwise direction (Fig. 5A). Due to the periodic boundary condition, bump activity repeatedly circulated the integrator. Concurrently, PV_1 and PV_2 neurons fired asynchronously and sparsely (Fig. 5B). SST_1 neurons were quiescent (Fig. 5C), but SST_2 neurons, which received excitation from Pyr via topographic connections, mimicked Pyr activity (Fig. 5D). This SST_2 activity prevented bump activity from propagating in the counterclockwise direction via asymmetrical feedback inhibition onto Pyr neurons. Because the propagation speed was constant, the location of bump activity was proportional to the duration of stimulus, when the strength of the sensory input was constant. As in the discrete integrator, when we replaced all the depressing synapses

with the static ones, bump activity stays in the same location instead of travelling as shown in Figs. 5E and F.

Next, we tested whether this network could perform lossless integration. As in the discrete neural integrator, stimulus inputs were introduced at time=100 and 300 ms and time=800-1000 ms. For simplicity, we did not consider the onset input at 800 ms because this input had no impact on network dynamics in the discrete integrator (Fig. 3A). As seen in Fig. 6A, bump activity cascaded through the network until there was a temporal gap in the stimulus inputs. Once we removed these stimulus inputs, bump activity remained in the same location. Then, it resumed moving from the previous location, as sensory inputs were reintroduced, consistent with lossless integration.

As in the discrete integrator, during the temporal gap in sensory input, the PV_1 and PV_2 neurons (Fig. 6B) became quiescent. As a result, the inhibition from the PV_1 and PV_2 neurons to the SST_1 neurons was reduced, which, thereby, increased SST_1 activity (Fig. 6C). The firing pattern of SST_2 neurons was comparable to that of the Pyr neurons (Fig. 6D). Because the SST_1 neurons were topographically connected to Pyr neurons, the SST_1 inhibited non-active Pyr neurons, which prevented bump activity from propagating to a new location. Together, this transforms the network into an effective attractor network, like the case for the discrete integrator (Fig. 3).

Finally, how sensitive was our model to the strength of the stimulus inputs (i.e., the amount of sensory evidence)? Neurophysiological experiments have clearly shown that the rate of accumulation of the sensory evidence is positively correlated with the strength of the stimulus inputs. Further, this rate of accumulation is accompanied by a decrease in reaction time (Gold and Shadlen, 2007). We asked if this continuous integrator could account for the correlation between reaction time and stimulus inputs. Specifically, we estimated how rapidly bump activity

propagated as a function of the strength of the stimulus input (i.e., the firing rate of stimulus input). Indeed, as shown in Fig. 6E, as we increased the strength of the stimulus inputs, the propagation speed of bump activity also increased, supporting that our integrator can, in part, explain the correlation between reaction time and the strength of sensory evidence.

Potential links to decision-making

Two different types of NIs have been proposed as models of neural integrators. Rate-code NIs have been proposed to account for ramping activity in LIP, whereas location-code NIs account for head-direction neurons or path integration for place cells (Goldman et al., 2009). However, our simulation results raise the possibility that cortical circuit-based location-code NIs can also explain the lossless integration of afferent inputs necessary for perceptual decision-making (Kiani et al., 2013; Liu et al., 2015). That is, our newly proposed integrator may also underlie LIP activity during decision-making. Our integrator did not directly replicate either ramping (Roitman and Shadlen, 2002) or stepping activity (Latimer et al., 2015), which were identified in LIP. Thus, we asked if this activity was encoded at the population level and could be readout via downstream target neurons.

To address this question, we created a downstream network that readout information from our integrator. Specifically, Pyr neurons in the continuous integrator were connected to a population of 5000 downstream neurons with coordinate-dependent probability p (Fig. 7A):

$$p = \frac{c}{4000} p_0, \tag{3}$$

where c is the coordinate of pre-synaptic Pyr neurons ranging from 1 to 4000. That is, the Pyr integrator neurons with the highest coordinate projected to those readout neurons with the highest connection probability p_0, whereas the connection probability between the first Pyr

neuron to the readout neurons was $p_0/4000$. The connectivity pattern between the integrator and these readout neurons is consistent with previous experimental findings (Perin et al., 2011), which demonstrated that the synaptic-connection probability depends on the distance between presynaptic and postsynaptic neurons: the probability decreased, as the spatial distance increased. In other words, we assumed that the distance between integrator neurons and readout neurons decreased as the neurons' coordinates increased. (Fig. 7A).

In our simulations, we tested two different levels of maximal connection probabilities: p_0=0.1 and p_0=1.0. Independent of the connection probability, we found that population activity increased (ramped) over time (Figs. 7B and C). The maximum activity in the readout neurons occurred when bump activity reached the last Pyr neurons in the integrator. This can be explained by the fact that, as the bump moved to higher coordinate Pyr integrator neurons, the number of depolarized readout neurons also increased, due to the connectivity pattern between the integrator and the readout neurons (Equation 3). For instance, when p_0=1, the first Pyr integrator neuron depolarized, on average, a single readout neuron but the last Pyr neuron depolarized all of the downstream neurons. Consequently, as the sensory stimulus unfolded over time, more and more readout neurons fired, and population activity ramped up**,** consistent with ramping LIP activity (Gold and Shadlen, 2007; Roitman and Shadlen, 2002)

We also probed the behavior of individual readout neurons. Fig. 7D and E show the firing-rate time courses of 5 randomly chosen readout neurons. Although population activity ramped up in both low and high connection-probability states (i.e., p_0=0.1 and p_0=1.0), individual neurons showed either ramping or transient activity depending on p_0. When p_0 was high, individual neuronal activity appeared to ramp up like population activity (Fig 7D). In contrast, when p_0 was low, individual downstream neuronal activity did not ramp up although the population activity

did exhibit ramping activity (Fig. 7E). More specifically, individual readout neurons had either active or non-active periods. This activity pattern is reminiscent with the stepping model of LIP activity (Latimer et al., 2015), in which the firing rates of individual neurons jump to a discrete value.

If individual neurons showed ramping activity, as in Fig. 7D, the firing rate should be positively correlated time. In contrast, the correlation will be negligible if individual cell activity does not increase smoothly over time. Thus, we further tested individual neuronal activity by regressing the firing rates of individual neurons with time (between 100 and 400 ms). As seen in Fig. 8A, the slopes and correlation coefficients were higher and the *p*-values (of the correlation) were lower when p=1.0 than when p=0.1. This result is consistent with the idea that, on average, individual neurons had more linear (ramping) activity when connections between the integrator and readout neurons were dense. Interestingly, when we looked at the distribution of weakly connected (p=0.1) downstream neurons in more detail, we found that ~1000/5000 neurons had significant slopes (Fig. 8B), suggesting that these neurons' firing rates increased over time. Thus, our model shows that stepping and ramping activity may not be isolated from each other. Instead, they form a continuum depending on the connections between the integrator and readout neurons.

Discussion

Perceptual decision-making has been thought to rely on the accumulation of sensory evidence (i.e., decision-variables) extracted from ambiguous sensory stimuli (LaBerge, 1962; Mazurek et al., 2003; Miller, 2015; Ratcliff, 1978; Ratcliff and Smith, 2004; Roitman and Shadlen, 2002; Smith and Ratcliff, 2004). For instance, the drift-diffusion model (DDM) assumes that a decision is initiated when accumulated evidence reaches a threshold for decision-making. It is generally thought that perceptual decision-making is instantiated through rate-code neural integrators (NIs), which are based on recurrent networks (Goldman et al., 2009; Wang, 2012). However, the degree to which rate-code NIs can explain perceptual decision-making is limited. For example, rate-code NIs become unstable when there is a temporal gap in the flow of incoming sensory evidence (Fig. 1), whereas behavioral studies indicate that participants act as 'perfect' integrators and are not affected by these temporal gaps (Kiani et al., 2013; Liu et al., 2015).

How then can the brain make reliable decisions even with temporal gaps? Our model predicts that the cortex can natively integrate sensory evidence and maintain accumulated evidence during temporal gaps. In our integrator, the location of bump activity represents the amount of presented sensory evidence; that is, our integrator shares the characteristics of location-code NIs previously proposed to account for head-direction neurons (Skaggs et al., 1995; Song and Wang, 2005); see below. In our simulations, bump activity in the integrator progressed through the network when sensory inputs were provided, but stayed at the same location when the sensory inputs were removed; the location of the bump is stable due to the inhibition of SST cells (Fig. 4). This indicates that our integrator can account for the robustness of perceptual decision-making to temporal gaps.

Comparison to other location code NIs

In terms of function, our model produces results equivalent to those of previously reported location-code NIs, which modelled head-direction neurons encoding the direction of an animal's head relative to its body and independent of its location in the environment. However, the underlying mechanisms are distinct.

In these earlier location-code NIs, the shift in the location of bump activity was realized by so-called "rotation" neurons, which employed either strictly excitatory neurons (Skaggs et al., 1995) or strictly inhibitory neurons (Song and Wang, 2005); these rotation neurons are located in the thalamic nucleus that receives inputs from the vestibular system. In contrast, our model showed that a cortical circuit, which consisted of excitatory pyramidal neurons and different types of inhibitory interneurons, can readily implement a location-code NI. More specifically, the two major inhibitory neuron types in the neocortex (Rudy et al., 2011) –PV and SST interneurons– make distinct contributions to this operation. PV neurons, which provide nonspecific feedback inhibition to pyramidal neurons (Bock et al., 2011; Ma et al., 2010), ensure that bump activity exists exclusively at a single location. On the other hand, SST neurons mediating lateral inhibition transformed the network into an effective attractor network capable of maintaining accumulated sensory evidence during the temporal gaps (Figs. 3C and 6C). We note that the sustained activity of SST neurons during temporal gaps is consistent with the empirical finding that SST cells are selectively activated during a delay period when a stimulus is removed and an animal needs to remember task-relevant information (Kim et al., 2016). In contrast to the role that interneurons and their inhibitory synapses played in our network model, depressing excitatory synapses made bump activity propagate through the network (Figs. 3D and 5E).

Together, our simulation results suggest that neurons and synapses in the neocortex are indeed suitable for controlling and maintaining the propagation of bump activity.

Links to sequential sampling models of decision-making

Sequential sampling models have been widely used to study decision-making processes because of their ability to account for the trade-off between decision-accuracy and reaction time (Ratcliff and Smith, 2004). The two main classes of sequential sampling models are diffusion and accumulator models (Ratcliff and Mckoon, 2008; Smith and Ratcliff, 2004). Although they share the same fundamental principle that evidence is accumulated for decision-making about ambiguous sensory information, they differ in the exact mechanisms (Ratcliff and Smith, 2004). Then, can our integrator explain the two main classes of sequential sampling models?

First, the accumulator model (a.k.a. race model) assumes that each piece of evidence supporting a possible choice is independently integrated and that a decision can be made if accumulated evidence hits the decision-bound (Miller, 2015; Ratcliff and Smith, 2004). This model can be implemented by using multiple location-code NIs, which independently integrate evidence. The decision-bound can be realized by assigning threshold neurons that project afferent inputs to readout neurons. If these connections between threshold neurons and readout neurons are strong enough, readout neurons will fire whenever bump activity arrives at the threshold neurons; that is, when the decision-bound crossing is detected.

