

Impact of Neuronal Properties on Network Coding: Roles of Spike Initiation Dynamics and Robust Synchrony Transfer

Stéphanie Ratté,¹ Sungho Hong,² Erik De Schutter,² and Steven A. Prescott^{1,*}

¹Neurosciences and Mental Health, The Hospital for Sick Children and Department of Physiology, University of Toronto, Toronto, ON M5G 1X8, Canada

²Computational Neuroscience Unit, Okinawa Institute of Science and Technology, Okinawa 904-0495, Japan

*Correspondence: steve.prescott@sickkids.ca

<http://dx.doi.org/10.1016/j.neuron.2013.05.030>

Neural networks are more than the sum of their parts, but the properties of those parts are nonetheless important. For instance, neuronal properties affect the degree to which neurons receiving common input will spike synchronously, and whether that synchrony will propagate through the network. Stimulus-evoked synchrony can help or hinder network coding depending on the type of code. In this Perspective, we describe how spike initiation dynamics influence neuronal input-output properties, how those properties affect synchronization, and how synchronization affects network coding. We propose that synchronous and asynchronous spiking can be used to multiplex temporal (synchrony) and rate coding and discuss how pyramidal neurons would be well suited for that task.

The synaptic connectivity between neurons comprising a network is critical for the operation of that network but so too are the intrinsic properties of the constituent neurons. When it comes to studying network operation, focus on the former has often trumped consideration of the latter. We will, in this Perspective, shift the focus to neuronal properties and address how those properties affect the collective activity within a network, particularly with respect to synchrony (for review of network properties affecting synchrony, see Kumar et al., 2010). To be clear, we will not consider synchrony associated with network oscillations; instead, we will focus on the sort of stimulus-driven synchrony considered to be a “trivial reflection of anatomical connectivity” insofar as it arises in neurons receiving common input (Singer, 1999). Despite its humble origins, such synchrony has fundamentally important consequences for network coding and has been the focus of much debate (Brette, 2012; Bruno, 2011; de la Rocha et al., 2007; Diesmann et al., 1999; Ermentrout et al., 2008; Estebanez et al., 2012; Hong et al., 2012; Ikegaya et al., 2004; Josić et al., 2009; Kumar et al., 2008; Ostojic et al., 2009; Panzeri et al., 2010; Renart et al., 2010; Rossant et al., 2011; Salinas and Sejnowski, 2001; Sharifi et al., 2013; Stanley, 2013). Does this synchrony help or hinder network coding? Neuronal properties are a crucial yet underappreciated component of the answer.

Neurons are often said to operate as integrators or as coincidence detectors based on how they process input (Abeles, 1982; König et al., 1996). Integrators can summate temporally dispersed (asynchronous) inputs, whereas coincidence detectors respond only to temporally coincident (synchronous) inputs. In other words, integrators and coincidence detectors are both sensitive to synchronous input, but coincidence detectors are selective for it. Selectivity is, as we will explain, derived from the dynamical mechanism responsible for transforming synaptic input into output spiking. Spike initiation dynamics also affect

whether sets of neurons that receive common synchronous input spike synchronously and whether or not that output synchrony is easily disrupted (Figure 1). Spike initiation dynamics thus control synchrony transfer—the degree to which synchronous input elicits synchronous output. The precision and robustness of synchrony transfer has critical implications for both rate- and synchrony-based coding.

The terms *integration* and *coincidence detection* serve to highlight differences in neuronal operation—and we will use these terms for that purpose—but they do not accurately depict how an average neuron operates. Some neurons are exquisitely specialized to operate in one or the other mode but most, including the average pyramidal neuron, operate somewhere in between. In that respect, operating mode is best conceptualized not as a dichotomy, but rather as a continuum with “pure” integration and “pure” coincidence detection at either end (Figure 2). Neurons operating in the midrange may exhibit traits of both operating modes, with certain traits manifesting more strongly than others depending on stimulus properties. Indeed, although they are suboptimal for integration or coincidence detection, the lack of specialization may allow pyramidal neurons to simultaneously employ both operating modes so as to encode different stimulus features in concert, thus enabling rate- and synchrony-based coding to be multiplexed.

Beyond emphasizing that operating mode represents a continuum, we also propose to refocus its definition around the concept of synchrony transfer: coincidence detectors not only detect synchrony, they also transfer synchrony more precisely and robustly than do integrators (Figure 1). After establishing the importance of synchrony transfer, we will explain its biophysical basis by identifying the neuronal factors upon which synchrony transfer depends, namely, selectivity for synchronous input and capacity to produce robust synchronous output. By regulating synchrony transfer via these neuronal factors, spike

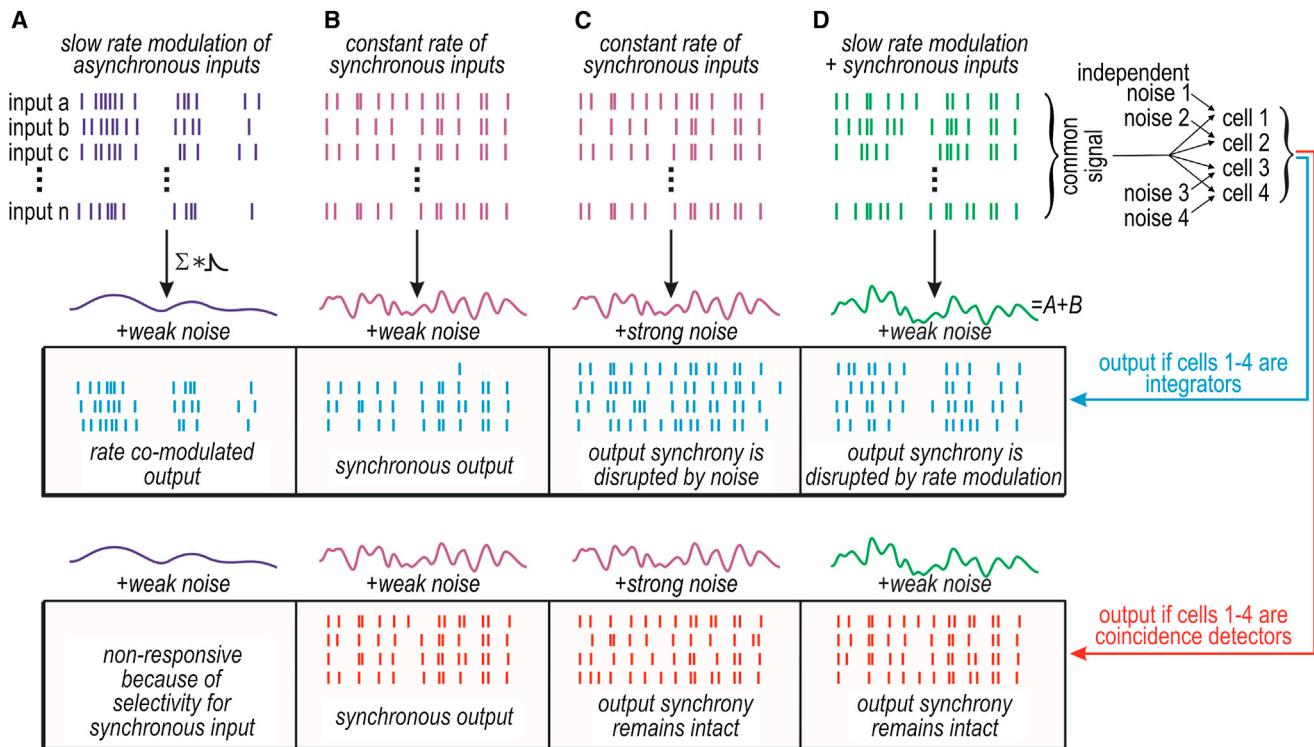


Figure 1. Synchrony Transfer Differs between Operating Modes

Top: the generation of differently shaped cumulative inputs based on the summation of input spike trains convolved with a synaptic conductance waveform. Stimulus and/or noise conditions differ between (A)–(D). Bottom: responses within a set of integrators or coincidence detectors receiving common (shared) input and independent noise. Unlike integrators, coincidence detectors respond selectively to synchronous input (compare A and B). Both integrators and coincidence detectors receiving common synchronous input will spike synchronously (B), but synchrony transfer is more robust among coincidence detectors, i.e., their output synchrony is less easily disrupted by strong independent noise (C) or by rate-modulated input (D). The robustness of synchrony transfer is a distinguishing feature of coincidence detectors.

initiation dynamics strongly influence whether a network encodes information by the timing of synchronous spikes and/or by the rate of asynchronous spikes.

Neural Coding Strategies: A Neuron's Perspective

Diverse candidate neural coding strategies have been identified (Perkel and Bullock, 1968). Those strategies are often divided into rate and temporal coding, but the division is not clear cut. The difference boils down to what timescale captures signal-dependent variations in spiking. The highest frequency (shortest timescale) encoded by the spike train can be inferred by analyzing the spike train with progressively smaller time windows to determine the window size at which mutual information between the spike train and the stimulus plateaus (Borst and Theunissen, 1999). The reciprocal of that time window represents the “sampling” rate, which, according to the Nyquist Theorem, should be at least twice the highest input frequency sampled by the neuron. Sampling rate relative to the spike rate determines whether the neural representation is sparse or dense, i.e., whether few (≤ 1) or many (>1) spikes can occur within each time window. Dense representations allow for spike counting, which is the basis for classic rate coding, whereas sparse representations do not (at least not within a single neuron on a single trial) and are thus often considered to imply temporal coding.

An important additional consideration for network coding is whether information is carried independently by each neuron or if information is available from the co-occurrence of spikes across two or more neighboring neurons—a correlation code (deCharms, 1998). We define *synchrony* as the co-occurrence of spikes within a time window narrow enough that only one spike per cell can occur within it (~5 ms), whereas *rate comodulation* is the cross-cell correlation of spike counts within broader time windows. A synchrony code is, therefore, a subtype of correlation coding—one that depends on precise spike timing. A synchrony code is also a subtype of temporal coding—one that depends on spike timing in one neuron relative to spike timing in neighboring neurons. Notably, if synaptic transmission is weak and unreliable (as is the case for many central synapses), synchrony is necessary for enabling brief inputs to activate the postsynaptic neuron (Stevens, 1994; Wang et al., 2010), which implies that synchrony is necessary for temporal coding.

In a network that exclusively utilizes rate coding, optimal coding occurs when neighboring neurons spike independently because correlations constitute redundancy, and redundancy usually reduces information capacity (Barlow, 1961; Gawne and Richmond, 1993; Mazurek and Shadlen, 2002; Sompolinsky et al., 2001; for review see Averbeck et al., 2006). Proponents of rate coding thus tend to view correlations, including synchrony,

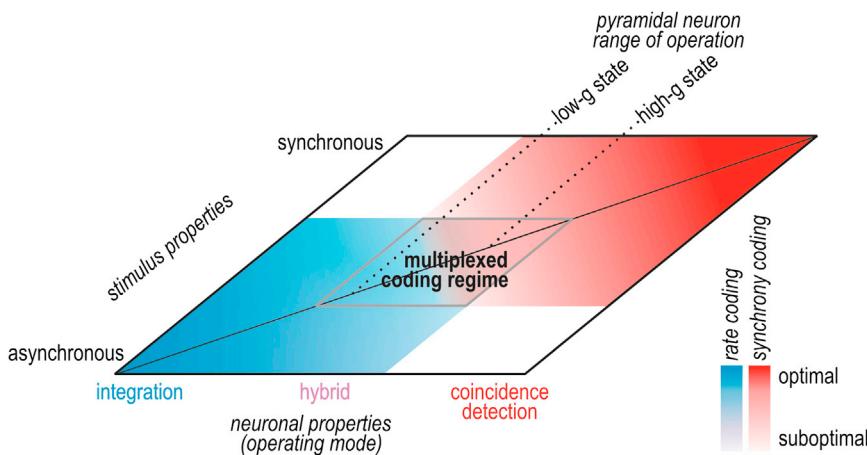


Figure 2. Neural Coding Depends Jointly on Neuronal Operating Mode and Stimulus Properties

Neuronal operating mode is represented as a continuum on one axis. Pyramidal neurons tend to operate in the middle range and can shift where they operate based on factors like conductance state. Input synchrony is represented on the other axis. Neural coding strategies are represented in blue (rate coding) and red (synchrony coding), and deeper colors represent better coding than paler colors. Pale regions overlap, revealing a regime in which a hybrid operating mode and multiplexed coding are possible.

as detrimental. Contrarily, synchrony between two neurons with overlapping receptive fields can lead to greater mutual information than if synchrony is ignored (Dan et al., 1998), meaning synchrony-encoded information can make up for, if not exceed, the reduction of rate-encoded information (Dan et al., 1998; Kenyon et al., 2004; Meister et al., 1995; Montani et al., 2007; Reich et al., 2001; Schnitzer and Meister, 2003), or so the proponents of synchrony coding would argue. Putting aside what proponents of either side think, we should ask what neurons think: to what inputs do they respond? Over what time window do they process input? After all, it is neurons that process information in the intact brain.

A single excitatory synaptic input typically causes only a small depolarization (<1 mV) in pyramidal and spiny stellate cells (Bruno and Sakmann, 2006; Mason et al., 1991; Sáez and Friedlander, 2009; Sayer et al., 1989; Song et al., 2005). Therefore, if a neuron sums input over a narrow time window (i.e., narrow enough that only one spike per presynaptic cell can occur within it), synchronous input from multiple presynaptic cells will be required to drive suprathreshold depolarization. On the other hand, if the neuron uses a broad time window (i.e., broad enough that multiple spikes per presynaptic cell can occur within it), suprathreshold depolarization can be driven by multiple inputs from just one presynaptic cell or via multiple presynaptic cells; the multicell input could be synchronous or asynchronous. Rerouting earlier definitions, coincidence detectors can be said to sum their inputs using a narrow time window, whereas integrators use a broad window (König et al., 1996). An integrator receiving synchronous input may appear to use a narrow window, but the window size is really a property of the neuron, not of the stimulus, which supports a neuron-centric definition of operating mode as opposed to a stimulus-centric one (Rudolph and Destexhe, 2003). The importance of a neuron-centric definition becomes clear when comparing synchrony transfer: integrators respond to synchronous input, but they do not transfer that synchrony as robustly as coincidence detectors do (see Figure 1).

Before proceeding, it is worth noting that simply having a spike threshold endows the neuron with sensitivity to the derivative of the input current or membrane potential (Agüera y Arcas and

Fairhall, 2003; Hong et al., 2007). In line with this, it has been shown that the simple threshold-and-fire model as well as leaky integrate-and-fire models can transfer synchrony under the appropriate stimulus conditions (Burak et al., 2009; Goedeke and Diesmann, 2008; Schultze-Kraft et al., 2013; Tchumatchenko et al., 2010). However, as Tchumatchenko et al. and Schultze-Kraft et al. note, this is true only for limited (and arguably unrealistic) stimulus conditions, i.e., high input synchrony driving large membrane potential fluctuations. In real neurons and in more sophisticated models whose spike initiation dynamics implement band-pass filtering, and which are therefore preferentially sensitive to relevant stimulus frequencies, the stimulus requirements for robust synchrony transfer are much less stringent (and more plausible).

Requirements for Synchrony Coding

Rate coding is broadly accepted as the pre-eminent coding strategy in the brain; by comparison, synchrony coding is contentious and often considered applicable only to particular systems like the auditory midbrain. We contend that synchrony coding occurs more broadly based on several lines of evidence. We will organize our discussion of that evidence around the 3-fold requirements for synchrony coding (Figure 3A): (1) principal neurons must have coincidence detector traits (in order to reliably transfer synchrony under realistic stimulus conditions), (2) they must receive synchronous input that contains information, and (3) they must produce synchronous output that can be decoded. Note that rate coding and synchrony coding are not mutually exclusive even though factors that facilitate one often do so at the expense of the other. The feasibility and utility of each coding strategy should be gauged independently, contrary to many past debates.

Requirement 1 is satisfied insofar as principal neurons can and do operate as *imperfect* coincidence detectors. This is suggested by their highly irregular spike trains since integration of multiple asynchronous inputs tends to produce regular spiking (Softky and Koch, 1993); but this speaks more to the input (see below), since even integrators receiving irregularly timed synchronous inputs will spike irregularly (Salinas and Sejnowski, 2000; Shadlen and Newsome, 1998; Stevens and Zador, 1998). More importantly, pyramidal neurons in the intact brain are constantly bombarded by synaptic input, so much so that they are chronically depolarized and shunted (Bernander et al.,

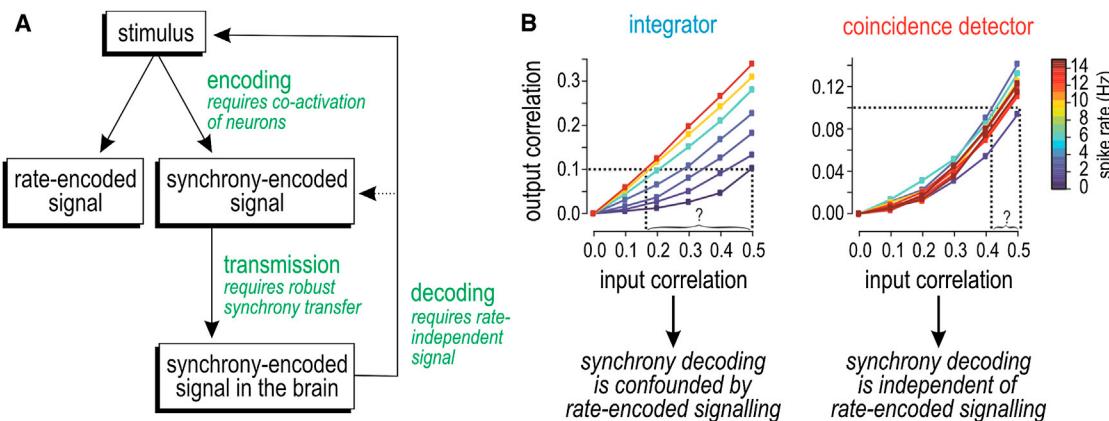


Figure 3. Requirements for Synchrony Coding and the Robustness of Synchrony Transfer to Spike-Rate Variation

(A) A synchrony-encoded signal arises from stimulus-dependent coactivation of neurons, which is not mutually exclusive of rate-encoded signaling. For synchrony-encoded signals to reach the CNS, they must be reliably transmitted across multiple synapses and must remain decodable in order to provide information about the original stimulus. Decodability relies on robust synchrony transfer.

(B) Graphs illustrate the challenge of decoding synchrony. Among integrators, the correlation input-output relationship varies with spike rate; consequently, for a given output correlation value, one cannot infer (decode) the input correlation without also knowing the spike rate. This suggests that synchrony coding cannot operate independently of rate coding and would necessitate a complicated decoding mechanism. However, among coincidence detectors, the input-output relationship is not confounded by variations in spike rate, meaning synchrony decoding from coincidence detectors is straightforward. (B) is modified from Hong et al. (2012).

1991; Destexhe and Paré, 1999; for review see Destexhe et al., 2003). Moreover, sensory input causes concomitant (albeit momentarily unbalanced) increases in both excitatory and inhibitory drive (Borg-Graham et al., 1998; Haider et al., 2013; Pouille et al., 2009; for review see Isaacson and Scanziani, 2011), which implies further increases in total conductance. The reduction in input resistance ($R = 1/g$) decreases neuronal sensitivity to constant and slowly fluctuating (low-frequency) inputs, but the concomitant reduction in the membrane time constant ($\tau = RC$) makes neurons relatively more sensitive to rapidly fluctuating (high-frequency) inputs. In addition, large membrane potential fluctuations driven by synaptic bombardment increase sensitivity to coincident inputs (Rossant et al., 2011). This tendency is enhanced by a nonlinear increase in adaptation that can further reduce sensitivity to slow input and thus enhance selectivity for fast input (Hong et al., 2012; Prescott et al., 2006, 2008b). The cumulative effect is that pyramidal neurons receiving realistic conductance-based background and stimulus-evoked inputs *in vivo*, and which therefore exist in a high-conductance state, behave more like coincidence detectors than is suggested by *in vitro* testing with artificial current-based stimuli (see also Azouz and Gray, 2000, 2003). To be clear, pyramidal neurons do not switch abruptly from one to the other operating mode but, instead, shift along a continuum (see Figure 2) and can exhibit reasonably strong coincidence detector traits.

Requirement 2 is satisfied insofar as principal neurons do receive synchronous input. For one, the cortex receives sensory input via synchronized activation of thalamocortical neurons (Alonso et al., 1996; Bruno and Sakmann, 2006) originating from the coactivation of primary sensory neurons (see below). Pyramidal neurons recorded *in vivo* exhibit irregular spiking (see above) driven by large fluctuations in membrane potential that, based on the small depolarization produced by unitary syn-

aptic events, can only be accounted for by some degree of synchrony among presynaptic cells (Destexhe and Paré, 1999; DeWeese and Zador, 2006). Indeed, cross-cell correlations in membrane potential (Lampl et al., 1999; Poulet and Petersen, 2008; Yu and Ferster, 2010) and spiking (Cohen and Kohn, 2011; deCharms and Merzenich, 1996; Jadhav et al., 2009; Smith and Kohn, 2008) have been documented *in vivo*. Spike cross-correlations are typically measured through pairwise comparisons; however, a postsynaptic neuron experiences correlations across its entire set of presynaptic neurons, which means that correlation values measured through pairwise comparisons must be scaled in order to infer the total input correlation. Very small pairwise correlations that have been reported as evidence for asynchrony (e.g., Ecker et al., 2010) can in fact belie large total input correlation (Rossant et al., 2011; Schneidman et al., 2006).

The origins of synchronous spiking dictate whether synchrony represents signal or noise. Realistic stimuli have spatiotemporal structure that enables them to coactivate neurons with adjacent or overlapping receptive fields; consequently, coactivation patterns can contain information about the stimulus (Brette, 2012; Dan et al., 1998; Meister et al., 1995). If coactivation patterns contain information, synchrony represents part of the signal. Although this does not prove that synchrony-encoded signals are decoded, nor can synchrony be labeled noise simply because it reduces the information decodable from rate-encoded signals; indeed, it would be equally unfair to label rate-encoded signals as noise because they compromise the decoding of synchrony-encoded signals (see below). That said, the aforementioned points do not rule out stimulus-independent synchrony that is truly noise (Mastronarde, 1989). What is arguably more important is that correlated spiking in higher brain areas has been observed to be stimulus dependent (Alonso et al., 1996; deCharms and Merzenich, 1996; Kohn and Smith,

2005; Temereanca et al., 2008), consistent with synchrony-encoded signals being successfully transmitted to the cortex.

Requirement 3 is satisfied insofar as synchrony-encoded signals are decodable depending on which type of cells carries the message. It has been suggested that synchrony decoding is implausible because of an “inextricable” link between output correlation and spike rate (de la Rocha et al., 2007). If synchrony transfer were to vary with spike rate, input correlation could not be unambiguously decoded from output correlation without that rate sensitivity being factored in, and indeed the synchrony-encoded information could be lost unrecoverably. However, although synchrony transfer is rate dependent among integrators (except under more extreme stimulus conditions; Schultze-Kraft et al., 2013), the same is not true for coincidence detectors (Figure 3B) (Hong et al., 2012; Tchumatchenko et al., 2010), which argues that synchrony-encoded messages carried by coincidence detectors are decodable. Hence, pyramidal neurons with coincidence detector traits should be able to produce synchronous output that is decodable.

These three requirements reflect upon the encoding, transmission, and decoding of synchrony-based signals. Encoding requires the structured coactivation of neurons. Decoding requires that synchrony-encoded signals are not conflated with other signals; in that respect, decodability depends on reliable transmission. Reliable transmission requires robust synchrony transfer. We must, therefore, understand what makes synchrony transfer robust. We will deconstruct the biophysical basis for robust synchrony transfer by considering two factors: (1) the selectivity of neurons for synchronous input and (2) their capacity to produce synchronous output. We will explain each factor in turn, linking both to spike initiation dynamics.

Selectivity for Synchronous Input

According to our neuron-centric definition of operating mode, integrators can summate asynchronous inputs, whereas coincidence detectors are excited uniquely by synchronous inputs (see Figure 1). In other words, coincidence detectors are selective for (i.e., tuned to) synchrony, whereas integrators are relatively untuned with respect to synchrony. Synchrony is reflected in spectral properties of the input: synchronous input has greater power at high frequencies and less power at low frequencies compared with asynchronous input of equivalent magnitude (i.e., with equivalent total power) (Destexhe et al., 2001). Putting two and two together, one might (correctly) postulate that integrators are tuned to lower frequencies, akin to a low-pass filter, whereas coincidence detectors are tuned to higher frequencies, akin to a high-pass filter, although the end result is a band-pass filter when the high-pass filter implemented by spike initiation is combined with the low-pass filter implemented by membrane capacitance.

Differential tuning reflects differences in neuronal excitability. A simple yet invaluable classification of excitability was provided by Hodgkin (1948) who identified three spiking patterns in response to sustained depolarization: Class 1 neurons can spike repetitively at an arbitrarily low rate and thus have a continuous frequency-current (f - I) curve, class 2 neurons cannot spike repetitively below a certain rate and thus have a discontinuous f - I curve, and class 3 neurons fire only one or a few spikes at

stimulus onset (Figure 4A). Each class of excitability is associated with differences in other response measures such as the phase response curve (Ermentrout, 1996) and spike-triggered average (Ermentrout et al., 2007; Mato and Samengo, 2008) (see below). In general, class 1 neurons exhibit integrator traits, whereas class 3 neurons and, to a lesser extent, class 2 neurons exhibit coincidence detector traits. Hodgkin’s classification thus provides a useful starting point for relating neuronal excitability with operating mode.

Differences in excitability reflect differences in spike initiation dynamics (Izhikevich, 2007; Prescott et al., 2008a; Rinzel and Ermentrout, 1998). “Dynamics” refers to how fast and slow currents interact to control spike initiation. Notably, currents with similar kinetics sum linearly whereas those with different kinetics interact nonlinearly. Therefore, net-fast and net-slow currents interact nonlinearly, and ultra-slow processes like adaptation currents or cumulative inactivation of sodium current can be treated as modulating the fast-slow interaction. Net-fast current is necessarily inward (depolarizing) at spike threshold. In class 1 excitability, net-slow current is also inward at perithreshold voltages and thus cooperates with fast current during spike initiation (Figure 4B). In class 2 and 3 excitability, net-slow current is outward (hyperpolarizing) at perithreshold voltages and thus competes with fast current during spike initiation. Class 2 excitability exists if fast inward current overpowers slow outward current when constant stimulation exceeds threshold. Class 3 excitability exists if fast inward current overpowers slow outward current only during a stimulus transient, which precludes repetitive spiking during sustained stimulation. Thus, on the basis of whether fast and slow currents cooperate or compete at perithreshold voltages, three classes of excitability arise from a continuum in the strength and direction of net-slow current. The strength of net-fast current (which depends on leak current) affects its competition with net-slow current, thus influencing the boundary between class 2 and 3 excitability (Lundstrom et al., 2008; Prescott et al., 2008a). In dynamical terms, it is the cooperative versus competitive nature of the interaction controlling spike initiation that distinguishes integration and coincidence detection. To be clear, net current depends on both activation and inactivation of contributing ion channels, meaning inactivation of an outward current has effects comparable to activation of an inward current if the two processes occur with similar kinetics and voltage dependency. Accordingly, and especially given that pyramidal neurons express a multitude of different ion channels, there are several distinct channel combinations that can implement equivalent spike initiation dynamics. That said, the interaction between membrane currents also depends on the stimulus waveform because subthreshold membrane currents are differentially activated or inactivated by stimuli with different kinetics. This speaks to the joint dependence of spiking on neuronal properties and stimulus properties (see below for discussion on filtering).

With respect to synaptic input, subthreshold inward current helps sustain the depolarization caused by excitatory inputs, thereby encouraging temporal summation (integration) in class 1 neurons; contrariwise, subthreshold outward current truncates the depolarization caused by excitatory inputs, thereby discouraging summation and allowing only coincident inputs that drive

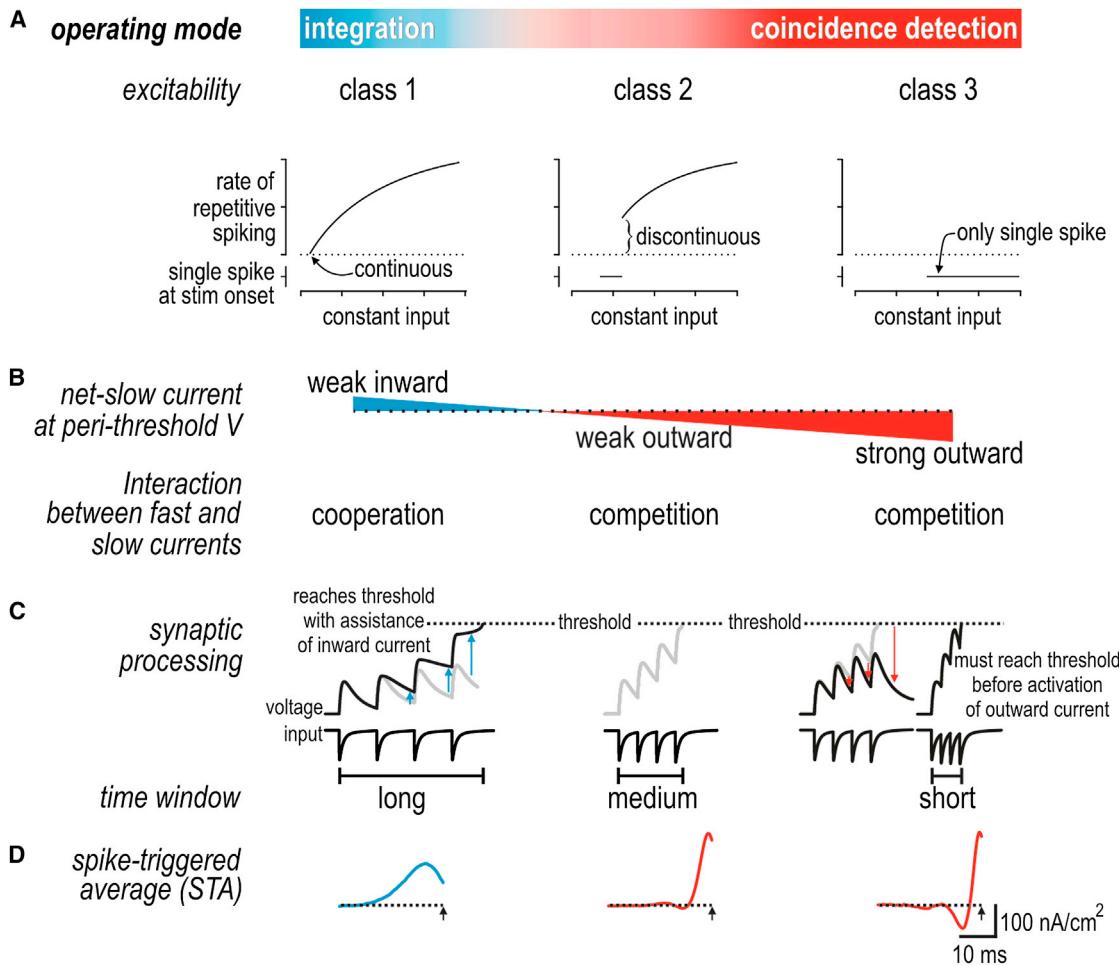


Figure 4. Spike Initiation Dynamics Control Operating Mode

(A) Classes 1, 2, and 3 of excitability are distinguished by the shape of the frequency-current curve defined by constant stimulation. Class 3 neurons (and class 2 neurons within a certain stimulus range) fire only one or a few spikes at stimulus onset. Those properties emerge from distinct nonlinear dynamical mechanisms that reflect whether fast and slow currents cooperate or compete during spike initiation.