Second, in the DDM, the amount of accumulated evidence for one choice either increases or decreases over time depending on fluctuations in presented evidence (Ratcliff and Mckoon, 2008). In our integrator, readout neuron activity can increase or decrease if readout neurons receive afferent inputs from two integrators. If one integrator provides feedforward excitation

and another provides inhibition onto the same readout neurons, readout-neuron activity either increases or decreases depending on which integrator provides stronger afferent inputs; the same scheme was indeed used in an earlier model (Mazurek et al., 2003). Alternatively, if the two integrators exclusively provide excitatory afferent inputs to two distinct sets of readout neurons, the mutual inhibition between the two sets of readout neurons can also replicate the fluctuation of accumulated evidence in DDM, as in an earlier model (Usher and McClelland, 2001).

In principle, the combination of readout neurons and location-code integrators can account for the most critical features of sequential sampling models. That is, our model does provide alternative potential mechanisms underlying sequential sampling models; furthermore, our model raises the possibility that sequential sampling models can be robust to temporal gaps in evidence.

Selection between stepping and ramping modes

Our model predicts that depending on the density of connections between the integrator and readout neurons, ramping or stepping activity will predominate (Figs. 7 and 8), mirroring the two types of neural activity seen in area LIP (Latimer et al., 2015; Roitman and Shadlen, 2002). More specifically, even with sparse connections (p=0.1), 80% of neurons show stepping activity, whereas 20% of neurons show ramping activity (Fig 8), suggesting that ramping and stepping activity likely coexist and that only the proportion of stepping/ramping neurons varies depending on task demands.

Importantly, an earlier empirical study (Purushothaman and Bradley, 2005) found evidence supporting the possibility that synaptic connections can be changed flexibly depending on the task. When animals performed fine-discrimination tasks, decisions relied on the highest precision neurons that carried the most relevant information in middle temporal areas (MT); that is,

decisions were based on a selective pooling of MT neurons. In contrast, low precision neurons' contributions were identified when animals performed coarse-discrimination tasks; that is, decisions were based on a broad and nonselective pooling of MT neurons. Indeed, the selective and the broad pooling are equivalent to sparse and dense connections in our model in terms of the number of connections to readout neurons. Therefore, we hypothesize that the brain can flexibly select between sparse and dense connections depending on behavioral demands and that such distinctively selected connectivity patterns may underlie different observations of LIP neuron activity during perceptual decision-making. Consistent with this hypothesis, stepping activity was observed from animals trained to report decisions after watching the entire stimulus presentation (i.e., the fixed viewing task), whereas ramping activity was observed from animals allowed to report decisions during the stimulus presentation (i.e., reaction-time task) (Latimer et al., 2015; Roitman and Shadlen, 2002; Shadlen et al., 2016).

Moreover, we note that sparse and dense connections make readout neurons behave distinctively, and thus they may be suitable for either fixed-viewing or reaction-time tasks. In fixed viewing tasks, animals watch the entire presentation before reporting decisions. That is, the integration is completed before decision-making, and thus animals' decisions can rely on the decision boundary-crossing (i.e., the arrival of the bump at the threshold neurons). Sparse connections would be optimal for the detection of boundary-crossing since they can reduce the wiring cost and minimize the false-alarm rate of readout neurons by reducing noise signals (afferent inputs independent of detection of threshold-crossing) to readout neurons. In reaction-time tasks, animals may need to use a relative threshold (a difference between evidences) to make decisions rapidly; if evidence difference is big enough, the choice would be apparent to them right away. However, with sparse connections, readout neurons become agnostic about the location of bump

activity before the bump arrives at threshold neurons. That is, sparse connections may be suboptimal in reaction time tasks. On the other hand, dense connections, which allow readout neurons to track the location of bump activity, should be chosen for decision-making using relative thresholds that may be necessary for reaction-time tasks.

Empirical evidence for our location-code NI and concluding remarks

Sequential activation, consistent with bump activity propagation in our model, has been observed in multiple brain regions (Pulvermuller and Shtyrov, 2009; Tang et al., 2008) including the visual cortex (Beggs and Plenz, 2003, 2004; Ikegaya et al., 2004; Sato et al., 2012; Xu et al., 2012), parietal cortex (Harvey et al., 2012) and frontal cortex (Seidemann et al., 1996). Although synfire chain (Diesmann et al., 1999) was proposed to account for such phenomena, most theoretical studies have been dedicated to studying the neural correlates of persistent activity observed in the prefrontal cortex due to its potential links to working memory and decision-making (Goldman et al., 2009; Wang, 2012). Recent theoretical studies raised the possibility that the sequential activation of neurons could also be the substrate of working memory (Goldman, 2009; Lundqvist et al., 2016; Rajan et al., 2016), reigniting interest in the mechanisms underlying sequential activation.

Although determining the exact mechanisms behind any cognitive functions remains difficult, we would like to underscore that our model showed that cortical circuits can natively switch between two seemingly distinct states, stable steady state (e.g., bump activity maintenance) and sequential activation state (e.g., bump activity propagation). In the model, the two major inhibitory cell types make this switching of states possible. Because these two cell types are major inhibitory neuron types in the neocortex (Rudy et al., 2011) and bump activity propagation exists in a wide range of brain regions, we speculate that the switching between the two states

may be one of the fundamental computing principles in the brain and plan to extend our current model to address this possibility.

Author Contributions

JHL, JT, SV and YEC designed research; JHL performed research and analyzed data; JHL ,JT, SV and YEC wrote the paper.

Acknowledgements

JHL wishes to thank the Allen Institute founders, Paul G. Allen and Jody Allen, for their vision, encouragement and support. YEC was support by funding from the NIDCD-NIH and Boucai Hearing Restoration Fund. We also want to thank Heather Hersh and Joshua Gold for helpful comments

References

Ardid, S., Wang, X.-J., and Compte, A. (2007). An integrated microcircuit model of attentional processing in the neocortex. J. Neurosci. *27*, 8486–8495.

Beggs, J.M., and Plenz, D. (2003). Neuronal avalanches in neocortical circuits. J. Neurosci. *23*, 11167–11177.

Beggs, J.M., and Plenz, D. (2004). Neuronal avalanches are diverse and precise activity patterns that are stable for many hours in cortical slice cultures. J. Neurosci. *24*, 5216–5229.

Beierlein, M., Gibson, J.R., and Connors, B.W. (2003). Two dynamically distinct inhibitory networks in layer 4 of the neocortex. J. Neurophysiol. *90*, 2987–3000.

Bock, D.D., Lee, W.-C.A., Kerlin, A.M., Andermann, M.L., Hood, G., Wetzel, A.W., Yurgenson, S., Soucy, E.R., Kim, H.S., and Reid, R.C. (2011). Network anatomy and in vivo physiology of visual cortical neurons. Nature *471*, 177–182.

Börgers, C., Epstein, S., and Kopell, N.J. (2008). Gamma oscillations mediate stimulus competition and attentional selection in a cortical network model. Proc. Natl. Acad. Sci. U. S. A. *105*, 18023–18028.

Cain, N., Barreiro, A.K., Shadlen, M.N., and Shea-Brown, E. (2013). Neural integrators for decision making: a favorable tradeoff between robustness and sensitivity. J. Neurophysiol. *109*, 2542–2559.

Compte, A., Brunel, N., Brunel, N., Goldman-Rakic, P.S., Goldman-Rakic, P.S., Wang, X.J., and Wang, X.J. (2000). Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. Cereb. Cortex *10*, 910–923.

Diesmann, M., Gewaltig, M.O., and Aertsen, a (1999). Stable propagation of synchronous spiking in cortical neural networks. Nature *402*, 529–533.

Ding, L., and Gold, J.I. (2012). Neural correlates of perceptual decision making before, during, and after decision commitment in monkey frontal eye field. Cereb. Cortex *22*, 1052–1067.

Ermentrout, B. (2007). XPPAUT. Scholarpedia *2*, 1399.

Ermentrout, G.B., and David, H.T. (2010). Mathematical Foundation of Neuroscience (springer).

Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. Trends Cogn. Sci. *9*, 474–480.

Gold, J., and Shadlen, M. (2007). The neural basis of decision making. Annu. Rev. Neurosci. *30*, 535–574.

Goldman, M.S. (2009). Memory without Feedback in a Neural Network. Neuron *61*, 621–634.

Goldman, M.S., Compte, A., and Wang, X.-J. (2009). Neural Integrator Models. Encycl. Neurosci. *6*, 165–178.

Harvey, C.D., Coen, P., and Tank, D.W. (2012). Choice-specific sequences in parietal cortex during a virtual-navigation decision task. Nature *484*, 62–68.

Hu, H., Gan, J., and Jonas, P. (2014). Interneurons. Fast-spiking, parvalbumin^{+} GABAergic interneurons: from cellular design to microcircuit function. Science *345*, 1255263.

Hubel, D.H., and Wiesel, T.N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J. Physiol. *160*, 106–154.2.

Hubel, D.H., and Wiesel, T.N. (1968). Receptive fields and functional architecture of monkey

striate cortex. J.Physiol.(London) *195*, 215–243.

Ikegaya, Y., Aaron, G., Cossart, R., Aronov, D., Lampl, I., Ferster, D., and Yuste, R. (2004). Synfire Chains and Cortical Songs: Temporal Modules of Cortical Activity. Sci. (New York, NY) *559*.

Kiani, R., Churchland, A.K., and Shadlen, M.N. (2013). Integration of direction cues is invariant to the temporal gap between them. J. Neurosci. *33*, 16483–16489.

Kim, J.N., and Shadlen, M.N. (1999). Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. Nat. Neurosci. *2*, 176–185.

Kim, D., Jeong, H., Lee, J., Ghim, J.W., Her, E.S., Lee, S.H., and Jung, M.W. (2016). Distinct Roles of Parvalbumin- and Somatostatin-Expressing Interneurons in Working Memory. Neuron *92*, 902–915.

Ko, H., Cossell, L., Baragli, C., Antolik, J., Clopath, C., Hofer, S.B., and Mrsic-Flogel, T.D. (2013). The emergence of functional microcircuits in visual cortex. Nature *496*, 96–100.

LaBerge, D. (1962). A recruitment theory of simple behavio. Psychometrika *27*, 375–396.

Latimer, K.W., Yates, J.L., Meister, M.L.R., Huk, A.C., Pillow, J.W., Shadlen, M.N., Kiani, R., Newsome, W.T., Gold, J.I., M., W.D., et al. (2015). Single-trial spike trains in parietal cortex reveal discrete steps during decision-making. Science *349*, 184–187.

Liu, A.S.K., Tsunada, J., Gold, J.I., and Cohen, Y.E. (2015). Temporal Integration of Auditory Information Is Invariant to Temporal Grouping Cues 1 , 2 , 3. eNeuro *2*.

Lundqvist, M., Rose, J., Herman, P., Brincat, S.L.L., Buschman, T.J.J., and Miller, E.K.K. (2016). Gamma and Beta Bursts Underlie Working Memory. Neuron *90*, 152–164.