(B) Differences in spike initiation dynamics can be ascribed to differences in the direction and magnitude of the net-slow current active at perithreshold potentials.

(C) Inward current helps sustain the depolarization caused by excitatory synaptic inputs, thereby lengthening the integration time window; outward current truncates depolarization, thereby shortening the integration time window.

(D) Differential processing is also evident in the shape of the spike-triggered stimulus average (STA).

fast suprathreshold depolarization (i.e., faster than outward current can activate) to elicit spiking in class 2 and 3 neurons (Figure 4C). In effect, the width of the integration time window is regulated by the strength and direction of subthreshold currents (Fricker and Miles, 2000; Gastrein et al., 2011; Prescott and De Koninck, 2005). Note that the delayed negative feedback implemented by voltage-dependent outward current in class 2 and 3 neurons has an effect very similar to that mediated by feed-forward synaptic inhibition, which is well recognized as a mechanism that limits the integration time window (e.g., Pouille and Scanziani, 2001; see also Ostojic et al., 2009). The difference lies in whether the negative feedback is a feature of the neuron or of the microcircuit.

An efficient way to assess signal processing characteristics, including the integration time window, is to measure the spike-triggered stimulus average (STA). This can be done by applying

noisy stimulation comprising a range of input frequencies and calculating the average stimulus waveform that precedes each spike; the noisy input can be constructed to reasonably approximate synaptic bombardment (Destexhe et al., 2001) and avoids having to repeat testing across multiple single-frequency inputs and combinations thereof (Rieke et al., 1997). The STA differs between class 1 (integrator) and class 2/3 (coincidence detector) neurons, being broad and monophasic in the former versus narrow and biphasic in the latter (Hong et al., 2012; Mato and Samengo, 2008) (Figure 4D). Duration of the positive phase reflects the integration time window. More generally, the STA reflects the stimulus features that drive spiking based on the recruitment of subthreshold membrane currents: a broad monophasic STA represents low-pass filtering (which confers tuning to low frequencies), whereas a narrow biphasic STA represents band-pass filtering (which confers tuning to higher frequencies).

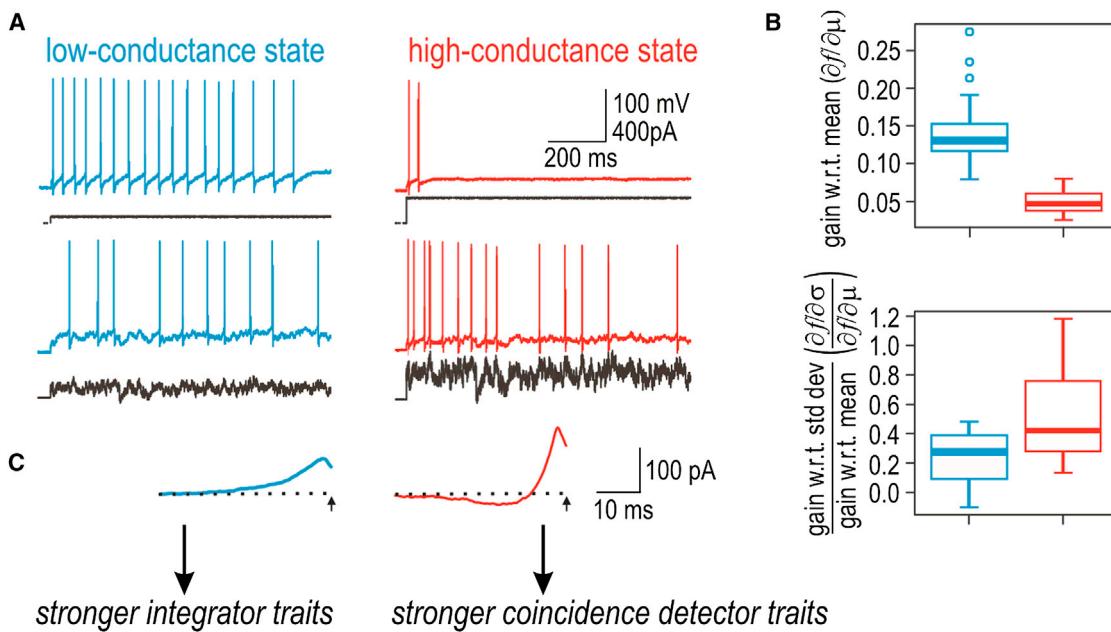


Figure 5. Pyramidal Neuron Operating Mode Is Intermediate and Modulable

(A) When tested in a low-conductance state, CA1 pyramidal neurons spike repetitively to constant stimulation and can maintain low spike rates, consistent with class 1 excitability. When the same neuron is tested in the high-conductance state (recreated via dynamic clamp), excitability is shifted toward class 2 excitability, as evidenced by a reduced tendency to maintain repetitive spiking during constant stimulation. Fluctuating stimuli can elicit vigorous spiking in either conductance state.

(B) The shift in excitability is accompanied by a shift in coding properties: neurons become less sensitive to the mean stimulus intensity (μ) and relatively more sensitive to the amplitude of stimulus fluctuations (σ), consistent with coincidence detector traits becoming more prominent in the high-conductance state.

(C) The shift is also accompanied by reshaping of the STA from a broad monophasic form to a narrower biphasic form. Modified from Hong et al. (2012).

The difference in signal processing is also evident in the spike-triggered stimulus correlation (STC) (Rieke et al., 1997) (see below).

Where do pyramidal neurons fit into this classification? Regular spiking CA1 pyramidal neurons exhibit class 1 excitability when tested in brain slices (Prescott et al., 2006, 2008b), but the synaptic bombardment experienced *in vivo* (see above) is predicted to encourage class 2/3 excitability by biasing the net-slow current at threshold in the outward direction. Using dynamic clamp to mimic synaptic bombardment in brain slices, voltage threshold undergoes a depolarizing shift because greater depolarization is needed to activate enough fast sodium channels to overwhelm the increased outward leak current. This, in turn, allows activation of other voltage-dependent outward currents and slow inactivation of inward currents, thereby biasing the net-slow current in the outward direction and encouraging class 2 excitability (Prescott et al., 2006, 2008b). Consequently, the same neuron that spikes repetitively during constant current injection in the low-conductance state often spikes only transiently when retested in the high-conductance state, although fluctuating stimuli can elicit vigorous spiking in either conductance state (Figure 5A). The shift in excitability, from class 1 to class 2, is associated with quantifiable changes in tuning: neurons become less sensitive to the stimulus mean and relatively more sensitive to the stimulus variance (Hong et al., 2012) (Figure 5B). The shift in operating mode is paralleled by reshaping of the STA (Figure 5C). Similarly, neocortical pyramidal cells tested *in vitro* tend to operate near the division between

class 1 and 2 excitability (Tateno et al., 2004; Tsubo et al., 2007) and can be made more class 2 excitable through enhanced adaptation (Stiefel et al., 2008). In general, adaptation currents and slow inactivation of inward currents can enhance sensitivity to the stimulus variance without completely nullifying responsiveness to the stimulus mean (Arsiero et al., 2007; Fernandez et al., 2011; Higgs et al., 2006; see also Lundstrom et al., 2009). These data show that pyramidal neurons exhibit coincidence detector traits and identify spike initiation dynamics as a key determinant of their operating mode.

Predicting Cross-Correlations through Reverse Correlation Analysis

Given a neuron's output spike train and its STA, reverse correlation can be used to predict its input. Conversely, how the neuron encodes its input can be modeled using its STA. By extension, if two neurons receive common input, the STA can be used to predict the correlated spiking driven by that input, and thus it can predict the cross-correlogram (CCG) (Figure 6A). More precisely, the shape of the CCG can be inferred by convolving the STAs from each neuron (Goldberg et al., 2004). It follows from their differently shaped STAs that the CCG for a pair of coincidence detectors is narrow and multiphasic, whereas the CCG for a pair of integrators is broad and monophasic (Hong et al., 2012; see also Barreiro et al., 2010, 2012).

However, the STA does not provide a sufficiently accurate description of neuronal response properties when the neuron is sensitive to multiple stimulus features. In this scenario, the

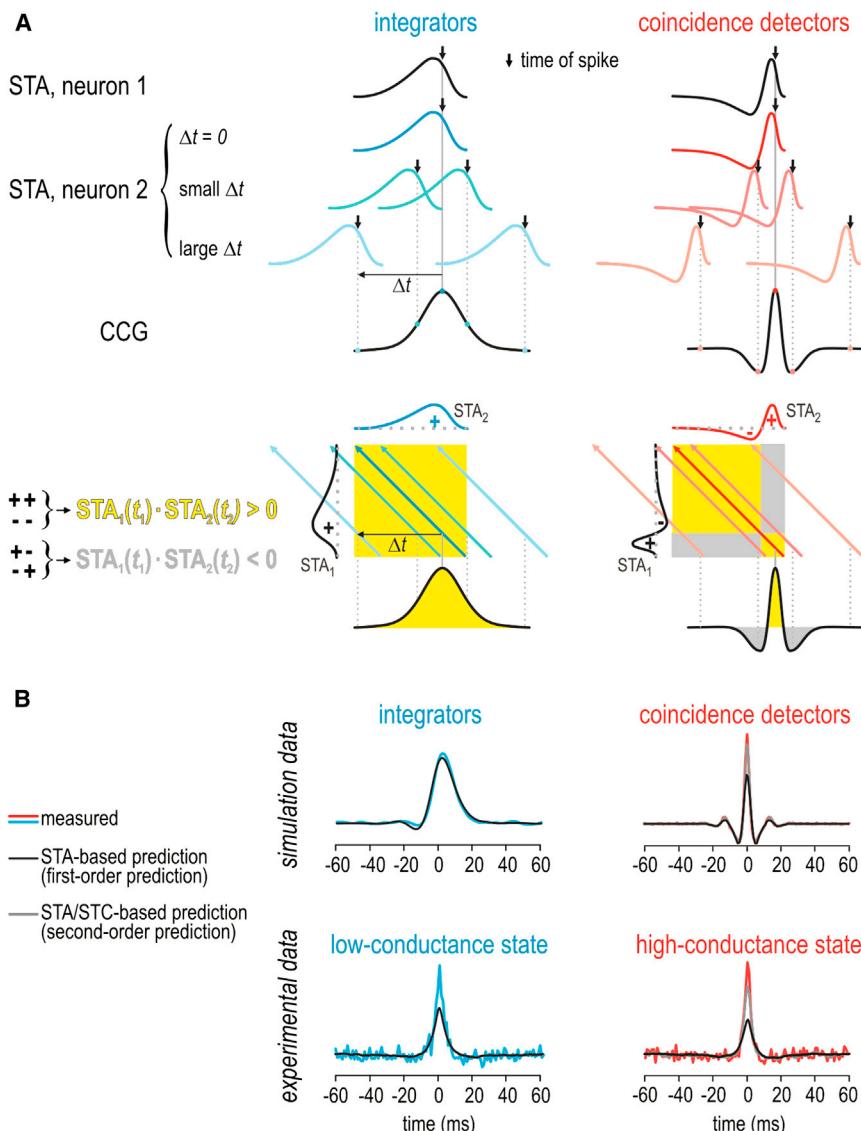


Figure 6. Predicting the Cross-Correlogram
 (A) The cross-correlogram (CCG) can be predicted by convolving the STAs of each neuron. Top: the STA in neuron 2 shifted by different Δt relative to the STA in neuron 1. For $\Delta t = 0$, the two STAs overlap perfectly, which corresponds to a high cross-correlation value. For large Δt , the cross-correlation drops to 0 as the STAs no longer overlap. For coincidence detectors, the cross-correlation can be negative for intermediate Δt if the STAs line up out of phase. Bottom: another depiction in which the STA in neuron 1 (STA_1) is plotted against the STA in neuron 2 (STA_2). Shading represents $STA_1(t_1) \cdot STA_2(t_2)$ with yellow corresponding to conditions in which the positive and negative components of STA_1 are in phase with the positive and negative components of STA_2 , and gray corresponding to conditions in which those components are out of phase. Colored arrows are the projections of STA_2 across $STA_1(t_1) \cdot STA_2(t_2)$ for the same Δt values shown in the top panel. The total cross-correlation represents the sum of $STA_1(t_1) \cdot STA_2(t_2)$ across that arrow.

(B) Examples of predicted CCGs for comparison with measured CCGs. The first-order prediction (based on the STA alone) provides a satisfactory fit to CCGs measured in the integrator model but does a poor job fitting the peak of CCGs measured from coincidence detector models. The “excess” synchrony was better accounted for by the second-order prediction (based on the STA and STC). For experimental data from CA1 pyramidal neurons, the second-order prediction becomes relatively more important when neurons are shifted toward the coincidence detector mode (i.e., in the high-conductance state) but is relevant even in the low-conductance state insofar as the first-order prediction is imperfect. This is consistent with pyramidal neurons operating in the middle range of the operating mode continuum.

information for building a good encoding model can be retrieved by the spike-triggered stimulus correlation (or equivalently the covariance; STC) (for details, see Schwartz et al., 2006). For reasons explained below, the STA-based encoding model provides a relatively good description of integrator response properties, whereas the multifeature model is needed to provide a similarly good description of coincidence detector response properties (Agüera y Arcas et al., 2003; Slep et al., 2005). By extension, the STC improves prediction of the CCG, but more so for coincidence detectors CCGs than for integrator CCGs (Hong et al., 2012). Notably, the multifeature model more accurately predicts the narrow central peak of the CCG that dominates the total correlation in coincidence detectors (Figure 6B).

Differential importance of the STC for predicting coincidence detector spiking compared with integrator spiking reflects upon the stimulus features that elicit spikes in each operating mode. In brief, integrators spike when the integrated stimulus in-

tensity exceeds some threshold; the STA accurately captures that feature selectivity. Stimulus intensity is also important for spike initiation in coincidence detectors, but the competitive dynamics render the process additionally (and nonlinearly) sensitive to the rate of change of stimulus intensity. The shape of the coincidence detector STA hints at the importance of abrupt depolarizing input, but the STC more accurately captures the sensitivity to rate of change of stimulus intensity, including how that sensitivity varies with stimulus intensity (for more detailed explanation, see Agüera y Arcas et al., 2003). In short, the STA is sufficient to distinguish integrator and coincidence detector operating modes and it can be used to qualitatively predict the shape of the CCG for pairs of neurons operating in either mode, but higher-order stimulus properties such as the STC become important in the case of coincidence detectors and provide quantitatively more accurate predictions.