Ma, W. -p., Liu, B. -h., Li, Y. -t., Josh Huang, Z., Zhang, L.I., and Tao, H.W. (2010). Visual Representations by Cortical Somatostatin Inhibitory Neurons--Selective But with Weak and Delayed Responses. J. Neurosci. *30*, 14371–14379.

Markram, H., Toledo-Rodriguez, M., Wang, Y., Gupta, A., Silberberg, G., and Wu, C. (2004). Interneurons of the neocortical inhibitory system. Nat. Rev. Neurosci. *5*, 793–807.

Mazurek, M.E., Roitman, J.D., Ditterich, J., and Shadlen, M.N. (2003). A Role for Neural Integrators in Perceptual Decision Making. Cereb. Cortex *13*, 1257–1269.

Miller, P. (2015). Decision Making Models (New York: Springer-Verlag New York).

Miller, K.D., and Fumarola, F. (2012). Mathematical Equivalence of Two Common Forms of Firing Rate Models of Neural Networks. Neural Comput. *24*, 25–31.

Perin, R., Berger, T.K., and Markram, H. (2011). A synaptic organizing principle for cortical neuronal groups. Proc. Natl. Acad. Sci. U. S. A. *108*, 5419–5424.

Pfeffer, C.K., Xue, M., He, M., Huang, Z.J., and Scanziani, M. (2013). Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. Nat. Neurosci. *16*, 1068–1076.

Pulvermuller, F., and Shtyrov, Y. (2009). Spatiotemporal signatures of large-scale synfire chains for speech processing as revealed by MEG. Cereb. Cortex *19*, 79–88.

Purushothaman, G., and Bradley, D.C. (2005). Neural population code for fine perceptual decisions in area MT. Nat. Neurosci. *8*, 99–106.

Rajan, K., Harvey, C.D., and Tank, D.W. (2016). Recurrent Network Models of Sequence Generation and Memory. Neuron *90*, 128–142.

Ratcliff, R. (1978). A theory of memory retrieval. Psychol. Rev. *85*, 59–108.

Ratcliff, R., and Mckoon, G. (2008). The Diffusion Decision Model: Theory and Data fpr Two-Choice Decision Tasks. Neural Comput. *20*, 873–922.

Ratcliff, R., and Smith, P.L. (2004). A Comparison of Sequential Sampling Models for Two-Choice Reaction Time. Psychol Rev. *111*, 333–367.

Roitman, J.D., and Shadlen, M.N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. J. Neurosci. *22*, 9475–9489.

Rudy, B., Fishell, G., Lee, S., and Hjerling-Leffler, J. (2011). Three groups of interneurons account for nearly 100% of neocortical GABAergic neurons. Dev. Neurobiol. *71*, 45–61.

Sato, T.K., Nauhaus, I., and Carandini, M. (2012). Traveling Waves in Visual Cortex. Neuron *75*, 218–229.

Seidemann, E., Meilijson, I., Abeles, M., Bergman, H., and Vaadia, E. (1996). Simultaneously recorded single units in the frontal cortex go through sequences of discrete and stable states in monkeys performing a delayed localization task. J. Neurosci. *16*, 752–768.

Shadlen, M.N., Kiani, R., Newsome, W.. T., Gold, J.I., Wolpert, D.M., Zylberberg, A., Ditterich, J., Lafuente, V., Yang, T., and Roitman, J.D. (2016). Comment on “Single-trial spike trains in parietal cortex reveal discrete steps during decision-making.” Science (80-.). *351*, 1406–1406.

Skaggs, W.E., Knierim, J.J., Kudrimoti, H.S., and McNaughton, B.L. (1995). A model of the neural basis of the rat’s sense of direction. Adv. Neural Inf. Process. Syst. *7*, 173–180.

Smith, P.L., and Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. Trends Neurosci. *27*, 161–168.

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.

Stringer, S.M., Trappenberg, T.P., Rolls, E.T., and de Araujo, I.E.T. (2002). Self-organizing continuous attractor networks and path integration: one-dimensional models of head direction cells. Network *13*, 217–242.

Tang, A., Jackson, D., Hobbs, J., Chen, W., Smith, J.L., Patel, H., Prieto, A., Petrusca, D., Grivich, M.I., Sher, A., et al. (2008). A Maximum Entropy Model Applied to Spatial and Temporal Correlations from Cortical Networks In Vitro. J. Neurosci. *28*, 505–518.

Usher, M., and McClelland, J.L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. Psychol. Rev. *108*, 550–592.

Wang, X.J. (2012). Neural dynamics and circuit mechanisms of decision-making. Curr. Opin. Neurobiol. 1–8.

Xu, S., Jiang, W., Poo, M.-M., and Dan, Y. (2012). Activity recall in a visual cortical ensemble. Nat. Neurosci. *15*, 449–455, S1-2.

Zhang, S., Xu, M., Kamigaki, T., Hoang Do, J.P., Chang, W.-C., Jenvay, S., Miyamichi, K., Luo, L., and Dan, Y. (2014). Long-range and local circuits for top-down modulation of visual cortex processing. Science (80-.). *345*, 660–665.

Figure Legends

Figure 1:The stability of a rate-code NI. (A), Structure of a rate-code NI with recurrent connections (r) and external input (*E*). The bifurcation analyses with *r* and *E* as bifurcaions parameters are shown in **(B)** and **(C)**, respectively.

Figure 2: The structure of the two versions of our integrator. (A), Connectivity between all 19 neuronal populations in the *discrete* integrator. **(B)**, interconnectivity between the 17 Pyr-SST populations; see **Methods** and Tables 1 and 2 for more details and parameters. The red arrows and blue circle-head arrows represent the excitatory and inhibitory connections within the network model, respectively. The thick and dashed black arrows represent onset and sustained stimulus inputs, respectively. **(C)**, Structure of *continuous* integrator. The five neuronal populations (Pyr, PV_1, PV_2, SST_1, and SST_2) interact with each other via connections shown in the figure. The red arrows and blue circle-head arrows represent the excitatory and inhibitory connections between individual neurons, respectively. In contrast, the double lined arrows (including red and blue) show connections between the populations. All connections between populations are randomly established. Sensory inputs are introduced to Pyr, PV_1 and PV_2 (dashed arrows). Periodic boundary condition is used to connect Pyr cells, as shown in the green arrow; see **Methods** and Table 3 for more details and parameters.

Figure 3: The responses of populations of the discrete integrator. (A), Spiking activity of Pyr neurons in all 17 populations; each population had 400 Pyr neurons. Each row in the plot shows the spike times of an individual Pyr neuron. 17 populations are shown in different colors; for clarity, we show color codes corresponding to a few populations in the inset. The red and black arrows show sensory-stimulus periods and the temporal gap between them, respectively. **(B)**, PV_1 and PV_2 activity, which indicates that PV activity is modulated by the sensory inputs. Both

PV populations contained 1088 PV neurons. **(C)**, SST neuron activity in all 17 populations; there are 16 SST neurons in each population. The same color scheme is used as in (A), and during the temporal gap, active SST and Pyr neurons have the same color, indicating that active SST and Pyr neurons belong to the same population. We also repeated the same simulation by replacing all depressing synapses with static ones. Pyr cell activity with static recurrent connections is shown in Panel **(D)**.

Figure 4: The bifurcation analysis of the discrete integrator with a firing rate model. We studied the interactions between two populations shown in **(A)**, using the firing rate model. Red and black lines represent stable and unstable steady solutions, respectively. **(B)-(D)** Firing rate of population 1 with respect to recurrent connections within a population (r) and external input to populations 1 and 2 (E_1, E_2), respectively. **(E)** and **(F)**, Fring rate of populations 1 and 2, respectively, in terms of mutual interactions (r_m).

Figure 5: The propagation of the bump activity during the stimulus period. (A)-(D), Spiking activity in Pyr, PV (PV_1 and PV_2), SST_1 and SST_2 neurons, respectively. During the stimulus presentation (100-1000 ms, marked as the red arrow), the location of bump propagates through the circular lattice: PV neurons fire asynchronously. SST_1 neurons are quiescent, whereas SST_2 neurons mimic Pyr neurons. We also show Pyr and SST1 activity in **(E)** and **(F)**, respectively.

Figure 6: Integration of synaptic inputs. (A)-(D), Raster plots of Pyr, PV, SST_1 and SST_2, respectively, when there was a gap between stimulus presentations. During the gap (300-500 ms, marked by the black arrow), SST_1 neurons became active, and the bump activity of Pyr neurons stayed at the same location. **(E)**, Travelling time of bump activity to traverse the network as a function of the strength of the sensory inputs (i.e., input firing-rate values to PV_1); see **Methods** for estimates of travelling time. The mean values and standard errors were calculated from 10

independent simulations. The x-axis shows the sensory inputs to PV_1 cells; Pyr cells receive 20% more spikes than PV1 cells.

Figure 7: Ramping activity in the downstream neurons. We connected Pyr to downstream neurons with the coordinate-dependent connection probability as shown in **(A)**. *c*, the coordinate of the individual neuron in Pyr, ranges from 1 to 4000. **(B)** and **(C),** Population activity of downstream neurons with higher and lower maximal connection probability, respectively; p_0 is set to 1.0 and 0.1 for high and low connection probability. 50 ms non-overlapped bins were used to calculate the population activity. In contrast, **(D)** and **(E)** show the activity of randomly chosen 5 individual neurons with higher and lower probability, respectively, in the same temporal bins.

Figure 8: Correlations between firing rates of individual cells and time. We performed linear regression analyses to see if individual cells' activity increases over time. **(A)**, Mean values and standard deviations of slopes, R^2 and p-values from 5000 readout neurons. The blue and red represent dense (p=1.0) and sparse (p=0.1) connections, respectively. The mean value of p-values is ~10^{-5} when p_0=1.0. **(B)**, Histogram of p-values.