Capacity to Produce Robust Synchronous Output

Previous discussions of operating mode have emphasized how neurons process their input. But to explain synchrony transfer, we must also consider how neurons produce their output and,

moreover, we must consider the output of multiple neurons in order to measure output synchrony. This would seem to require the difficult task of recording simultaneously from all the neurons whose output is to be cross-correlated; however, by replaying the same simulated synaptic input signal (along with different noise), one can collect many spike trains from individually recorded neurons and then cross-correlate their responses after alignment based on the common signal (de la Rocha et al., 2007; Hong et al., 2012; Reyes, 2003). We refer to this as a *virtual network* approach since the neurons, although not part of the same “real” network, are stimulated and analyzed as if they are part of the same “virtual” network. Notably, the input synchrony and the fraction of input that is shared across neurons are not only known, they are controlled by the experimenter. This approach is therefore very useful for studying how and why synchrony transfer differs between operating modes.

Synchronous spiking across a set of neurons requires that spike timing within each constituent neuron is temporally precise in relation to the input. Rapidly fluctuating input—the sort arising from presynaptic synchrony—drives more precisely timed spikes than constant or slowly fluctuating input (Bryant and Segundo, 1976; Cecchi et al., 2000; Galán et al., 2008; Mainen and Sejnowski, 1995; Nowak et al., 1997). Those data demonstrate that spike timing can be precise on the basis of input and thus support a stimulus-centric definition of operating mode (Schultze-Kraft et al., 2013), but neuronal properties are nonetheless critical. By being less sensitive to mean stimulus intensity, coincidence detectors exhibit better spike-timing precision than integrators firing at an equivalent average rate (Prescott et al., 2006; Prescott and Sejnowski, 2008). Indeed, several studies have linked stronger outward membrane current with increased precision (Berry and Meister, 1998; Billimoria et al., 2006; Schreiber et al., 2004; Svirskis and Rinzel, 2003), whereas inward currents or slowly inactivating outward currents have the opposite effect (Barreiro et al., 2012; Cudmore et al., 2010; Fricker and Miles, 2000). Specifically, band-pass filtering in coincidence detectors attenuates low-frequency input such that repetitive spiking is prevented or reduced and membrane potential is, in a sense, “clamped” below threshold. Rapid stimulus fluctuations elicit spikes (because they are not attenuated) and the timing of those spikes is very precise (see above). The critical point is this: because fluctuation-driven spikes are not superimposed on repetitive mean-driven spiking, spike timing is more tightly linked to stimulus fluctuation timing (Prescott and Sejnowski, 2008). Unlike in integrators, the rate of spiking in pure coincidence detectors reflects the rate of synchronous suprathreshold inputs, not the amplitude of a slow, rate-encoded signal (see Figure 1) (König et al., 1996)—this explains the rate insensitivity of synchrony transfer among coincidence detectors (Figure 3B). But once again bear in mind that pyramidal neurons operate in a middle range and can exhibit mean-driven and fluctuation-driven spiking. The two spike “types” can coexist so long as timing of the latter is not strongly corrupted by the former and so long as a decoder can ultimately separate the two. We will address both issues below.

Beyond being insensitive to spike rate, synchrony transfer must also be robust to noise. Indeed, it has been shown that a small perturbation can elicit an extra spike in the recipient cell,

which in turn elicits extra spikes in multiple postsynaptic cells, resulting in large stimulus-independent (i.e., noisy) variations in membrane potential that disrupt spike timing (London et al., 2010). London et al. did not, however, demonstrate that perturbations elicit synchronous spikes; that would require that the perturbation occurs synchronously across multiple neurons (which is conceivable) and that the recipient neurons are all simultaneously close to threshold (which is doubtful) so that the input is not only received simultaneously, but it also elicits spikes simultaneously. Without synchronous activation of multiple presynaptic cells, postsynaptic coincidence detectors would not be activated, or at least a set of coincidence detectors would not be activated synchronously. As a result, asynchronous perturbation-driven spiking will be curtailed, not amplified, within a network of coincidence detectors. In this regard, it is noteworthy that London et al. used integrator-type model neurons in their simulations and that their experiments, although conducted *in vivo*, seemed to emphasize the low-conductance state (e.g., reported values of input resistance are comparable to those in Destexhe et al., 2001 before synaptic bombardment); this may reflect the inclusion of the down state that exists during anesthesia but that is absent during wakefulness (e.g., Constantinople and Bruno, 2011) and/or the exclusion of sensory evoked activity that would increase conductance (see above).

To better understand why synchrony transfer by coincidence detectors is robust, consider the following hypothetical scenario with numbers based loosely on published data (e.g., Wang et al., 2010). A set of 50 coactivated neurons synapse onto a postsynaptic coincidence detector that requires only 30 synchronous excitatory inputs to achieve suprathreshold depolarization. This means that only 30 of the 50 presynaptic neurons must spike simultaneously in order to excite the postsynaptic neuron. The other 20 presynaptic neurons need not be activated or their spikes could be lost to noise without compromising postsynaptic activation (Zador, 1998)—we refer to this as an *excess synchrony safety margin* (Figure 7A). By “lost spikes,” we mean spikes that would have been elicited by the signal but are absent because of the effects of noise. On the other hand, the likelihood of noise simultaneously coactivating 30 presynaptic neurons is arguably quite low (see above)—we refer to this as the *minimum synchrony safety margin*. In other words, synchrony-driven spiking will not be easily disrupted or confused with noise-driven spiking in the presence of these safety margins. An important conclusion is that temporal coding is more robust when it uses synchronous spikes among multiple neurons rather than isolated spikes in single neurons—this seems obvious but is routinely overlooked.

Beyond affecting the probability of signal-driven spikes, noise could also compromise synchrony by jittering the timing of signal-driven spikes. Intriguingly, spike timing in coincidence detectors is protected against jitter. This quality control mechanism can be understood from the shapes of the STA and CCG (Figure 7B). Consider another hypothetical scenario in which two neurons spike synchronously. The STA provides an estimate of the common signal that triggered those spikes. Next, consider what would happen if neuron 2 received a perturbation. The perturbation would almost certainly jitter spike timing in neuron 2, but it might also reduce the probability that neuron 2 even

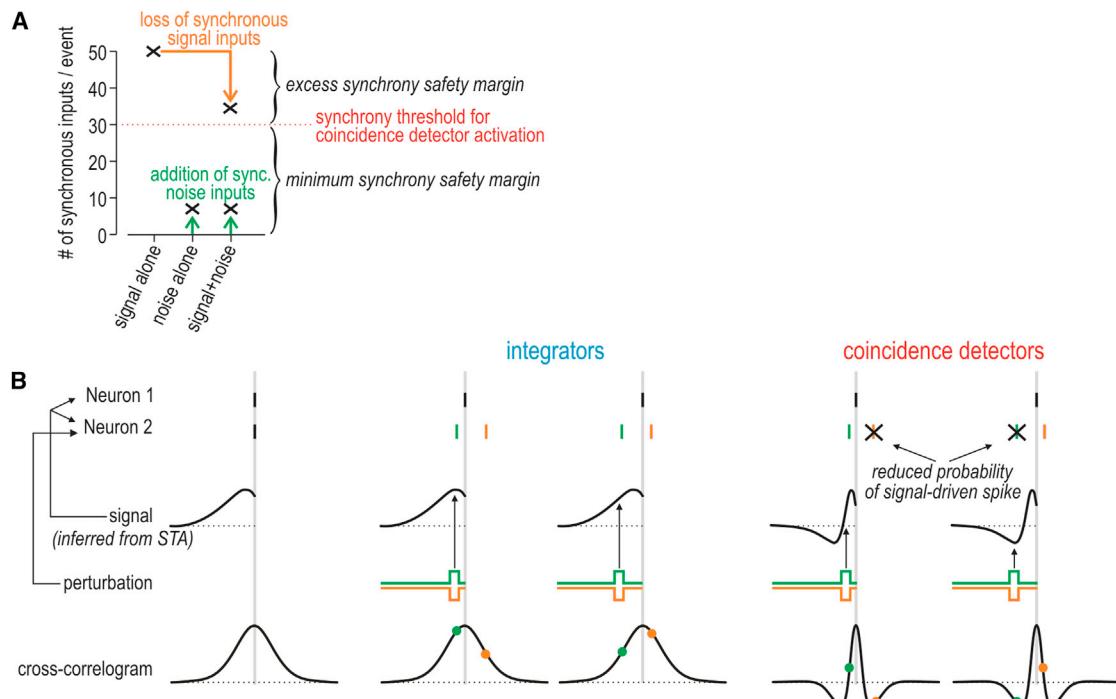


Figure 7. Robustness of Synchrony Transfer to Noise

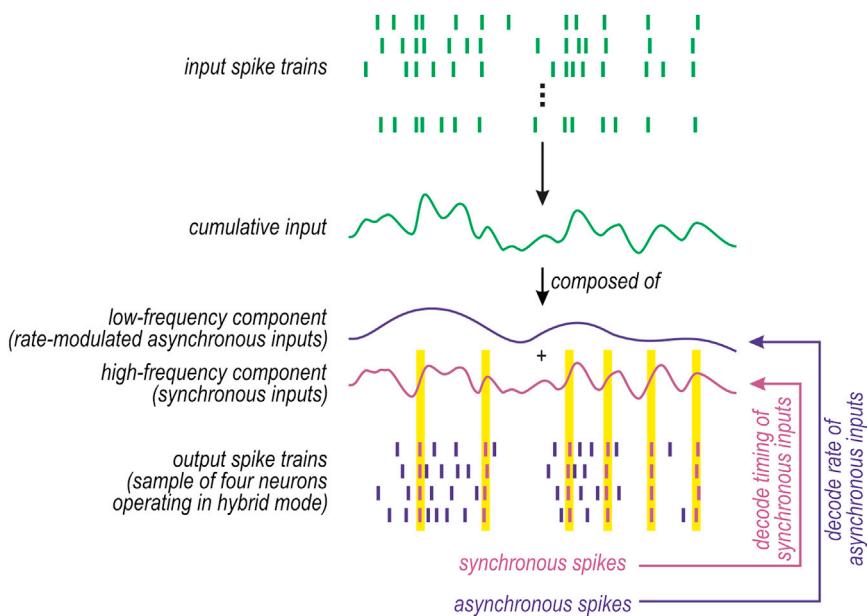
(A) Graph depicts synchronous input to a hypothetical postsynaptic coincidence detector that requires 30 synchronous inputs for activation. Without noise, the signal coactivates 50 presynaptic neurons. This means that up to 20 synchronous inputs could fail to occur (e.g., because of the effects of noise) without compromising activation of the postsynaptic neuron—this constitutes the excess synchrony safety margin. On the other hand, noise would have to coactivate at least 30 presynaptic neurons in order to activate the postsynaptic neuron, which is unlikely—this constitutes the minimum synchrony safety margin. The former safety margin reduces false negatives, whereas the latter reduces false positives with respect to correctly detecting the input signal. Integrators, by definition, have a lower synchrony threshold, which implies a smaller minimum synchrony safety margin.

(B) If two neurons spike more synchronously than expected by chance, they probably receive common input (signal) and we can infer the shape of that input based on the STA. Furthermore, if one neuron spikes, the CCG tells us the probability that the other neuron will spike. If neuron 2 receives a brief perturbation, its spike (shown in the same color as the perturbation) is jittered relative to the spike in the other neuron. In an integrator, because the CCG peak is so broad, a jittered spike will still tend to fall near the peak of the CCG (as shown by colored dots). In a coincidence detector, by comparison, even moderate jittering can shift the timing of the anticipated spike such that it coincides with one of the troughs surrounding the narrow peak of the multiphasic CCG, which implies that the probability of spiking falls to below-chance levels and that the spike will probably be “lost.” Thus, the spike initiation dynamics that are characteristic of coincidence detectors implement a *quality control mechanism*, wherein precision is maintained at the expense of reliability. The CCG troughs also ensure that tightly synchronized spikes are clearly distinguishable from asynchronous spikes because the probability of loosely synchronized spiking is very low.

spikes. In an integrator, the timing of the perturbation relative to the broad monophasic STA is relatively unimportant in this regard; in a coincidence detector, on the other hand, timing of the perturbation relative to the narrow biphasic STA has important consequences. The reduced probability of signal-driven spiking is most easily understood from the CCG, which shows the probability that neuron 2 will spike at times shortly before or after the spike in neuron 1. If a perturbation in neuron 2 jitters the anticipated signal-driven spike such that its timing coincides with either trough (negative phase) of the CCG, the probability of that spike occurring will be reduced to below-chance levels. In other words, noise is more likely to cause “lost” spikes than to cause strongly jittered spikes in coincidence detectors; the signal-driven spikes that remain will be temporally precise and therefore well synchronized. This quality control mechanism, which trades off reliability for precision, makes sense if an excess synchrony safety margin can accommodate the lost spikes.

Compared with the broad CCG characteristic of integrators, the narrow peak of coincidence detector CCGs indicates more precise synchronization. Furthermore, the adjacent troughs

seen in coincidence detector CCGs indicate correlated quiescence around the synchronous spikes; in other words, if neuron 2 does not spike within a couple of milliseconds of the spike in neuron 1 (during the CCG peak), it is less likely than chance to spike at slightly longer times (during the CCG troughs). Those troughs thus represent a boundary separating synchronous input-driven spikes from asynchronous input-driven spikes: the former are well synchronized, the latter are asynchronous, and there are few marginally synchronized spikes whose origin is ambiguous. Correctly identifying synchronous and asynchronous output spikes is important inasmuch as it can allow a decoder to distinguish spikes driven by a common signal from those driven by independent noise: the former are synchronous, whereas the latter are not. Similarly, it would allow a decoder to distinguish spikes driven by a common synchrony-encoded signal from those driven by a common rate-encoded signal: the former are synchronous, whereas the latter are not (which is not to exclude rate comodulation). The last point leads to the idea of multiplexing, but first, we must compare our claims against quantitative analysis of synchrony transfer.

**Figure 8. Multiplexed Coding**

Top rasters depict input comprising synchronous inputs plus rate-modulated asynchronous inputs. Bottom rasters depict output spike trains in four postsynaptic neurons operating in hybrid mode. Synchronous inputs elicit synchronous output spikes (purple), whereas rate-modulated asynchronous inputs elicit rate-modulated asynchronous output spikes (blue). By comparison, pure coincidence detectors would not respond to the asynchronous inputs (see Figure 1A) and pure integrators would not respond synchronously to synchronous inputs because of their rate-modulated asynchronous spiking (see Figure 1D).

When measured synchrony transfer is compared against the synchrony transfer predicted by reverse correlation analysis, output correlation among idealized integrators is accounted for by the first-order prediction (based on the STA), whereas coincidence detectors spike more synchronously than expected (Hong et al., 2012). “Excess” or unpredicted output correlation among coincidence detectors is concentrated at the center of the CCG (see Figure 6B), consistent with a failure of the STA to predict highly synchronized spiking that can be corrected by incorporating STC-based analysis. Those results speak to the importance of the rate of change of stimulus intensity in eliciting precisely synchronized spiking. Although rather obvious, that conclusion can be overlooked if oversimplified neuron models are used. Hong et al. (2012) found that pyramidal neurons were sensitive to stimulus variance in both the low- and high-conductance states and were simply more sensitive in the latter, consistent with operation in the midrange of the operating mode continuum. One should note that the comparison between predicted and measured cross-correlation was conducted using a broad range of stimulus intensities and noise conditions, the implication being that stimulus-dependent synchrony can persist despite stimulus-dependent modulation of the mean spike rate and can be properly analyzed for different stimulus parameters.

From Hybrid Operating Mode to Multiplexed Coding

By not being optimized for integration or coincidence detection, pyramidal neurons exhibit traits of both operating modes and could, therefore, be said to use a hybrid mode. This raises the question of whether a hybrid operating mode conveys benefits that justify the lack of specialization. We propose that a hybrid operating mode allows rate and synchrony codes to be multiplexed (Figure 2). Multiplexing refers to the transmission of more than one signal via a single communication channel and can increase information capacity (Lathi and Ding, 2009). Single

neurons in sensory systems have been shown to achieve multiplexing via temporal scale (frequency) division, wherein different signals are allocated to pass bands that span nonoverlapping frequencies (for review, see Panzeri et al., 2010). In the scenario considered here, synchrony-encoded signals (with power concentrated at high frequencies) are encoded by synchronous spiking, whereas asynchronous rate-encoded signals (with power concentrated at lower frequencies) are encoded by asynchronous rate-modulated spiking (Figure 8). The distinctly represented signals can coexist if synchrony transfer is robust to rate-modulated spiking. The safety margins and spike timing quality control mechanism described in Figure 7 represent biologically straightforward ways to maintain the distinction between synchronous and asynchronous spikes; in engineering terms, those mechanisms could be said to implement guard bands that separate the two pass bands.

Past studies have demonstrated rate coding multiplexed with temporal coding that depends on intrinsically generated network oscillations (Friedrich et al., 2004; Huxter et al., 2003; Mazzoni et al., 2011). Our proposed form of multiplexing more closely matches that described by Riehle et al. (1997) in the motor cortex and by Steinmetz et al. (2000) in the somatosensory cortex (see also Estebanez et al., 2012), where transient synchronization occurs independently of rate modulation but in relation to external and internal events, including attention. This form of multiplexing is also supported by our observation that precise synchrony can exist over a broad range of spike rates driven by different mean stimulus intensities (Hong et al., 2012). One potential argument against multiplexing is that recorded spike trains tend to exhibit only weak pairwise correlations. However, when cross-correlating the output spike trains of two neurons that are part of a multiplexing set—indeed, not all cross-correlated cell pairs will participate in the same set—synchronous spikes may occur only rarely compared with asynchronous spikes. This “dilution” will result in small cross-correlation values, but this does not rule out that precisely synchronized spikes occur, it simply means that those synchronous spikes are well hidden and necessitate careful analysis (Grün, 2009). We predict that synchrony-encoded signaling requires higher-order correlations—that synchrony among n neurons is greater than extrapolated from pairwise correlations—in order to support an excess

synchrony safety margin. Indeed, despite being difficult to quantify (see Staude et al., 2010), such correlations do exist (Ohiorenhan et al., 2010; Shimazaki et al., 2012; Yu et al., 2011).

A set of hybrid mode neurons can discriminate between a shared synchrony-encoded signal and independent noise by responding synchronously to the former and asynchronously to the latter. The same set can discriminate between a shared synchrony-encoded signal and a shared asynchronous rate-encoded signal by, again, responding synchronously to the former and asynchronously to the latter. Distinguishing between the rate-encoded signal and noise relies on noise being independent across the neurons so that it can be averaged out. Shared noise thus represents a problem for rate coding and synchrony coding; in that regard, both coding strategies could benefit from decorrelation mechanisms such as balanced excitation/inhibition (Renart et al., 2010) and inhibitory feedback (Tetzlaff et al., 2012), the only caveat being that signal-dependent correlations must persist while spurious correlations arising from noise are eliminated. This may come down to signal-dependent correlations being of higher order than noise-based correlations, consistent with the minimum synchrony safety margin.

Conclusions and Future Directions

Spike initiation dynamics differ between neurons and can be modulated within a given neuron, e.g., by changes in the total membrane conductance. Those dynamics represent how a multitude of membrane currents interact to control spike initiation. Although we have focused here on how spike initiation dynamics affect stimulus-driven synchrony within a feedforward network, these same dynamics are known to affect synchronization and oscillations within recurrently connected networks (Hansel et al., 1995). By restricting which inputs elicit spikes and which do not, spike initiation dynamics confer tuning to different stimulus parameters. Coincidence detectors function as band-pass filters that are tuned to high-input frequencies whose power represents the degree of input synchrony. Integrators function as low-pass filters tuned to lower frequencies and are thus relatively untuned with respect to synchrony. Real pyramidal neurons function somewhere in between.

Rapidly fluctuating (synchronous) inputs can produce precisely timed spikes within a single neuron, which translates into synchronous spiking across a set of neurons who share that input. In this respect, both integrators and coincidence detectors can respond to synchronous input with synchronous output. However, the two operating modes differ in how robust that output synchrony is to background noise (Figures 1C and 7) and to variations in firing rate (Figures 1D and 3B). Coincidence detectors transfer synchrony more robustly because of their spike initiation dynamics. For this reason, operating mode is best defined according to synchrony transfer.

Furthermore, operating mode must be treated as a continuum if we are to accurately describe the functioning of real neurons. Indeed, the longstanding debate over whether pyramidal neurons operate as integrators or coincidence detectors can be resolved by agreeing that they exhibit traits of both; moreover, those traits are modulated by factors like conductance state and are variably manifested by stimuli with different spectral properties. Once this is recognized, it becomes obvious that

pyramidal neurons are suboptimal when it comes to integration or coincidence detection and, by extension, that they are suboptimal at rate and synchrony coding. However, a hybrid operating mode—one that exploits elements of both integration and coincidence detection—may enable multiplexing of rate and synchrony coding, thereby allowing pyramidal neurons to achieve higher total information capacity than if they used one or the other code optimally.

Several issues arise from this Perspective. For instance, which neuron models can capture the essential differences between integrator and coincidence detector operating mode? Conductance-based neuron models can exhibit either operating mode based on parameter values (Lundstrom et al., 2008; Prescott et al., 2008a). This is similarly true for more sophisticated integrate-and-fire (IF) models such as the adaptive exponential IF model (Brette and Gerstner, 2005; for review, see Brunel, 2010). In principle, stimulus-dependent variations in the voltage trajectory toward threshold can be replaced with stimulus-dependent variations in threshold (Yamauchi et al., 2011). What is important is that the model includes different timescales so that intrinsic processes can interact with timescales present in the input, thus enabling inputs with power at lower or higher frequencies to preferentially elicit spikes. In this regard, the STA is invaluable in describing how stimulus properties and intrinsic neuron properties interact. Rather than pronouncing here on which models succeed or fail to capture different operating modes, we recommend that models be tested by measuring their STA under a broad range of stimulus conditions.

Beyond determining which models are most appropriate, it is important to experimentally determine where different types of neurons fall along the operating mode continuum, whether the population is tightly or broadly distributed along the continuum, etc. Like for models, the STA is a valuable descriptor of neuronal response properties. For neurons falling within the middle range, can they operate in a hybrid mode and achieve multiplexed coding under certain stimulus conditions? Under what stimulus conditions? Another broad and important set of questions includes how neurons operating in different modes function within different network architectures.

To conclude, spike initiation dynamics regulate synchrony transfer properties, and synchrony transfer properties regulate network coding strategies; therefore, spike initiation dynamics regulate network coding strategies. An accurate and complete understanding of network coding demands that we give greater consideration to neuronal properties, especially to spike initiation dynamics.

ACKNOWLEDGMENTS

This work was supported by NIH grant NS076706 (S.A.P.) and the Okinawa Institute of Science and Technology School Corporation (E.D.S.). S.A.P. is also a Rita Allen Scholar in Pain and the 53rd Mallinckrodt Scholar. We thank Dan Simons for his constructive feedback.

REFERENCES

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

- Agüera y Arcas, B., and Fairhall, A.L. (2003). What causes a neuron to spike? *Neural Comput.* 15, 1789–1807.
- Agüera y Arcas, B., Fairhall, A.L., and Bialek, W. (2003). Computation in a single neuron: Hodgkin and Huxley revisited. *Neural Comput.* 15, 1715–1749.
- Alonso, J.M., Usrey, W.M., and Reid, R.C. (1996). Precisely correlated firing in cells of the lateral geniculate nucleus. *Nature* 383, 815–819.
- Arsiero, M., Lüscher, H.R., Lundstrom, B.N., and Giugliano, M. (2007). The impact of input fluctuations on the frequency-current relationships of layer 5 pyramidal neurons in the rat medial prefrontal cortex. *J. Neurosci.* 27, 3274–3284.
- Averbeck, B.B., Latham, P.E., and Pouget, A. (2006). Neural correlations, population coding and computation. *Nat. Rev. Neurosci.* 7, 358–366.
- Azouz, R., and Gray, C.M. (2000). Dynamic spike threshold reveals a mechanism for synaptic coincidence detection in cortical neurons in vivo. *Proc. Natl. Acad. Sci. USA* 97, 8110–8115.
- Azouz, R., and Gray, C.M. (2003). Adaptive coincidence detection and dynamic gain control in visual cortical neurons in vivo. *Neuron* 37, 513–523.
- Barlow, H.B. (1961). Possible principles underlying the transformation of sensory messages. In *Sensory Communication*, W. Rosenblith, ed. (Cambridge, MA: MIT Press).
- Barreiro, A.K., Shea-Brown, E., and Thilo, E.L. (2010). Time scales of spike-train correlation for neural oscillators with common drive. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 81, 011916.
- Barreiro, A.K., Thilo, E.L., and Shea-Brown, E. (2012). A-current and type I/type II transition determine collective spiking from common input. *J. Neurophysiol.* 108, 1631–1645.
- Bernander, O., Douglas, R.J., Martin, K.A., and Koch, C. (1991). Synaptic background activity influences spatiotemporal integration in single pyramidal cells. *Proc. Natl. Acad. Sci. USA* 88, 11569–11573.
- Berry, M.J., 2nd, and Meister, M. (1998). Refractoriness and neural precision. *J. Neurosci.* 18, 2200–2211.
- Billimoria, C.P., DiCaprio, R.A., Birmingham, J.T., Abbott, L.F., and Marder, E. (2006). Neuromodulation of spike-timing precision in sensory neurons. *J. Neurosci.* 26, 5910–5919.
- Borg-Graham, L.J., Monier, C., and Frégnac, Y. (1998). Visual input evokes transient and strong shunting inhibition in visual cortical neurons. *Nature* 393, 369–373.
- Borst, A., and Theunissen, F.E. (1999). Information theory and neural coding. *Nat. Neurosci.* 2, 947–957.
- Brette, R. (2012). Computing with neural synchrony. *PLoS Comput. Biol.* 8, e1002561.
- Brette, R., and Gerstner, W. (2005). Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. *J. Neurophysiol.* 94, 3637–3642.
- Brunel, N. (2010). Modeling point neurons: from Hodgkin-Huxley to integrate-and-fire. In *Computational Modeling Methods for Neuroscientists*, E. De Schutter, ed. (Cambridge, MA: MIT Press).
- Bruno, R.M. (2011). Synchrony in sensation. *Curr. Opin. Neurobiol.* 21, 701–708.
- Bruno, R.M., and Sakmann, B. (2006). Cortex is driven by weak but synchronously active thalamocortical synapses. *Science* 312, 1622–1627.
- Bryant, H.L., and Segundo, J.P. (1976). Spike initiation by transmembrane current: a white-noise analysis. *J. Physiol.* 260, 279–314.
- Burak, Y., Lewallen, S., and Sompolinsky, H. (2009). Stimulus-dependent correlations in threshold-crossing spiking neurons. *Neural Comput.* 21, 2269–2308.
- Cecchi, G.A., Sigman, M., Alonso, J.M., Martínez, L., Chialvo, D.R., and Mugnaini, M.O. (2000). Noise in neurons is message dependent. *Proc. Natl. Acad. Sci. USA* 97, 5557–5561.
- Cohen, M.R., and Kohn, A. (2011). Measuring and interpreting neuronal correlations. *Nat. Neurosci.* 14, 811–819.
- Constantinople, C.M., and Bruno, R.M. (2011). Effects and mechanisms of wakefulness on local cortical networks. *Neuron* 69, 1061–1068.
- Cudmore, R.H., Fronzaroli-Molinieres, L., Giraud, P., and Debanne, D. (2010). Spike-time precision and network synchrony are controlled by the homeostatic regulation of the D-type potassium current. *J. Neurosci.* 30, 12885–12895.
- Dan, Y., Alonso, J.M., Usrey, W.M., and Reid, R.C. (1998). Coding of visual information by precisely correlated spikes in the lateral geniculate nucleus. *Nat. Neurosci.* 1, 501–507.
- de la Rocha, J., Doiron, B., Shea-Brown, E., Josić, K., and Reyes, A. (2007). Correlation between neural spike trains increases with firing rate. *Nature* 448, 802–806.
- deCharms, R.C. (1998). Information coding in the cortex by independent or coordinated populations. *Proc. Natl. Acad. Sci. USA* 95, 15166–15168.
- deCharms, R.C., and Merzenich, M.M. (1996). Primary cortical representation of sounds by the coordination of action-potential timing. *Nature* 381, 610–613.
- Destexhe, A., and Paré, D. (1999). Impact of network activity on the integrative properties of neocortical pyramidal neurons in vivo. *J. Neurophysiol.* 81, 1531–1547.
- Destexhe, A., Rudolph, M., Fellous, J.M., and Sejnowski, T.J. (2001). Fluctuating synaptic conductances recreate in vivo-like activity in neocortical neurons. *Neuroscience* 107, 13–24.
- Destexhe, A., Rudolph, M., and Paré, D. (2003). The high-conductance state of neocortical neurons in vivo. *Nat. Rev. Neurosci.* 4, 739–751.
- DeWeese, M.R., and Zador, A.M. (2006). Non-Gaussian membrane potential dynamics imply sparse, synchronous activity in auditory cortex. *J. Neurosci.* 26, 12206–12218.
- Diesmann, M., Gewaltig, M.O., and Aertsen, A. (1999). Stable propagation of synchronous spiking in cortical neural networks. *Nature* 402, 529–533.
- Ecker, A.S., Berens, P., Keliris, G.A., Bethge, M., Logothetis, N.K., and Tolias, A.S. (2010). Decorrelated neuronal firing in cortical microcircuits. *Science* 327, 584–587.
- Ermentrout, B. (1996). Type I membranes, phase resetting curves, and synchrony. *Neural Comput.* 8, 979–1001.
- Ermentrout, G.B., Galán, R.F., and Urban, N.N. (2007). Relating neural dynamics to neural coding. *Phys. Rev. Lett.* 99, 248103.
- Ermentrout, G.B., Galán, R.F., and Urban, N.N. (2008). Reliability, synchrony and noise. *Trends Neurosci.* 31, 428–434.
- Estebanez, L., El Boustani, S., Destexhe, A., and Shultz, D.E. (2012). Correlated input reveals coexisting coding schemes in a sensory cortex. *Nat. Neurosci.* 15, 1691–1699.
- Fernandez, F.R., Broicher, T., Truong, A., and White, J.A. (2011). Membrane voltage fluctuations reduce spike frequency adaptation and preserve output gain in CA1 pyramidal neurons in a high-conductance state. *J. Neurosci.* 31, 3880–3893.
- Fricker, D., and Miles, R. (2000). EPSP amplification and the precision of spike timing in hippocampal neurons. *Neuron* 28, 559–569.
- Friedrich, R.W., Habermann, C.J., and Laurent, G. (2004). Multiplexing using synchrony in the zebrafish olfactory bulb. *Nat. Neurosci.* 7, 862–871.
- Galán, R.F., Ermentrout, G.B., and Urban, N.N. (2008). Optimal time scale for spike-time reliability: theory, simulations, and experiments. *J. Neurophysiol.* 99, 277–283.
- Gastrein, P., Campanac, E., Gasselin, C., Cudmore, R.H., Bialowas, A., Carlier, E., Fronzaroli-Molinieres, L., Ankrum, N., and Debanne, D. (2011). The role of hyperpolarization-activated cationic current in spike-time precision and intrinsic resonance in cortical neurons in vitro. *J. Physiol.* 589, 3753–3773.
- Gawne, T.J., and Richmond, B.J. (1993). How independent are the messages carried by adjacent inferior temporal cortical neurons? *J. Neurosci.* 13, 2758–2771.