Figures

Figure 1

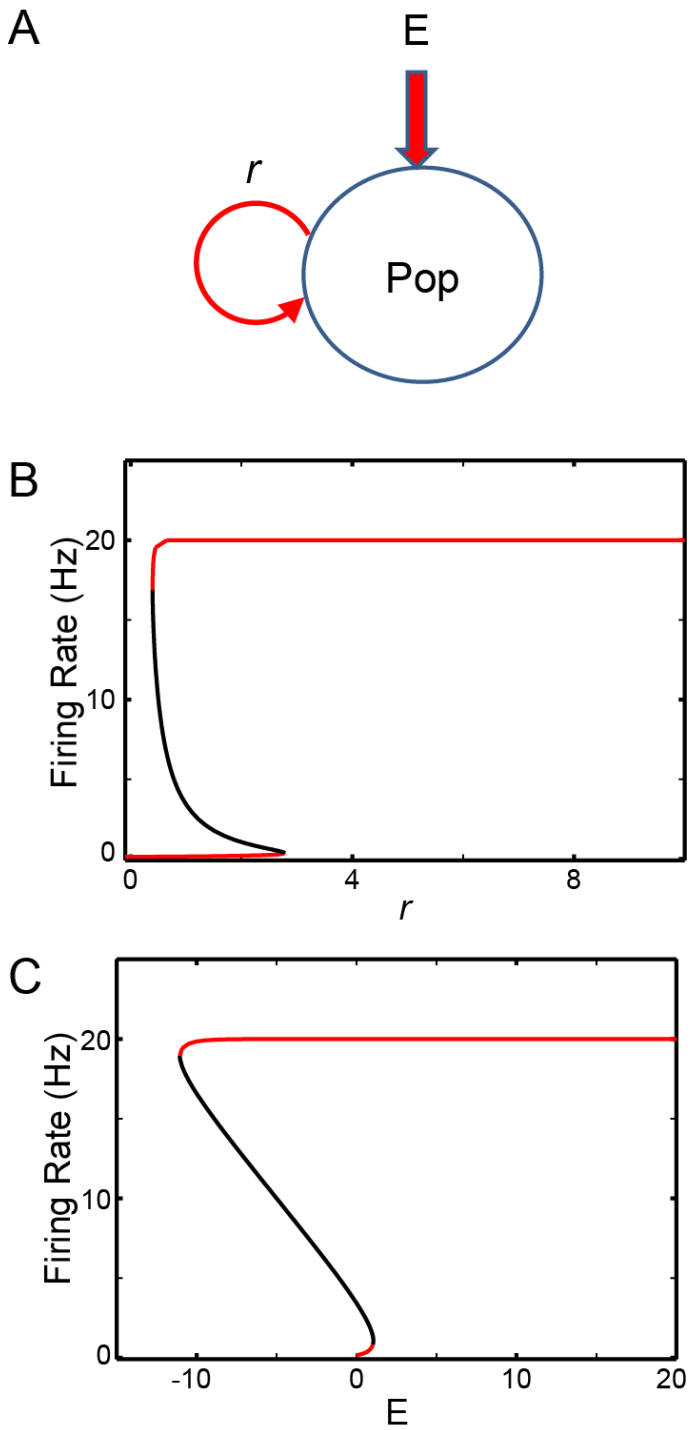


Figure 2

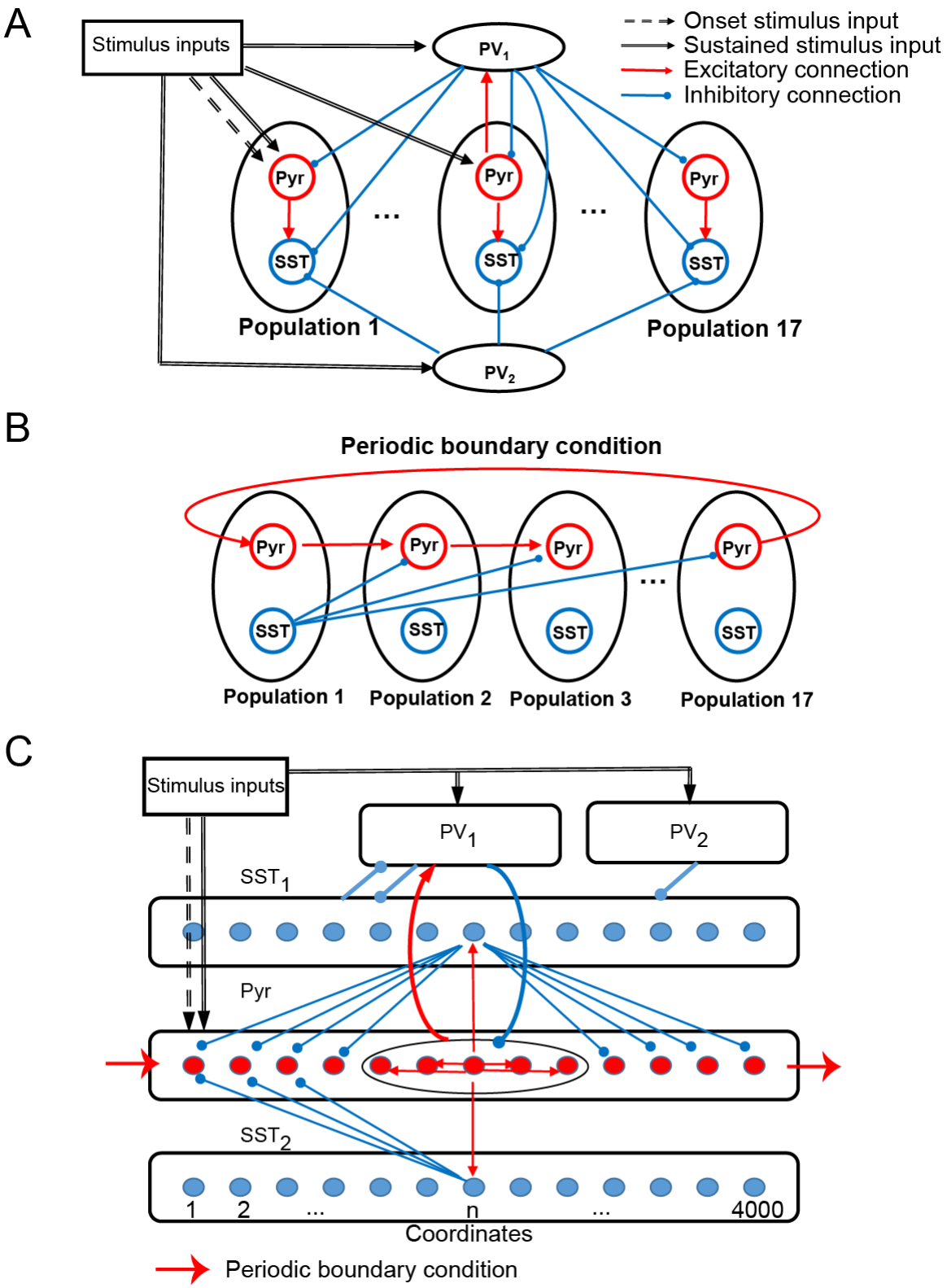

A
Stimulus inputs
PV1
PV2
Onset stimulus input
Sustained stimulus input
Excitatory connection
Inhibitory connection
Pyr
SST
Population 1
Population 17
B
Periodic boundary condition
Population 1
Population 2
Population 3
Population 17
C
Stimulus inputs
PV1
PV2
SST1
Pyr
SST2
1
2
...
n
...
4000
Coordinates
Periodic boundary condition

Figure 3

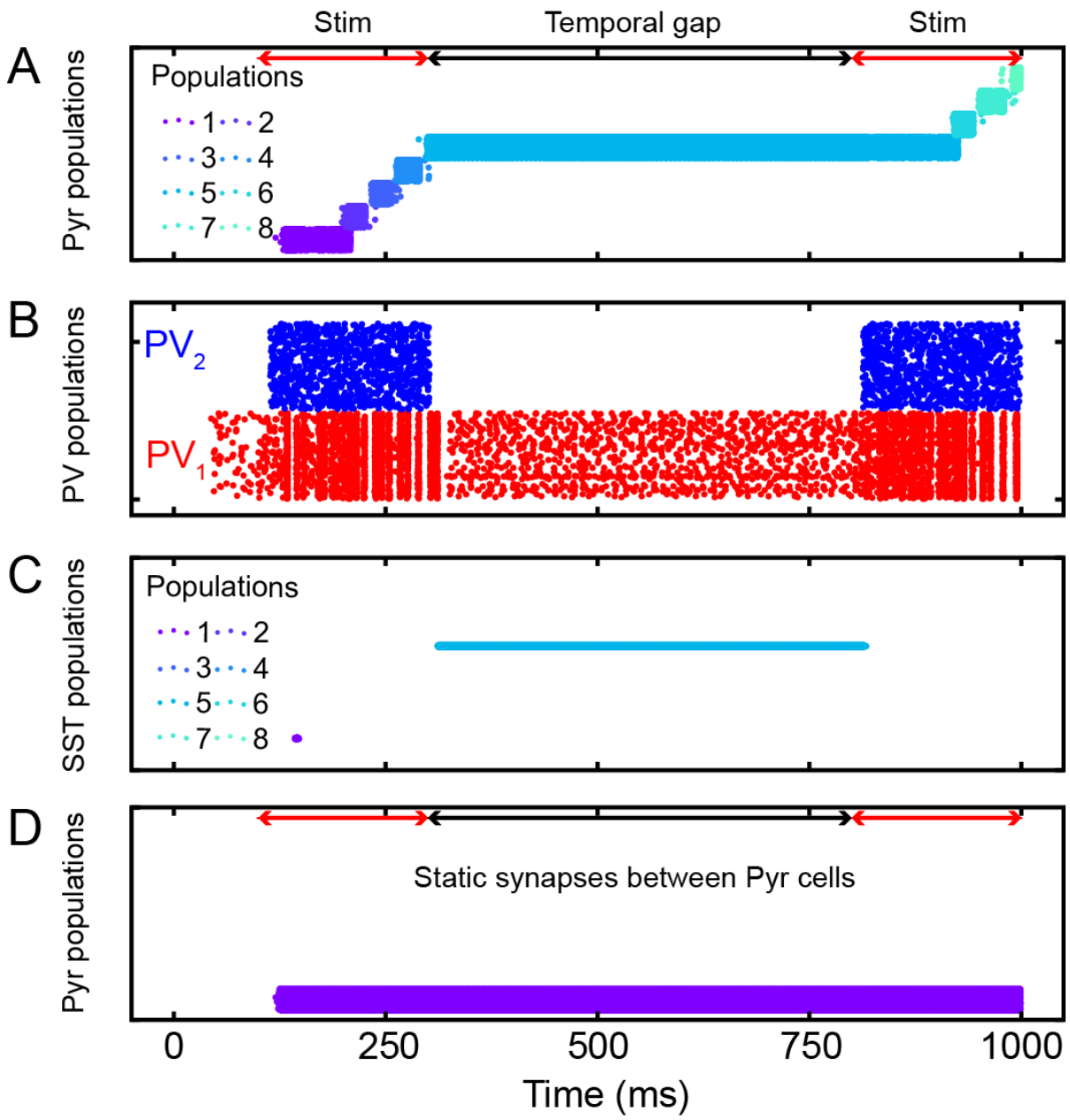

Stim
Temporal gap
Stim
A
Pyr populations
Populations
1
2
3
4
5
6
7
8
B
PV populations
PV2
PV1
C
SST populations
Populations
1
2
3
4
5
6
7
8
D
Pyr populations
Static synapses between Pyr cells
0
250
500
750
1000
Time (ms)