- Goedeke, S., and Diesmann, M. (2008). The mechanism of synchronization in feed-forward neuronal networks. *New J. Phys.* 10, 015007.
- Goldberg, J.A., Rokni, U., Boraud, T., Vaadia, E., and Bergman, H. (2004). Spike synchronization in the cortex/basal-ganglia networks of Parkinsonian primates reflects global dynamics of the local field potentials. *J. Neurosci.* 24, 6003–6010.
- Grün, S. (2009). Data-driven significance estimation for precise spike correlation. *J. Neurophysiol.* 101, 1126–1140.
- Haider, B., Häusser, M., and Carandini, M. (2013). Inhibition dominates sensory responses in the awake cortex. *Nature* 493, 97–100.
- Hansel, D., Mato, G., and Meunier, C. (1995). Synchrony in excitatory neural networks. *Neural Comput.* 7, 307–337.
- Higgs, M.H., Slee, S.J., and Spain, W.J. (2006). Diversity of gain modulation by noise in neocortical neurons: regulation by the slow afterhyperpolarization conductance. *J. Neurosci.* 26, 8787–8799.
- Hodgkin, A.L. (1948). The local electric changes associated with repetitive action in a non-medullated axon. *J. Physiol.* 107, 165–181.
- Hong, S., Agüera y Arcas, B., and Fairhall, A.L. (2007). Single neuron computation: from dynamical system to feature detector. *Neural Comput.* 19, 3133–3172.
- Hong, S., Ratté, S., Prescott, S.A., and De Schutter, E. (2012). Single neuron firing properties impact correlation-based population coding. *J. Neurosci.* 32, 1413–1428.
- Huxter, J., Burgess, N., and O'Keefe, J. (2003). Independent rate and temporal coding in hippocampal pyramidal cells. *Nature* 425, 828–832.
- 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. *Science* 304, 559–564.
- Isaacson, J.S., and Scanziani, M. (2011). How inhibition shapes cortical activity. *Neuron* 72, 231–243.
- Izhikevich, E.M. (2007). *Dynamical Systems in Neuroscience* (Cambridge, MA: MIT Press).
- JadHAV, S.P., Wolfe, J., and Feldman, D.E. (2009). Sparse temporal coding of elementary tactile features during active whisker sensation. *Nat. Neurosci.* 12, 792–800.
- Josić, K., Shea-Brown, E., Doiron, B., and de la Rocha, J. (2009). Stimulus-dependent correlations and population codes. *Neural Comput.* 21, 2774–2804.
- Kenyon, G.T., Theiler, J., George, J.S., Travis, B.J., and Marshak, D.W. (2004). Correlated firing improves stimulus discrimination in a retinal model. *Neural Comput.* 16, 2261–2291.
- Kohn, A., and Smith, M.A. (2005). Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. *J. Neurosci.* 25, 3661–3673.
- König, P., Engel, A.K., and Singer, W. (1996). Integrator or coincidence detector? The role of the cortical neuron revisited. *Trends Neurosci.* 19, 130–137.
- Kumar, A., Rotter, S., and Aertsen, A. (2008). Conditions for propagating synchronous spiking and asynchronous firing rates in a cortical network model. *J. Neurosci.* 28, 5268–5280.
- Kumar, A., Rotter, S., and Aertsen, A. (2010). Spiking activity propagation in neuronal networks: reconciling different perspectives on neural coding. *Nat. Rev. Neurosci.* 11, 615–627.
- Lampl, I., Reichova, I., and Ferster, D. (1999). Synchronous membrane potential fluctuations in neurons of the cat visual cortex. *Neuron* 22, 361–374.
- Lathi, B.P., and Ding, Z. (2009). *Modern Digital and Analog Communication Systems*, Fourth Edition (Oxford: Oxford University Press).
- London, M., Roth, A., Beeren, L., Häusser, M., and Latham, P.E. (2010). Sensitivity to perturbations *in vivo* implies high noise and suggests rate coding in cortex. *Nature* 466, 123–127.
- Lundstrom, B.N., Hong, S., Higgs, M.H., and Fairhall, A.L. (2008). Two computational regimes of a single-compartment neuron separated by a planar boundary in conductance space. *Neural Comput.* 20, 1239–1260.
- Lundstrom, B.N., Famulare, M., Sorensen, L.B., Spain, W.J., and Fairhall, A.L. (2009). Sensitivity of firing rate to input fluctuations depends on time scale separation between fast and slow variables in single neurons. *J. Comput. Neurosci.* 27, 277–290.
- Mainen, Z.F., and Sejnowski, T.J. (1995). Reliability of spike timing in neocortical neurons. *Science* 268, 1503–1506.
- Mason, A., Nicoll, A., and Stratford, K. (1991). Synaptic transmission between individual pyramidal neurons of the rat visual cortex *in vitro*. *J. Neurosci.* 11, 72–84.
- Mastronarde, D.N. (1989). Correlated firing of retinal ganglion cells. *Trends Neurosci.* 12, 75–80.
- Mato, G., and Samengo, I. (2008). Type I and type II neuron models are selectively driven by differential stimulus features. *Neural Comput.* 20, 2418–2440.
- Mazurek, M.E., and Shadlen, M.N. (2002). Limits to the temporal fidelity of cortical spike rate signals. *Nat. Neurosci.* 5, 463–471.
- Mazzoni, A., Brunel, N., Cavallari, S., Logothetis, N.K., and Panzeri, S. (2011). Cortical dynamics during naturalistic sensory stimulations: experiments and models. *J. Physiol. Paris* 105, 2–15.
- Meister, M., Lagnado, L., and Baylor, D.A. (1995). Concerted signaling by retinal ganglion cells. *Science* 270, 1207–1210.
- Montani, F., Kohn, A., Smith, M.A., and Schultz, S.R. (2007). The role of correlations in direction and contrast coding in the primary visual cortex. *J. Neurosci.* 27, 2338–2348.
- Nowak, L.G., Sanchez-Vives, M.V., and McCormick, D.A. (1997). Influence of low and high frequency inputs on spike timing in visual cortical neurons. *Cereb. Cortex* 7, 487–501.
- Ohiorhenuan, I.E., Mechler, F., Purpura, K.P., Schmid, A.M., Hu, Q., and Victor, J.D. (2010). Sparse coding and high-order correlations in fine-scale cortical networks. *Nature* 466, 617–621.
- Ostojic, S., Brunel, N., and Hakim, V. (2009). How connectivity, background activity, and synaptic properties shape the cross-correlation between spike trains. *J. Neurosci.* 29, 10234–10253.
- Panzeri, S., Brunel, N., Logothetis, N.K., and Kayser, C. (2010). Sensory neural codes using multiplexed temporal scales. *Trends Neurosci.* 33, 111–120.
- Perkel, D.H., and Bullock, T.H. (1968). Neural coding. *Neurosci. Res. Prog. Sum.* 3, 405–527.
- Pouille, F., and Scanziani, M. (2001). Enforcement of temporal fidelity in pyramidal cells by somatic feed-forward inhibition. *Science* 293, 1159–1163.
- Pouille, F., Marin-Burgin, A., Adesnik, H., Atallah, B.V., and Scanziani, M. (2009). Input normalization by global feedforward inhibition expands cortical dynamic range. *Nat. Neurosci.* 12, 1577–1585.
- Poulet, J.F., and Petersen, C.C. (2008). Internal brain state regulates membrane potential synchrony in barrel cortex of behaving mice. *Nature* 454, 881–885.
- Prescott, S.A., and De Koninck, Y. (2005). Integration time in a subset of spinal lamina I neurons is lengthened by sodium and calcium currents acting synergistically to prolong subthreshold depolarization. *J. Neurosci.* 25, 4743–4754.
- Prescott, S.A., and Sejnowski, T.J. (2008). Spike-rate coding and spike-time coding are affected oppositely by different adaptation mechanisms. *J. Neurosci.* 28, 13649–13661.
- Prescott, S.A., De Koninck, Y., and Sejnowski, T.J. (2006). Nonlinear interaction between shunting and adaptation controls a switch between integration and coincidence detection in pyramidal neurons. *J. Neurosci.* 26, 9084–9097.
- Prescott, S.A., De Koninck, Y., and Sejnowski, T.J. (2008a). Biophysical basis for three distinct dynamical mechanisms of action potential initiation. *PLoS Comput. Biol.* 4, e1000198.

- Prescott, S.A., Ratté, S., De Koninck, Y., and Sejnowski, T.J. (2008b). Pyramidal neurons switch from integrators in vitro to resonators under in vivo-like conditions. *J. Neurophysiol.* 100, 3030–3042.
- Reich, D.S., Mechler, F., and Victor, J.D. (2001). Independent and redundant information in nearby cortical neurons. *Science* 294, 2566–2568.
- Renart, A., de la Rocha, J., Bartho, P., Hollender, L., Parga, N., Reyes, A., and Harris, K.D. (2010). The asynchronous state in cortical circuits. *Science* 327, 587–590.
- Reyes, A.D. (2003). Synchrony-dependent propagation of firing rate in iteratively constructed networks in vitro. *Nat. Neurosci.* 6, 593–599.
- Riehle, A., Grün, S., Diesmann, M., and Aertsen, A. (1997). Spike synchronization and rate modulation differentially involved in motor cortical function. *Science* 278, 1950–1953.
- Rieke, F., Warland, D., de Ruyter van Steveninck, R.R., and Bialek, W. (1997). *Spikes: Exploring the Neural Code* (Cambridge, MA: MIT Press).
- Rinzel, J., and Ermentrout, G.B. (1998). Analysis of neural excitability and oscillations. In *Methods in Neuronal Modeling: From Ions to Networks*, C. Koch and I. Segev, eds. (Cambridge, MA: MIT Press), pp. 251–291.
- Rossant, C., Leijon, S., Magnusson, A.K., and Brette, R. (2011). Sensitivity of noisy neurons to coincident inputs. *J. Neurosci.* 31, 17193–17206.
- Rudolph, M., and Destexhe, A. (2003). Tuning neocortical pyramidal neurons between integrators and coincidence detectors. *J. Comput. Neurosci.* 14, 239–251.
- Sáez, I., and Friedlander, M.J. (2009). Synaptic output of individual layer 4 neurons in guinea pig visual cortex. *J. Neurosci.* 29, 4930–4944.
- Salinas, E., and Sejnowski, T.J. (2000). Impact of correlated synaptic input on output firing rate and variability in simple neuronal models. *J. Neurosci.* 20, 6193–6209.
- Salinas, E., and Sejnowski, T.J. (2001). Correlated neuronal activity and the flow of neural information. *Nat. Rev. Neurosci.* 2, 539–550.
- Sayer, R.J., Redman, S.J., and Andersen, P. (1989). Amplitude fluctuations in small EPSPs recorded from CA1 pyramidal cells in the guinea pig hippocampal slice. *J. Neurosci.* 9, 840–850.
- Schneidman, E., Berry, M.J., 2nd, Segev, R., and Bialek, W. (2006). Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature* 440, 1007–1012.
- Schnitzer, M.J., and Meister, M. (2003). Multineuronal firing patterns in the signal from eye to brain. *Neuron* 37, 499–511.
- Schreiber, S., Fellous, J.M., Tiesinga, P., and Sejnowski, T.J. (2004). Influence of ionic conductances on spike timing reliability of cortical neurons for suprathreshold rhythmic inputs. *J. Neurophysiol.* 91, 194–205.
- Schultze-Kraft, M., Diesmann, M., Grün, S., and Helias, M. (2013). Noise suppression and surplus synchrony by coincidence detection. *PLoS Comput. Biol.* 9, e1002904.
- Schwartz, O., Pillow, J.W., Rust, N.C., and Simoncelli, E.P. (2006). Spike-triggered neural characterization. *J. Vis.* 6, 484–507.
- Shadlen, M.N., and Newsome, W.T. (1998). The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J. Neurosci.* 18, 3870–3896.
- Sharifi, N., Benda, J., and Lindner, B. (2013). Information filtering by synchronous spikes in a neural population. *J. Comput. Neurosci.* 34, 285–301.
- Shimazaki, H., Amari, S., Brown, E.N., and Grün, S. (2012). State-space analysis of time-varying higher-order spike correlation for multiple neural spike train data. *PLoS Comput. Biol.* 8, e1002385.
- Singer, W. (1999). Neuronal synchrony: a versatile code for the definition of relations? *Neuron* 24, 49–65, 111–125.
- Slee, S.J., Higgs, M.H., Fairhall, A.L., and Spain, W.J. (2005). Two-dimensional time coding in the auditory brainstem. *J. Neurosci.* 25, 9978–9988.
- Smith, M.A., and Kohn, A. (2008). Spatial and temporal scales of neuronal correlation in primary visual cortex. *J. Neurosci.* 28, 12591–12603.
- Softky, W.R., and Koch, C. (1993). The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *J. Neurosci.* 13, 334–350.
- Sompolinsky, H., Yoon, H., Kang, K., and Shamir, M. (2001). Population coding in neuronal systems with correlated noise. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 64, 051904.
- Song, S., Sjöström, P.J., Reigl, M., Nelson, S., and Chklovskii, D.B. (2005). Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS Biol.* 3, e68.
- Stanley, G.B. (2013). Reading and writing the neural code. *Nat. Neurosci.* 16, 259–263.
- Staude, B., Grün, S., and Rotter, S. (2010). Higher-order correlations in non-stationary parallel spike trains: statistical modeling and inference. *Front. Comput. Neurosci.* 4, 16.
- Steinmetz, P.N., Roy, A., Fitzgerald, P.J., Hsiao, S.S., Johnson, K.O., and Niebur, E. (2000). Attention modulates synchronized neuronal firing in primate somatosensory cortex. *Nature* 404, 187–190.
- Stevens, C.F. (1994). Neuronal communication. Cooperativity of unreliable neurons. *Curr. Biol.* 4, 268–269.
- Stevens, C.F., and Zador, A.M. (1998). Input synchrony and the irregular firing of cortical neurons. *Nat. Neurosci.* 1, 210–217.
- Stiefel, K.M., Gutkin, B.S., and Sejnowski, T.J. (2008). Cholinergic neuromodulation changes phase response curve shape and type in cortical pyramidal neurons. *PLoS ONE* 3, e3947.
- Svirskis, G., and Rinzel, J. (2003). Influence of subthreshold nonlinearities on signal-to-noise ratio and timing precision for small signals in neurons: minimal model analysis. *Network* 14, 137–150.
- Tateno, T., Harsch, A., and Robinson, H.P. (2004). Threshold firing frequency-current relationships of neurons in rat somatosensory cortex: type 1 and type 2 dynamics. *J. Neurophysiol.* 92, 2283–2294.
- Tchumatchenko, T., Malyshev, A., Geisel, T., Volgushev, M., and Wolf, F. (2010). Correlations and synchrony in threshold neuron models. *Phys. Rev. Lett.* 104, 058102.
- Temereanca, S., Brown, E.N., and Simons, D.J. (2008). Rapid changes in thalamic firing synchrony during repetitive whisker stimulation. *J. Neurosci.* 28, 11153–11164.
- Tetzlaff, T., Helias, M., Einevoll, G.T., and Diesmann, M. (2012). Decorrelation of neural-network activity by inhibitory feedback. *PLoS Comput. Biol.* 8, e1002596.
- Tsubo, Y., Takada, M., Reyes, A.D., and Fukai, T. (2007). Layer and frequency dependencies of phase response properties of pyramidal neurons in rat motor cortex. *Eur. J. Neurosci.* 25, 3429–3441.
- Wang, H.P., Spencer, D., Fellous, J.M., and Sejnowski, T.J. (2010). Synchrony of thalamocortical inputs maximizes cortical reliability. *Science* 328, 106–109.
- Yamauchi, S., Kim, H., and Shinomoto, S. (2011). Elemental spiking neuron model for reproducing diverse firing patterns and predicting precise firing times. *Front Comput Neurosci* 5, 42.
- Yu, J., and Ferster, D. (2010). Membrane potential synchrony in primary visual cortex during sensory stimulation. *Neuron* 68, 1187–1201.
- Yu, S., Yang, H., Nakahara, H., Santos, G.S., Nikolić, D., and Plenz, D. (2011). Higher-order interactions characterized in cortical activity. *J. Neurosci.* 31, 17514–17526.
- Zador, A. (1998). Impact of synaptic unreliability on the information transmitted by spiking neurons. *J. Neurophysiol.* 79, 1219–1229.