Figure 4

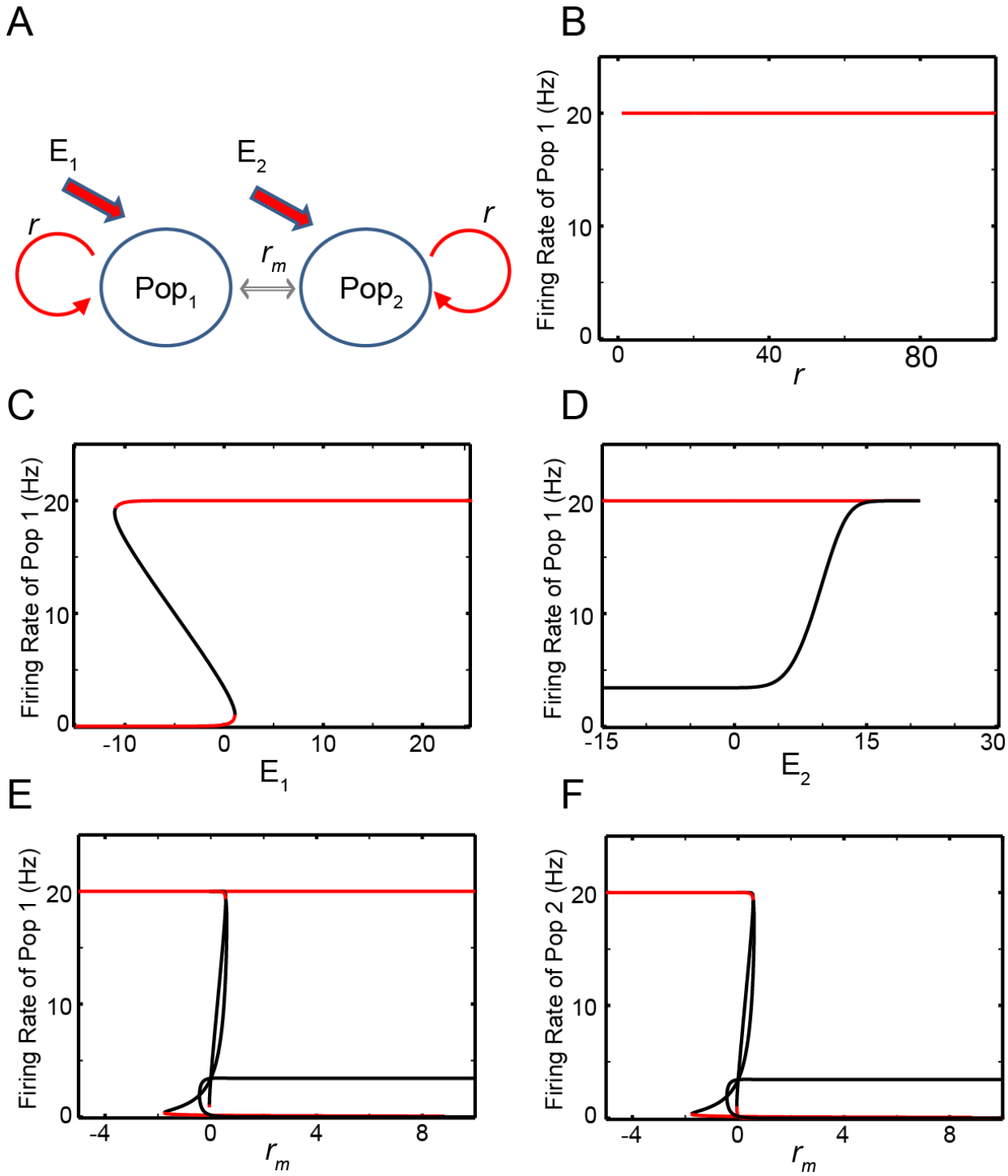

A
E1
E2
r
r
rm
Pop1
Pop2
B
Firing Rate of Pop 1 (Hz)
20
10
0
0
40
80
r
C
Firing Rate of Pop 1 (Hz)
20
10
0
-10
0
10
20
E1
D
Firing Rate of Pop 1 (Hz)
20
10
0
-15
0
15
30
E2
E
Firing Rate of Pop 1 (Hz)
20
10
0
-4
0
4
8
rm
F
Firing Rate of Pop 2 (Hz)
20
10
0
-4
0
4
8
rm

Figure 5

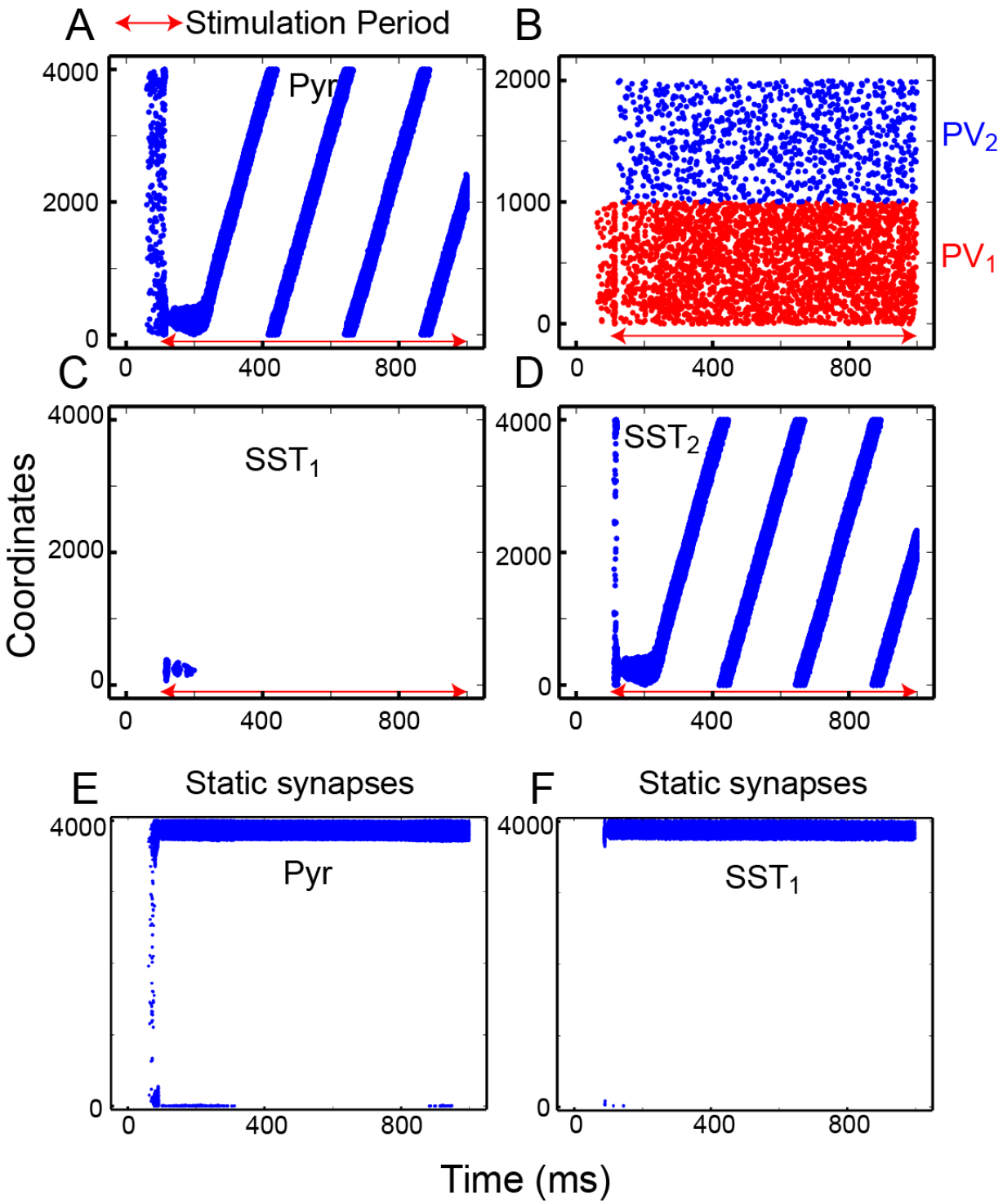

A
Stimulation Period
Pyr
B
PV2
PV1
C
SST1
D
SST2
Coordinates
E
Static synapses
Pyr
F
Static synapses
SST1
Time (ms)

Figure 6

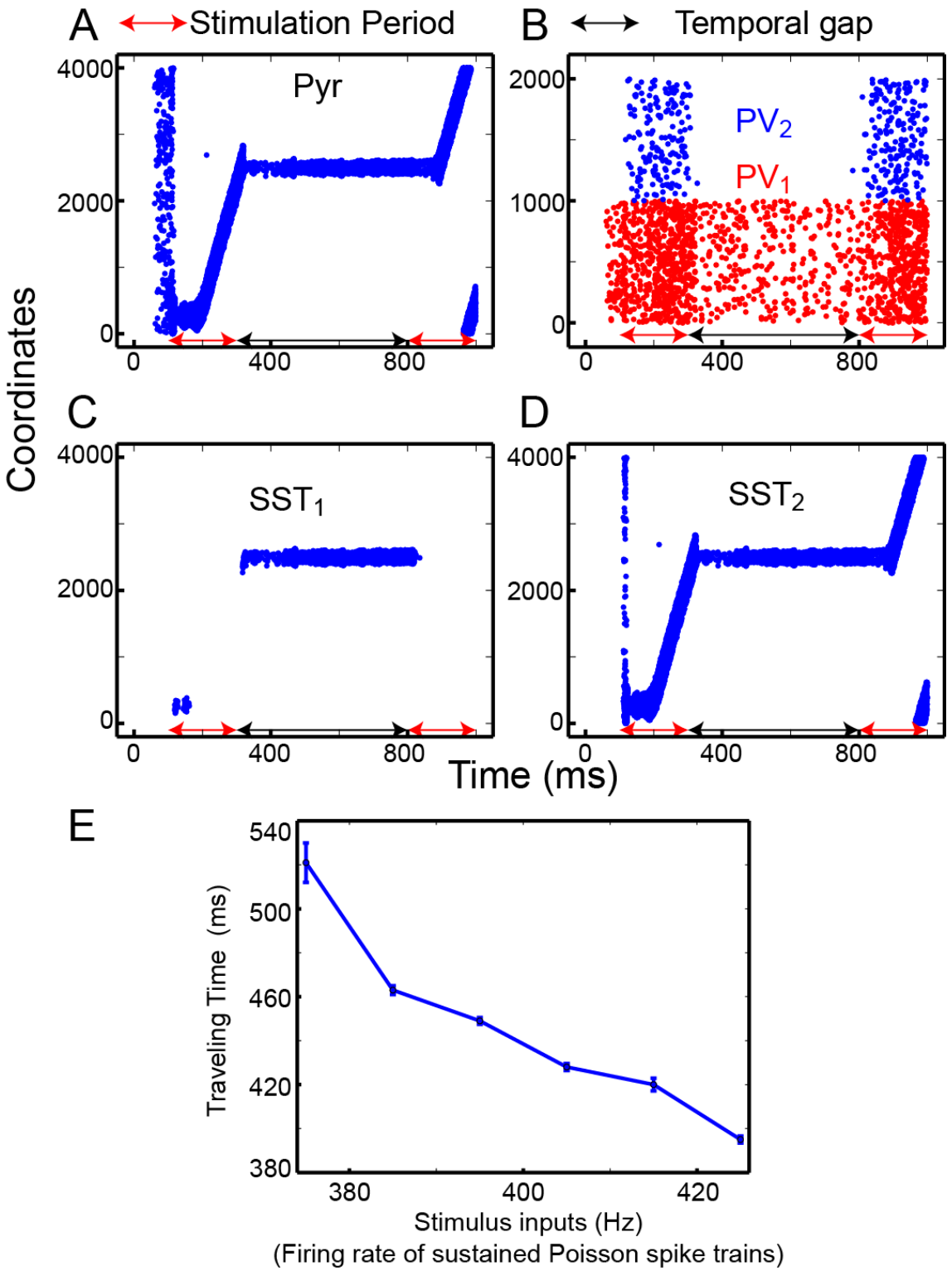

A
Stimulation Period
B
Temporal gap
Pyr
PV2
PV1
C
D
SST1
SST2
Coordinates
Time (ms)
E
Traveling Time (ms)
Stimulus inputs (Hz)
(Firing rate of sustained Poisson spike trains)