Single Neuron Firing Properties Impact Correlation-Based Population Coding

Sungho Hong,¹ Stéphanie Ratté,² Steven A. Prescott,^{2*} and Erik De Schutter^{1,3*}

¹Computational Neuroscience Unit, Okinawa Institute of Science and Technology, Okinawa 904-0411, Japan, ²Department of Neurobiology and Pittsburgh Center for Pain Research, University of Pittsburgh, Pittsburgh, Pennsylvania 15213, and ³Theoretical Neurobiology, University of Antwerp, 2610 Antwerp, Belgium

Correlated spiking has been widely observed, but its impact on neural coding remains controversial. Correlation arising from comodulation of rates across neurons has been shown to vary with the firing rates of individual neurons. This translates into rate and correlation being equivalently tuned to the stimulus; under those conditions, correlated spiking does not provide information beyond that already available from individual neuron firing rates. Such correlations are irrelevant and can reduce coding efficiency by introducing redundancy. Using simulations and experiments in rat hippocampal neurons, we show here that pairs of neurons receiving correlated input also exhibit correlations arising from precise spike-time synchronization. Contrary to rate comodulation, spike-time synchronization is unaffected by firing rate, thus enabling synchrony- and rate-based coding to operate independently. The type of output correlation depends on whether intrinsic neuron properties promote integration or coincidence detection: “ideal” integrators (with spike generation sensitive to stimulus mean) exhibit rate comodulation, whereas ideal coincidence detectors (with spike generation sensitive to stimulus variance) exhibit precise spike-time synchronization. Pyramidal neurons are sensitive to both stimulus mean and variance, and thus exhibit both types of output correlation proportioned according to which operating mode is dominant. Our results explain how different types of correlations arise based on how individual neurons generate spikes, and why spike-time synchronization and rate comodulation can encode different stimulus properties. Our results also highlight the importance of neuronal properties for population-level coding insofar as neural networks can employ different coding schemes depending on the dominant operating mode of their constituent neurons.

Introduction

Neurons in many brain areas exhibit correlated spiking but the role of those correlations remains controversial (Singer, 1993; Zohary et al., 1994; Engel et al., 1997; Gerstner et al., 1997; Shadlen and Movshon, 1999; Treisman, 1999; Salinas and Sejnowski, 2001; Palanca and DeAngelis, 2005; Averbeck et al., 2006; Schneidman et al., 2006; Wolfe et al., 2010). Noise correlations are generally thought to degrade coding efficiency (Averbeck et al., 2006) [with exceptions (Cafaro and Rieke, 2010)], but signal-dependent correlations could conceivably carry information. However, the feasibility of correlation-based coding has been called into question by the observation that output correlation varies with firing rate despite no change in input correlation (de la Rocha et al., 2007). If such a correlation–rate relationship

always existed, input correlation could not be unambiguously decoded from output correlation, and transferred correlations would become meaningless (see Fig. 1). Importantly, correlations range from precise spike-time synchronization (on a millisecond timescale) to coarse rate comodulation (on a timescale up to seconds). We hypothesized that different types of correlation may differ fundamentally in how they are generated and what information they convey.

Propagation of correlated spiking depends on how individual neurons respond to correlated input (sensitivity to correlation) and whether groups of neurons respond with correlated output (transfer of correlation) such that postsynaptic neurons themselves receive correlated input (Abeles, 1991; Aertsen et al., 1996; Reyes, 2003). With respect to sensitivity to correlation, a critical factor is whether neurons operate as integrators or coincidence detectors: integrators respond to temporally dispersed inputs, whereas coincidence detectors respond selectively to rapid depolarization caused by temporally coincident (synchronous) inputs (Abeles, 1982; König et al., 1996). Operating mode (i.e., integration vs coincidence detection) reflects interplay between stimulus kinetics and spike threshold mechanism (see Results). With respect to the transfer of correlation, integrators spike repetitively at a rate proportional to their time-averaged input, whereas coincidence detectors respond to each suprathreshold input with an isolated spike—these spiking patterns are conducive to rate and temporal coding, respectively (Mainen and Sejnowski, 1995; König

Received July 20, 2011; revised Nov. 15, 2011; accepted Nov. 28, 2011.

Author contributions: S.H., S.A.P., and E.D.S. designed research; S.H., S.R., and S.A.P. performed research; S.H. and S.A.P. analyzed data; S.H., S.R., S.A.P., and E.D.S. wrote the paper.

This work was supported by the Okinawa Institute of Science and Technology Promotion Corporation (S.H., E.D.S.) and NIH Grant NS076706 (S.A.P.). S.A.P. is also a Rita Allen Foundation Scholar in Pain and the 53rd Mallinckrodt Scholar.

The authors declare no competing financial interests.

*S.A.P. and E.D.S. contributed equally to this work.

Correspondence should be addressed to Sungho Hong, Computational Neuroscience Unit, Okinawa Institute of Science and Technology, 7542 Onna, Onna-son, Okinawa 904-0411, Japan. E-mail: shong@oist.jp.

DOI:10.1523/JNEUROSCI.3735-11.2012

Copyright © 2012 the authors 0270-6474/12/321413-16\$15.00/0

et al., 1996; Salinas and Sejnowski, 2001; Schreiber et al., 2004; Prescott et al., 2006; Prescott and Sejnowski, 2008; Tiesinga et al., 2008). Importantly, the temporal precision with which an individual neuron spikes should affect how synchronously a group of such neurons will spike to shared inputs, which bears directly on transfer of synchrony. Notably, neurons may be highly specialized for one operating mode, but most exhibit a context-dependent combination of modes (Maex et al., 2000; Rudolph and Destexhe, 2003; Prescott et al., 2006; Hong et al., 2008).

Using computer simulations and dynamic-clamp experiments, we identified different types of output correlation and investigated how transfer of each type of correlation depends on the intrinsic properties of cell pairs receiving correlated input. We show that pairs of “realistic” coincidence detectors exhibit rate comodulation and spike-time synchronization, whereas pairs of realistic integrators exhibit only rate comodulation. Synchrony, unlike rate comodulation, depends on spike-timing rather than rate. Rate- and synchrony-based coding are thus shown to operate independently.

Materials and Methods

Stimulus preparation

We constructed stimulus waveforms using the same procedure as de la Rocha et al. (2007). The stimulus $I(t)$ for a given trial was the linear summation of two Ornstein–Uhlenbeck (OU) processes (Uhlenbeck and Ornstein, 1930) described by the following:

$$I(t) = \mu + \sigma(\sqrt{c}\zeta_c(t) + \sqrt{1-c}\zeta_i(t)), \quad (1)$$

where μ and σ are the mean and SD of the stimulus, and c is the input correlation (i.e., the fraction of fluctuating input shared between neurons) (see Fig. 1A). The common component $\zeta_c(t)$ was instantiated once and applied to all trials, whereas the independent component $\zeta_i(t)$ was randomly updated for each trial. Each OU process was formed by the following:

$$d\zeta = -\frac{\zeta}{\tau} dt + N_\tau \xi \sqrt{dt}, \quad (2)$$

where $\xi(t)$ is Gaussian white noise with zero mean and unit variance. Sampling rates were 10 and 5 kHz for experiments and simulations, respectively. $N_\tau = (2/\tau)^{1/2}$ is a normalization constant that makes $\zeta(t)$ have unit variance. A correlation time $\tau = 5$ ms was used unless otherwise indicated.

Model neurons and simulation procedures

Two conductance-based neuron models were used. We modeled the integrator as a Morris–Lecar (ML) model with type 1 excitability (Prescott et al., 2008a) and the coincidence detector as a Hodgkin–Huxley low-sodium (HHLS) model with type 3 excitability (Lundstrom et al., 2008). Equations for the ML model are as follows:

$$\begin{aligned} C \frac{dV}{dt} = & -g_L(V - E_L) - g_{Na}m_\infty(V)(V - E_{Na}) \\ & - g_Kn(V - E_K) + I(t), \end{aligned} \quad (3)$$

where

$$\begin{aligned} \frac{dn}{dt} &= \frac{\phi}{\tau_n}(n_\infty - n), \\ m_\infty(V) &= 0.5(1 + \tanh(V - V_1)/V_2), \\ n_\infty(V) &= 0.5(1 + \tanh(V - V_3)/V_4), \\ \tau_n(V) &= 1/\cosh(0.5(V - V_3)/V_4). \end{aligned} \quad (4)$$

and $g_{Na} = 20$ mS/cm², $g_K = 20$ mS/cm², $g_L = 2$ mS/cm², $\phi = 0.15$, $V_1 = -1.2$ mV, $V_2 = 18$ mV, $V_3 = 0$ mV, and $V_4 = 10$ mV. The membrane capacitance per area C was 2 μ F/cm² and the surface area was 100 μ m².

Table 1. Stimulus conditions used for simulation of the integrator (ML), coincidence detector (HHLS), and FT models

Model	τ (ms)	Parameter	No. of samples	Minimum	Maximum
Integrator	5	c	5	0.1	0.5
		μ (pA)	12	345	375
		σ^2 (pA ²)	10	100	400
	50	c	5	0.1	0.5
		μ (pA)	10	330	400
		σ^2 (pA ²)	10	100	400
Coincidence	5	c	5	0.1	0.5
		μ (pA)	18	-150	200
		σ^2 (pA ²)	8	1000	2500
	50	c	5	0.1	0.5
		μ (pA)	25	-100	100
		σ^2 (pA ²)	10	400	4400
FT	5	c	5	0.1	0.5
		μ (pA)	1	0	0
		σ^2 (pA ²)	30	100	7000

Equations for the HHLS model are as follows:

$$\begin{aligned} C \frac{dV}{dt} = & -g_L(V - E_L) - g_{Na}m^3h(V - E_{Na}) \\ & - g_Kn^4(V - E_K) + I(t), \end{aligned} \quad (5)$$

with activation variables m , n , and h governed by the following:

$$\begin{aligned} \tau(V) \frac{dz}{dt} &= z_\infty(V) - z, \\ \tau_z &= \frac{1}{\alpha + \beta}, \quad z_\infty = \frac{\alpha}{\alpha + \beta}, \quad z = m, n, h. \end{aligned} \quad (6)$$

where

$$\begin{aligned} \alpha_m &= \frac{0.1(V + 40)}{1 - \exp[-0.1(V + 40)]}, \\ \beta_m &= 4 \exp[-0.556(V + 65)], \\ \alpha_h &= 0.07 \exp[-0.05(V + 65)], \\ \beta_h &= \frac{1}{1 + \exp[-0.1(V + 35)]}, \\ \alpha_n &= \frac{0.01(V + 55)}{1 - \exp[-0.1(V + 55)]}, \\ \beta_n &= 0.125 \exp[-0.0125(V + 65)]. \end{aligned} \quad (7)$$

$g_{Na} = 41$ mS/cm², $g_K = 79$ mS/cm², $g_L = 0.3$ mS/cm², and the membrane capacitance $C = 1$ μ F/cm² and the surface area was 100 μ m².

The filter-and-threshold (FT) model consisted of three components: a linear filter to transform input to voltage, a voltage threshold, and an afterhyperpolarization (AHP). For the filter, we used the time derivative of a 15-ms-long Blackman filter, which was normalized to transform an input with variance 1 pA² to an output with a variance 0.1 mV². The threshold was 1 mV and the AHP inserted for each spike had -0.5 mV amplitude and 30 ms decay time.

All simulations in conductance-based models were performed in NEURON (Hines and Carnevale, 1997). Simulations in the FT model were performed using custom Python scripts. All code will be made available on ModelDB. Each stimulus condition (c , μ , σ^2) was repeated 2–10 times for 30 min of simulated time. All (c , μ , σ^2) combinations used are summarized in Table 1. For Figure 5, the 100–200 simulation runs conducted for each model resulted in a very large amount of data, making the calculation of correlations (from up to ~40,000 pairs) computationally challenging and the results difficult to present; therefore, we selected

only pairs having the same stimulus mean, which resulted in 1000 and 2500 pairs for the integrator and coincidence detector, respectively. This criterion was not applied to experimental data (see Fig. 7) since there were fewer trials.

Slice preparation and electrophysiology

Experimental protocols were approved by the University of Pittsburgh Institutional Animal Care and Use Committee and have been previously described (Prescott et al., 2006). Briefly, adult male Sprague Dawley rats were anesthetized with intraperitoneal injection of sodium pentobarbital (50–75 mg/kg) and perfused intracardially with ice-cold oxygenated (95% O₂ and 5% CO₂) sucrose-substituted artificial CSF (ACSF) containing the following (in mm): 252 sucrose, 2.5 KCl, 2 CaCl₂, 2 MgCl₂, 10 glucose, 26 NaHCO₃, 1.25 NaH₂PO₄, and 5 kynurenic acid. The brain was rapidly removed and sectioned coronally to give 300-μm-thick slices, which were kept in normal oxygenated ACSF (126 mm NaCl instead of sucrose and without kynurenic acid) at room temperature until recording.

Slices were transferred to a recording chamber constantly perfused with oxygenated (95% O₂ and 5% CO₂) ACSF heated to 31 ± 1°C. Pyramidal neurons in the CA1 region of hippocampus were recorded in the whole-cell configuration with >70% series resistance compensation using an Axopatch 200B amplifier (Molecular Devices). Membrane potential (after correction for the liquid junction potential of 9 mV) was adjusted to -70 mV through tonic current injection. Intracellular recording solution contained the following (in mm): 125 KMeSO₄, 5 KCl, 10 HEPES, and 2 MgCl₂, 4 ATP (Sigma-Aldrich), 0.4 GTP (Sigma-Aldrich), as well as 0.1% Lucifer yellow; pH was adjusted to 7.2 with KOH. Pyramidal morphology was confirmed with epifluorescence after recording. All experiments were performed in 10 μM bicuculline methiodide (Sigma-Aldrich), 10 μM CNQX (6-cyano-7-nitroquinoxaline-2,3-dione) (Sigma-Aldrich), and 40 μM D-AP-5 (D-2-amino-5-phosphonovaleric acid) (Ascent Scientific) to block background synaptic activity.

Stimuli (see above) were injected into the recorded neurons through the patch pipette. To manipulate spike threshold mechanism (see Results), an artificial “shunt” conductance ($E_{\text{shunt}} = -70$ mV, $g_{\text{shunt}} = 10$ nS) was applied via dynamic clamp implemented with a Digidata 1200A ADC/DAC board (Molecular Devices) and DYNCLAMP2 software (Pinto et al., 2001) running on a dedicated processor as previously described (Prescott et al., 2006; Prescott and De Koninck, 2009); update rate was 10 kHz. Traces were low-pass filtered at 2 kHz and digitized at 10 kHz using a CED 1401 computer interface (Cambridge Electronic Design).