Figure 7

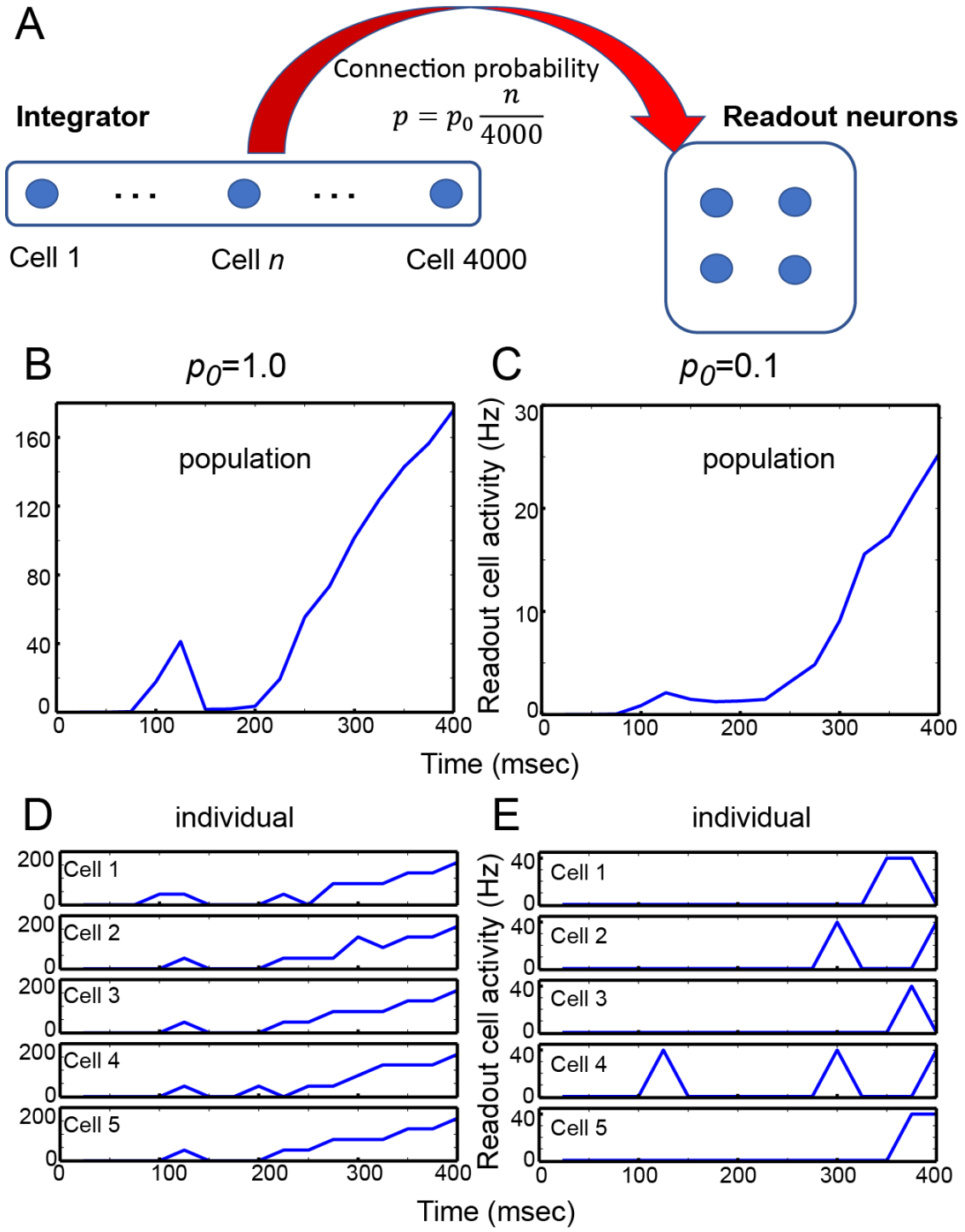
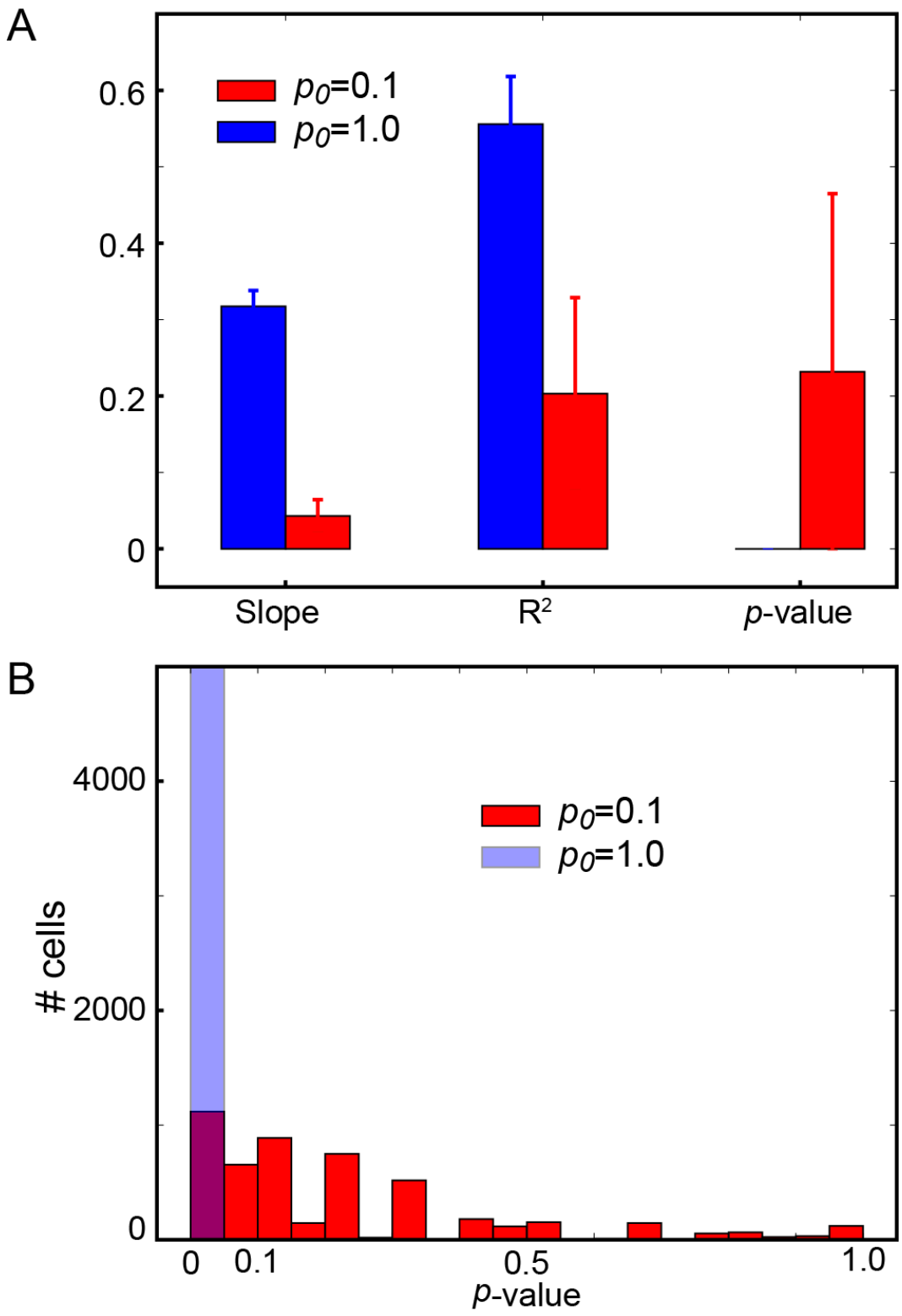

A
Connection probability
$p = p_0 \frac{n}{4000}$
Integrator
Readout neurons
Cell 1
Cell n
Cell 4000
B
p_0=1.0
C
p_0=0.1
population
population
Readout cell activity (Hz)
Time (msec)
D
individual
E
individual
Cell 1
Cell 2
Cell 3
Cell 4
Cell 5
Readout cell activity (Hz)
Time (msec)

Figure 8


A
0.6
0.4
0.2
0
p_0=0.1
p_0=1.0
Slope
R^2
p-value
B
4000
2000
0
cells
p_0=0.1
p_0=1.0
0
0.1
0.5
1.0
p-value

Methods

In this study, we developed discrete and continuous versions of a neural integrator. Both of these circuit networks were implemented within the NEST environment (Gewaltig and Diesmann, 2007), a peer-reviewed, freely-available simulation package. In both versions of the integrator, all of the neurons were leaky integrate-and-fire (LIF) neurons. The excitatory and inhibitory neurons within an integrator formed excitatory and inhibitory connections onto a set of 'target' neurons. All of the integrator neurons and in the target neurons had identical internal dynamics; specifically, each presynaptic spike induced an abrupt increase in a neuron's membrane potential that decayed exponentially. These neurons were implemented using the native NEST model iaf_psc_exp (Gewaltig and Diesmann, 2007). Table 1 shows the exact parameters used for the neurons and synapses in both neural integrators.

The structure of the *discrete* integrator

The structure of the discrete integrator is summarized in Figs. 2A and B. As seen in Fig. 2A, the discrete integrator consisted of 19 different neuronal populations. 17 of these neuronal populations were modeled to contain 400 pyramidal (Pyr) neurons and 16 somatostatin (SST) neurons. Within each of these 17 populations, Pyr neurons formed excitatory synapses with both Pyr and SST neurons. These 17 populations were topographically organized: Pyr neurons within a population had unidirectional excitatory connections with the adjacent population (e.g., population 2 projected to population 3 but not to population 1). We had a periodic boundary condition in which the (last) population 17 connected to the (first) population 1; see Fig. 2B. In contrast, SST neurons formed inhibitory connections with Pyr neurons in all of the other populations. Recurrent connections between Pyr neurons within a population had depressing

synapses, but all of the other synaptic connections were static. We implemented these depressing synapses using the Tsodyks-Markram model included in the NEST distribution (Table 1).

The two remaining populations each had 1088 parvalbumin (PV) neurons. All of the Pyr neurons had excitatory connections with the PV neurons in one population (PV_1) but not with those in the second PV population (PV_2). Both PV_1 and PV_2 neurons made non-specific inhibitory connections with Pyr and SST neurons; see Table 2 for the connection probability. These two PV populations simulated feedback and feedforward inhibition between Pyr neurons.

The structure of the *continuous* integrator

The continuous integrator was composed of a population of Pyr neurons, two PV populations (PV_1 and PV_2), and two populations of SST neurons (SST_1 and SST_2); see Fig. 4. Table 3 lists the parameters of these neuronal populations. In this network, 4000 Pyr, SST_1 and SST_2 neurons were distributed in a circular lattice, each of which had unique coordinate between 1-4000. We arbitrarily set the coordinates to increase in the clockwise direction. The neuronal numbers were arbitrary and were not constrained by the ratio of excitatory to inhibitory neurons, which is roughly 4:1. It should be noted that it is straightforward to extend this network model to include more excitatory neurons: instead of a single Pyr neuron at each coordinate, a small population of Pyr neurons at each coordinate can be instantiated without changing any of the details of the network structure.

Pyr neurons were mutually connected, via excitatory connections, to their neighboring Pyr neurons when the difference between their coordinates was $\leq\pm200$, which is similar to a distance-dependent connection probability (Perin et al., 2011). These connections were established with a periodic boundary condition: Pyr neuron 4000 and Pyr neuron 1 were

mutually connected. The PV_1 and PV_2 populations were independent of this circular lattice (see Fig. 4).

Pyr neurons interacted with the PV_1, SST_1 and SST_2 populations in distinct ways. First, the pattern of connectivity between the Pyr and PV_1 populations was randomly generated. Second, a Pyr neuron projected only to those SST_1 and SST_2 neurons that had the same coordinates (i.e., a one-to-one topographic mapping). The connection strength was designed to be just strong enough for a single Pyr "spike" to cause a SST_1 or SST_2 neuron to fire (Table 3), like a single layer-5 pyramidal-neuron spike can induce SST-expressing Martinotti neurons to fire (Silberberg and Markram, 2007). SST_1 and SST_2 neurons also had inhibitory connections with Pyr neurons but the connectivity rules differed. (1) SST_1 neurons formed connections only with those Pyr neurons in which the SST_2-and-Pyr difference was ≥ 200. (2) SST_2 neurons formed connections only with those Pyr neurons with lower coordinate values. PV_2 neurons randomly inhibited SST_1 neurons, and the connection probability is shown in Table 3. In our continuous integrator, all of the excitatory synapses were depressing, whereas all of the inhibitory synapses were static.

External inputs for both integrators

The excitability of each neuron depended on the sum of its synaptic inputs from all of the other neurons in the network and from external inputs, which were modeled as Poisson spike trains. Tables 2 and 3 show the neuron-specific rates of these external inputs. Sensory information was simulated as Poisson spike trains. In the model, sensory information consists of transient and sustained sensory information. The transient sensory information represented the transient cell responses observed in the sensory systems including retina, lateral geniculate nucleus and cortex (Cleland et al., 1971; de la Rocha et al., 2008; Piscopo et al., 2013; De Valois et al., 2000), and we assumed that this transient activity helped to ensure that bump activity always initiated in the

same location in the network. Transient inputs were introduced to the first 400 and 100 Pyr neurons in the discrete and continuous integrators, respectively. The duration of transient inputs was 100 msec, whereas the sustained inputs formed projections with all Pyr, PV_1 and PV_2 neurons during the entire stimulus. One can think of these sensory inputs as originating from the thalamus; though, we did not explicitly model the properties of thalamic neurons.

Travelling time for the bump

Using the continuous integrator, we examined if the propagation speed of the bump depended on the input strength by calculating the time course of the last 400 Pyr neurons (i.e., those with 400 highest coordinates). Specifically, we generated an event-related spike histogram using non-overlapping 10 ms bins of spiking data. “Travelling time of the bump” was defined as the time, relative to stimulus onset, when the number of spikes in a single bin exceeds the sum of the mean plus two standard deviations of the number of spikes during the simulation period.

Table 1: Neural parameters for neurons and synapses. When a spike arrived, the membrane potential instantly jumped to a new value, which was determined by its capacitance (*C*) and time constant (τ_m). When the membrane potential was higher than the spike threshold, the membrane potential was reset to the predefined value (V_{reset}). Without any external input, the membrane potential relaxes back its the resting membrane potentials (E_L). Synaptic events decay exponentially with a 2-ms time constant (τ_{syn}). All synapses have a 1.5 ms delay unless otherwise stated; the only exception is given in Table 2. For depressing synapses, we selected the parameters (U and τ_{ref}) given below.

Neuronal Parameters		Synaptic parameters	
C (membrane capacitance)	1 pF	τ_{syn}	2.0 ms
V_{th} (spike threshold)	20 mV	delay	1.5
τ_m (Membrane time constant)	20 ms	U	0.5
E_L (resting membrane potential)	0 mV	τ_{ref}	200 ms for discrete integrator 50 ms for continuous integrator
V_{reset} (reset after spiking)	0 mV		

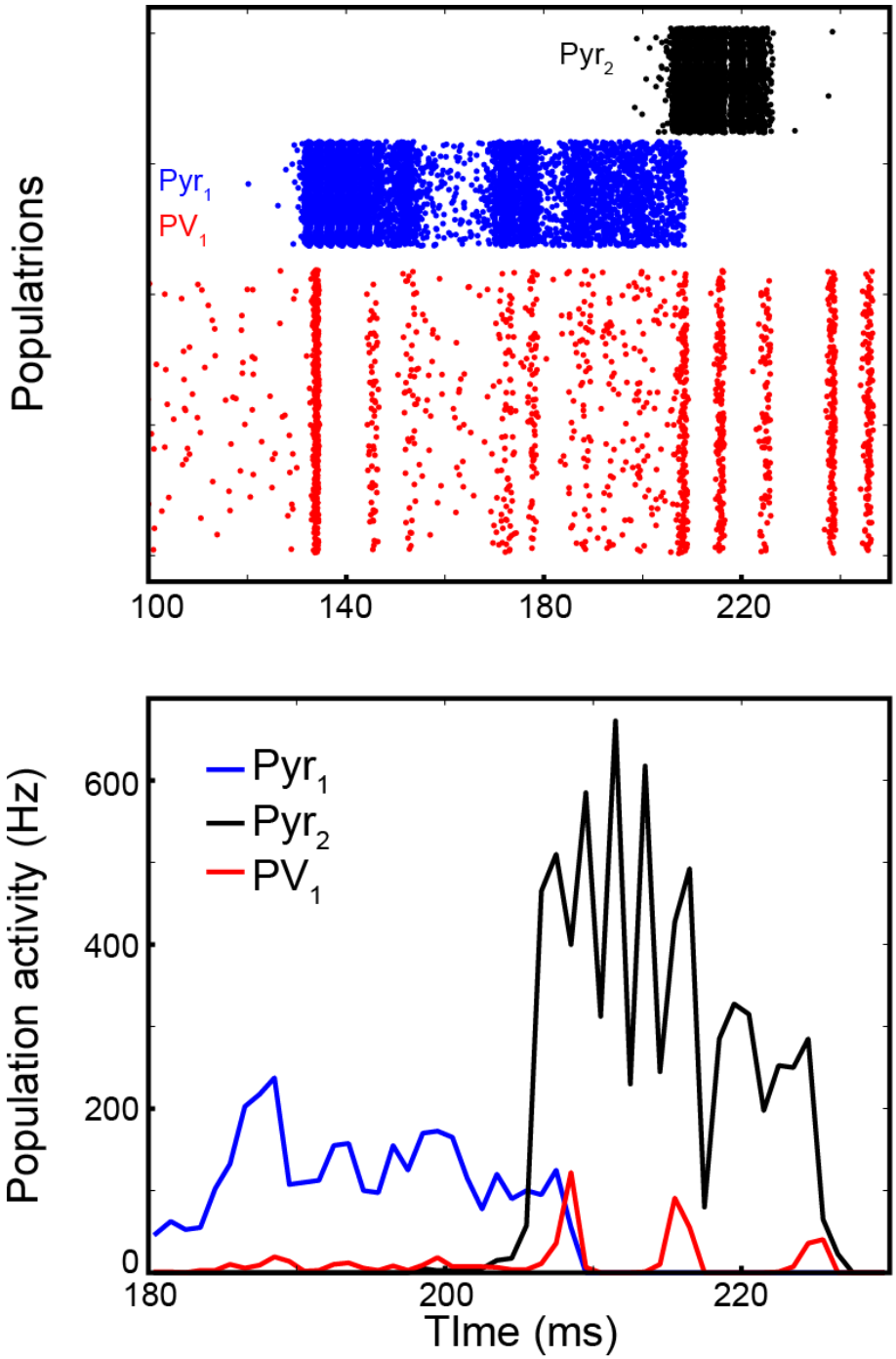
Table 2: The parameters used for the discrete integrator. We connected populations by specifying connection probabilities and synaptic connection strengths. The presynaptic neurons are shown before the arrow. The first value in the parentheses is the connection probability. The connection strengths follow the Gaussian distributions. The mean values are the second value in the parentheses, and the standard deviations are chosen to be 10% of the mean. The excitatory and inhibitory connections cannot be less or higher than 0, respectively; when they violate this condition, we set them to 0.

	Total Number	Background inputs (Hz)	Stimulus input (Hz; sustained)
Pyr	6800	2,800	2000
PV_1	1088	4,500	2000
PV_2	1088	N/A	2000
SST	544	3,200	N/A
Connectivity within populations (connection probability, strength (pA))			
Pyr→Pyr	(1.0 , 1.8)	Pyr→SST	(0.4, 0.96)
PV_1→PV_1	(0.3, -0.72)	PV_1→PV_1	(0.1, -0.72)
Connectivity across populations (connection probability, strength (pA))			
Pyr→Pyr	(0.2, 0.12) *delay 10 ms	PV_2→SST	(1.0, -6.0)
Pyr→PV_1	(0.2, 0.12)	SST→Pyr	(1.0,-4.8)
PV_1→Pyr	(0.2, -1.08)	SST→PV_1	(0.3, -0.6)
PV_1→SST	(0.3, -0.6)		
Connection strength for background and stimulus inputs (pA)			
Pyr	0.12	PV_2	0.36
PV_1	0.12	SST	0.12
Onset stimulus input			
Target	Pyr neurons in population 1	Firing rate	1000 Hz

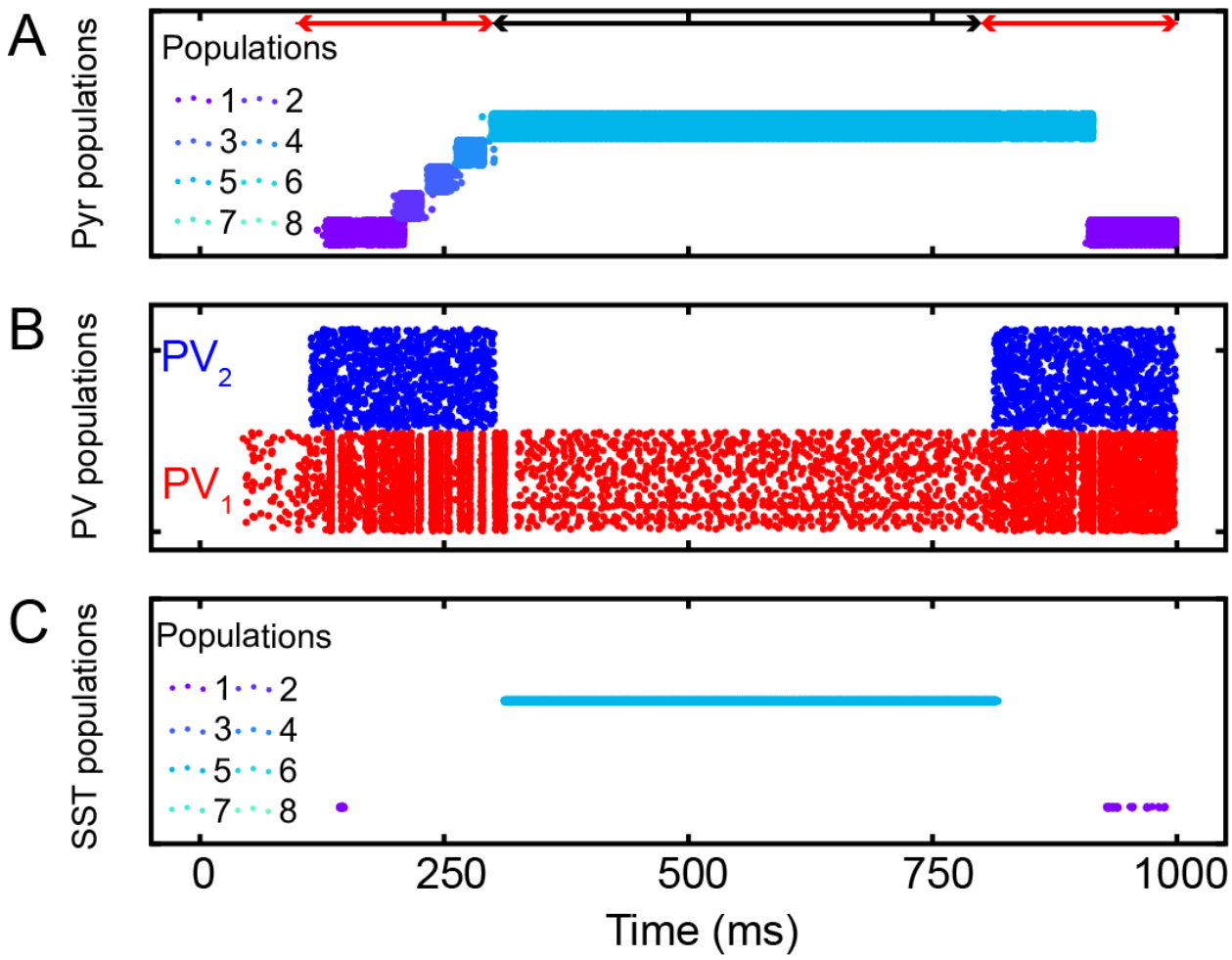
Table 3: The parameters used in the continuous integrator. Due to the lack of population structure, we connected neurons by specifying the number of inward connections to a specific neuron. Thus, we display the number of presynaptic neurons below.

	Total Number	Background inputs (Hz)	Stimulus input (Hz)
Pyr	4000	4,000	486
PV_1	1000	4,000	405
PV_2	1000	3,000	405
SST_1	4000	2,000	N/A
SST_2	4000	2,000	N/A
Connectivity (Number of presynaptic neurons, strength (pA))			
Pyr→Pyr	(400 , 0.12)	PV_1→SST_1	(100,-0.72)
Pyr→PV_1	(400, 0.12)	PV_2→SST_1	(1000, -0.72)
Pyr→SST_1	(1, 10.8)	SST_1→Pyr	(3600, -0.72)
Pyr→SST_2	(1, 10.8)	SST_1→PV_1	(100, -0.72)
PV_1→Pyr	(100, -0.72)	SST_2→Pyr	(400, -0.72)
Connection strength for background and stimulus inputs (pA)			

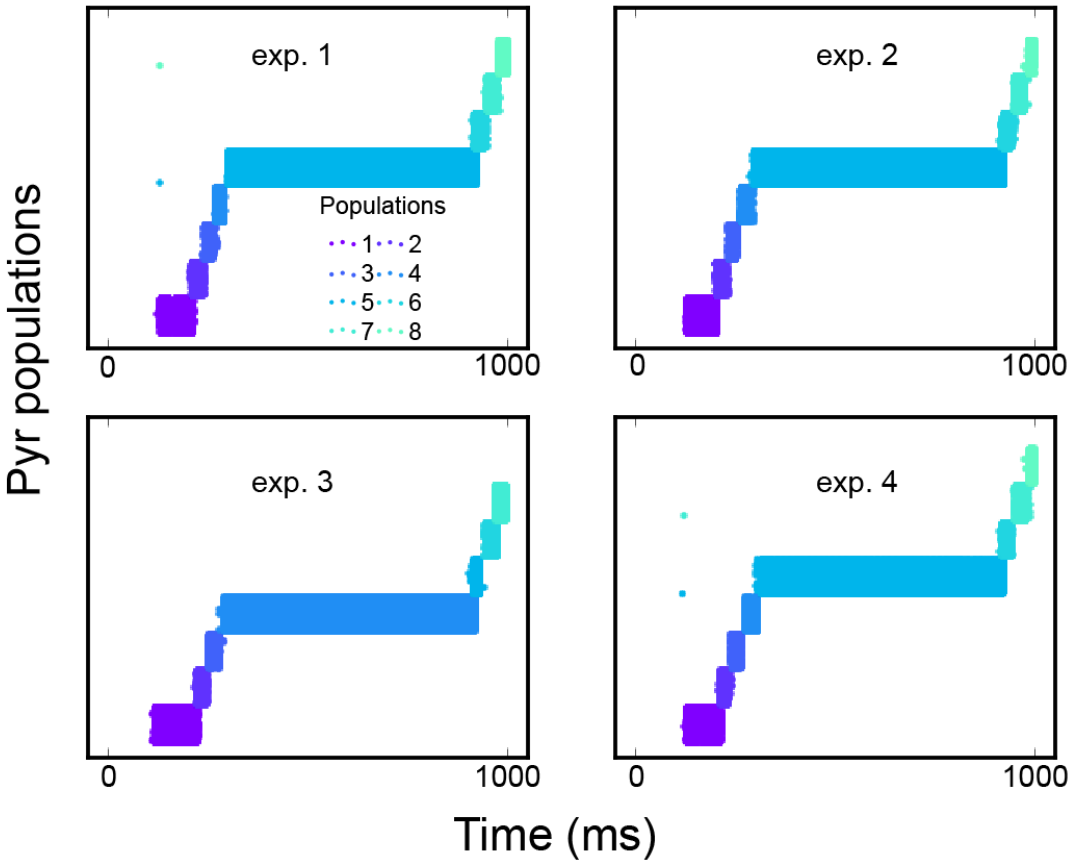
Supplemental Figures



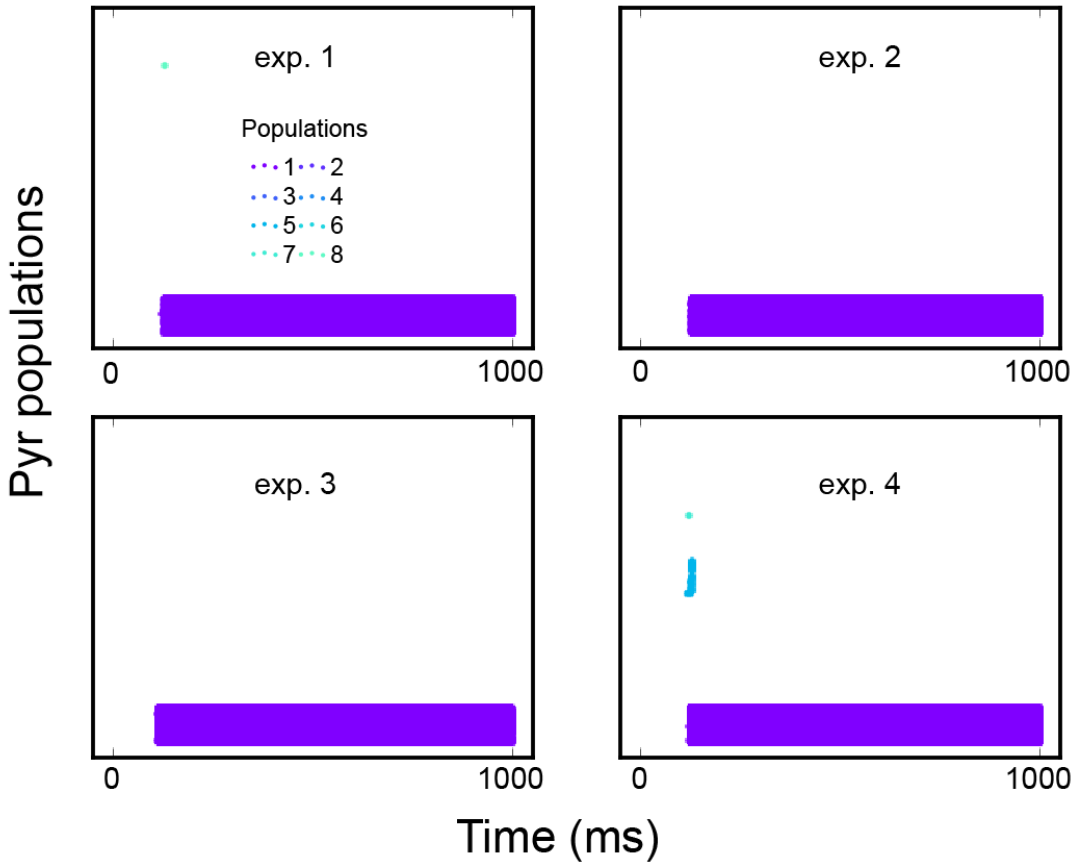
Supplemental Figure 1: Raster plots of Pyr_1, Pyr_2 and PV_1 populations during the transition of bump activity. (A), Action potentials of Pyr1, Pyr2 and PV1 are shown in blue, black and red, respectively. (B), Population activity rate calculated using 1-msec non-overlapping bins.



Supplemental Figure 2: The responses of populations without inhibition of PV_1 to Pyr_1 populations during the temporal gap. We repeated the same experiment shown in Fig. 3, but removed inhibition from PV1 to Pyr1 populations at 500 ms. That is, the first half of the responses are identical to those shown in Fig. 3. But the last half are different. As expected, population 1 was reactivated by transient inputs, when stimulus inputs were reintroduced; this destroyed the accumulated information. Conversely, lateral inhibition ensures the network accurately encode evidence accumulated.



Supplemental Figure 3: Lossless integration with depressing synapses. We instantiated 4 more networks and tested Pyr cell responses. In all four simulations, Pyr responses are consistent.



Supplemental Figure 4: Failure of lossless integration with static synapses. We instantiated 4 more networks and tested Pyr cell responses. In all four experiments, we observed that only population 1 active during the entire simulation.