Reverse correlation analysis

For simulation and experimental data, we calculated spike-triggered averages (STAs) and covariance of stimuli (STC). The STA is simply the average of the set of stimuli that led to spikes subtracted from the mean of the prior stimulus distribution (i.e., the distribution of all stimuli independent of spiking output). Here, to remove the ambiguities caused by the temporal correlations, we used the fluctuating part of the unfiltered stimulus, $I(t) = \sigma(\sqrt{\zeta}(t) + \sqrt{1 - \zeta}(t))$; in other words, we used $\xi(t)$ instead of $\zeta(t)$ (see Eq. 2). Therefore, we have the following:

$$\text{STA}(t) = \langle I(t_{\text{spike}} - t) \rangle_{\text{spike}} - \langle I \rangle_{\text{prior}}, \quad (8)$$

The time window for the STA was 200 ms before each spike, which captured most of the STA power. In a similar way, the STC and spike-triggered correlation of stimuli (STCor) $Q(t, t')$ are given by Bialek and de Ruyter Van Steveninck (1988) as follows:

$$\begin{aligned} \text{STC}(t, t') &= \text{Cov}_{\text{spike}} - \text{Cov}_{\text{prior}} \\ &= \langle \{I(t_{\text{spike}} - t) - \text{STA}(t)\} \\ &\quad \langle I(t_{\text{spike}} - t') - \text{STA}(t') \rangle_{\text{spike}} - \text{Cov}_{\text{prior}} \rangle \\ &= \langle I(t_{\text{spike}} - t)I(t_{\text{spike}} - t') \rangle_{\text{spike}} - \text{STA}(t)\text{STA}(t') \\ &\quad - \langle I(-t)I(-t') \rangle_{\text{prior}} \\ Q(t, t') &= \text{STA}(t)\text{STA}(t') + \text{STC}(t, t'), \end{aligned} \quad (9)$$

where Cov_{spike} and Cov_{prior} are the covariance matrices of the spike-triggered and prior stimuli, respectively. The STA and STCor were used for predicting a cross-correlogram (CCG) in the first- and second-order by Equations 29 and 31, respectively (see below, Predicting CCGs from reverse correlation analysis). Predicted spike train covariance and correlation were computed from the CCG in the same way as the measured CCGs (see below).

Calculation of the measured CCG and correlation

We computed the CCGs of each spike train as follows. We started by building a spike train from the spike times with a $\Delta t = 1$ ms time bin and computed the CCGs via the time-averaged unbiased empirical correlation function (Perkel et al., 1967): when a spike train for neuron i and k th repetition is $y_{i,k}(t)$, the cross-correlogram is given by the following:

$$\begin{aligned} \text{CC}(\tau) &= \text{CCF}(\tau) - \text{CCF}_{\text{shuffle}}(\tau), \\ \text{CCF}(\tau) &= \langle y_1(0)y_2(\tau) \rangle \\ &= \frac{1}{N_{\text{repeat}}\Delta t(L - |\tau|)} \sum_{k=1}^{N_{\text{repeat}}} \int_0^{L-\tau} y_{1,k}(s)y_{2,k}(s + \tau)ds, \\ \text{CCF}_{\text{shuffle}}(\tau) &= \langle y_1(0)y_2(\tau) \rangle_{\text{shuffle}} \\ &= \frac{1}{N_{\text{repeat}}\Delta t(L - |\tau|)} \sum_{k=1}^{N_{\text{repeat}}} \int_0^{L-\tau} y_{1,k}(s)y_{2,k+1}(s + \tau)ds, \end{aligned} \quad (10)$$

where N_{repeat} is the number of repetitions, and L is the length of the spike trains and $y_{i,N_{\text{repeat}}+1} = y_{i,1}$.

From Equation 10, we computed the correlation of the spike-counts with the time window of size T as follows:

$$n_{i,k}(t) = \int_0^T y_{i,k}(t + \tau)d\tau = \int_{-\infty}^{\infty} y_{i,k}(t + \tau)\Theta_T(\tau)d\tau. \quad (11)$$

Θ_T is a window function giving $\Theta_T(t) = 1$ if $0 \leq t \leq T$ and $\Theta_T(t) = 0$ otherwise. The shift correlator computes the covariance as follows:

$$C_T = \langle n_1(t)n_2(t) \rangle - \langle n_1(t)n_2(t) \rangle_{\text{shuffle}}, \quad (12)$$

and the correlation coefficient is given by $\rho_T = C_T/(C_1C_2)^{1/2}$ where the auto-covariance of each neuron C_i ($i = 1, 2$) was calculated in the same way. Equations 10 and 12 are related as follows:

$$\begin{aligned} \langle n_1(t)n_2(t) \rangle &= \int_{-\infty}^{\infty} \text{CCF}(\tau)\Phi_T(\tau)d\tau \\ \langle n_1(t)n_2(t) \rangle_{\text{shuffle}} &= \int_{-\infty}^{\infty} \text{CCF}(\tau)_{\text{shuffle}}\Phi_T(\tau)d\tau, \end{aligned} \quad (13)$$

where $\Phi_T(t)$ is a triangular window function such that $\Phi_T(t) = |T - t|$ if $-T \leq t \leq T$ and $\Phi_T(t) = 0$ otherwise (Bair et al., 2001). Therefore, C_T as well as ρ_T were computed from CCGs via Equation 13.

In our analysis, we always computed the full CCGs and tried to analyze their behavior both at long and short timescales. The time window size T was not as important as in de la Rocha et al. (2007); therefore, in every case, we used $T = 200$ ms and dropped T from the notations.

Statistical analysis

We calculated the variance of the covariance, $\text{Var}[C]$, by computing the bootstrap statistics of $\langle y_1(0)y_2(\tau) \rangle_{\text{shuffle}}$ in Equation 10; at each step, we constructed a resampled cross-correlogram $\langle y_1(0)y_2(\tau) \rangle_{\text{resample}}$ by random resampling from $\langle y_1(0)y_2(\tau) \rangle_{\text{shuffle}}$ and computed $\langle n_1n_2 \rangle_{\text{resample}}$ by Equation 12. We collected 400 $\langle n_1n_2 \rangle_{\text{resample}}$ with which the variance

became stable. The $\text{Var}[C]$ is taken as this resample variance and C is also recalibrated with the resample mean as follows:

$$\begin{aligned} \text{Var}[C] &= \text{Var}[\langle n_1 n_2 \rangle_{\text{resample}}], \\ C \rightarrow C &- E[\langle n_1 n_2 \rangle_{\text{resample}}]. \end{aligned} \quad (14)$$

For experimental data in Figure 7B–D, t scores were calculated from C and $\text{Var}[C]$. However, as for the simulation data, we could arbitrarily add more repetitions to suppress $\text{Var}[C]$ down to sufficiently low level. For the analytic prediction, $\text{Var}[C]$ could be computed in a similar way, but it was always insignificant (<1% of the covariance) for both experimental and simulation data.

In Figures 5 and 7, we evaluated the predictive power of the predicted C in terms of the coefficient of determination, R^2 , defined by the following:

$$R^2 = 1 - \frac{\sum [\text{Measured } C/c] - (\text{Predicted } C/c)]^2}{\sum [(\text{Measured } C/c) - E[(\text{Measured } C/c)]]^2}. \quad (15)$$

By definition, $R^2 = 1$ signifies the perfect match, while the prediction has no correlation with the measured C when $R^2 = 0$. In some cases, $R^2 < 0$, and this signifies that the predictions in fact diverge away from the data with a different average.

Predicting CCGs from reverse correlation analysis

Here, we derive the first- and second-order prediction presented as Equations 29–32 in Results by using the reverse correlation analysis of the linear–nonlinear (LN) cascade model (Victor and Shapley, 1980; Meister and Berry, 1999). The LN model is composed of two stages: first, the stimulus $I(t)$ was linear filtered by the relevant features of the model $\{\varepsilon_\alpha\}$ ($\alpha = 1, 2, \dots, D$), and the probability to spike at the given time bin $[t, t + \Delta t]$, $P(\text{spike}|x)$ where

$$x_\alpha = \int_0^\infty \varepsilon_\alpha(t') I(t - t') dt'. \quad (16)$$

Here, as above (see Reverse correlation analysis), we use the unfiltered stimulus $I(t) = \sigma(\sqrt{c}\xi_c(t) + \sqrt{1-c}\xi_i(t))$. Since the actual injected stimulus and $I(t)$ are linearly related, STA and STCor have the same temporal correlation as the actual stimulus, and, consequently, the temporal correlation naturally shows up in the predicted CCGs; for example, compare predicted CCGs for $\tau = 5$ and 50 ms in Figure 5B.

We follow the same derivation as in the study by de la Rocha et al. (2007): the common noise part $\delta I = \sqrt{c}\sigma\xi_c$ will be regarded as the perturbation on top of $I_0 = I(t)|_{c=0}$ and the comodulated part of the firing rate will be determined by Wiener series expansion in δI .

First-order prediction (see Results, Eqs. 29, 30). The firing rate change induced by a small perturbation $I_0 \rightarrow I_0 + \delta I$ can be approximated as follows:

$$P(\text{spike}) \approx P(\text{spike})_{\delta I=0} \left(1 + \frac{1}{\sigma^2} \int_0^\infty \text{STA}(t') \delta I(t - t') dt' \right), \quad (17)$$

where $\delta I(t)$ has Gaussian statistics with a variance σ^2 (Rieke et al., 1999; Hong et al., 2008). STA represents the spike-triggered average of stimuli. Then the correlation function of $q(t) = P(\text{spike at } t)$ is as follows:

$$\begin{aligned} &\langle (q_1(0) - \langle q_1 \rangle)(q_2(\tau) - \langle q_2 \rangle) \rangle \\ &= \frac{\langle q_1 \rangle \langle q_2 \rangle}{\sigma_1^2 \sigma_2^2} \int_0^\infty \text{STA}_1(t') \text{STA}_2(t') \langle \delta I_1(-t') \delta I_2(\tau - t') \rangle dt' dt'' \\ &= \frac{c \langle q_1 \rangle \langle q_2 \rangle}{\sigma_1^2 \sigma_2^2} \int_0^\infty \text{STA}_1(t') \text{STA}_2(t' + \tau) dt', \end{aligned} \quad (18)$$

where $\langle \delta I_1(0) \delta I_2(\tau) \rangle = c \sigma_1 \sigma_2 \delta(\tau)$ if $\delta I_i = \sqrt{c} \delta \xi_c$. This is essentially the same derivation of the correlation–gain relationship based on the linear response theory (de la Rocha et al., 2007; Hong et al., 2008) since the STA is the linear kernel relating the stimulus and firing rate as in Equation 17.

We now define the firing rate $v(t) = q(t)/\Delta t$ and the predicted cross-correlogram for each pair as follows:

$$\text{CCG}_{\text{first order}}(t) = c \frac{\nu_1 \nu_2}{\sigma_1 \sigma_2} \int \text{STA}_1(t') \text{STA}_2(t + t') dt', \quad (19)$$

which is Equation 29 in Results. Furthermore, if we use the identity (Chialvo et al., 1997; Hong et al., 2008),

$$\sigma \frac{\partial v}{\partial \mu} = \frac{v}{\sigma} \int \text{STA}(t) dt, \quad (20)$$

the firing rate covariance is given by the following:

$$\begin{aligned} C_{\text{first order}} &= \int d\tau \langle (v_1(0) - \langle v_1 \rangle)(v_2(\tau) - \langle v_2 \rangle) \rangle \\ &= c \frac{\nu_1 \nu_2}{\sigma_1 \sigma_2} \int \int \text{STA}_1(t) \text{STA}_2(t + t') dt dt' \\ &= c \frac{\nu_1 \nu_2}{\sigma_1 \sigma_2} \left(\int \text{STA}_1(t) dt \right) \left(\int \text{STA}_2(t'') dt'' \right) \\ &= c \sigma_1 \sigma_2 \frac{\partial v_1}{\partial \mu_1} \frac{\partial v_2}{\partial \mu_2}. \end{aligned} \quad (21)$$

where we changed the variables in the third line as $t'' = t + t'$. The last line is equivalent to the original relationship between the correlation and gain (see Results, Eq. 30) (de la Rocha et al., 2007; Shea-Brown et al., 2008).

Second-order prediction (see Results, Eqs. 31, 32). When we include the second-order term in Equation 17, we obtain the following (Rieke et al., 1999; Hong et al., 2008):

$$\begin{aligned} P(\text{spike}) &\approx P(\text{spike})_{\delta I=0} \left(1 + \frac{1}{\sigma^2} \int_0^\infty \text{STA}(t') \delta I(t - t') dt' \right. \\ &\quad \left. + \frac{1}{2! \sigma^4} \int_0^\infty \int_0^\infty Q(t_1, t_2) \delta I(t - t_1) \delta I(t - t_2) dt_1 dt_2 \right), \end{aligned} \quad (22)$$

where $Q(t_1, t_2)$ is the STCor in Equation 9. The contribution to the CCG, Equation 31 in Results, is obtained in a straightforward way, as follows:

$$\delta \text{CCG}_{\text{second order}}(t) = c^2 \frac{\nu_1 \nu_2}{2 \sigma_1^2 \sigma_2^2} \int \int Q_1(t_1, t_2) Q_2(t_1 + t, t_2 + t) dt_1 dt_2. \quad (23)$$

Note that we excluded the self-contraction to generate a proper Wiener series (Rieke et al., 1999).

In particular, when we have the same input condition, $\mu_1 = \mu_2 = \mu$ and $\sigma_1 = \sigma_2 = \sigma$, for the same type of neuron as in Figures 3 and 6, the peak height at $t = 0$ is given by the following:

$$\begin{aligned} CCG_{\text{first+second order}}(0) &= c \frac{\nu^2}{\sigma^2} \int \text{STA}(t)^2 dt \\ &\quad + c^2 \frac{\nu^2}{2\sigma^4} \iint Q(t_1, t_2)^2 dt_1 dt_2. \end{aligned} \quad (24)$$

Furthermore, we consider the case when the neuron is effectively well described by a single preferred spike-evoking stimulus feature (PSESF), $\varepsilon(t)$, such as when the neuron functions almost as a pure integrator or a pure coincidence detector. In this case, $Q(t_1, t_2)$ can be written as $Q(t_1, t_2) = Y(t_1)Y(t_2)$, where $Y(t) \propto \varepsilon(t)$, and therefore

$$\begin{aligned} \delta CCG_{\text{second order}}(0) &= c^2 \frac{\nu^2}{2\sigma^4} \iint Y(t_1)^2 Y(t_2)^2 dt_1 dt_2 \\ &= c^2 \frac{\nu^2}{2\sigma^4} \left(\int Q(t_1, t_1) dt_1 \right) \left(\int Q(t_2, t_2) dt_2 \right). \end{aligned} \quad (25)$$

From $\int Q(t_1, t_1) dt_1 = \sigma^2(\partial\nu/\partial\sigma)/\nu$ (Hong et al., 2008), the predicted peak height becomes

$$CCG_{\text{first+second order}}(0) = CCG_{\text{first order}}(0) + \frac{1}{2} c^2 \sigma^2 \left(\frac{\partial\nu}{\partial\sigma} \right)^2, \quad (26)$$

which is equivalent to Equation 32 in Results. Therefore, the firing rate gain with respect to stimulus variance contributes to the CCG peak height.

Note in the case of a single PSESF that the input mean and variance sensitivity of the firing rate are related as follows (Hong et al., 2008):

$$\frac{\partial\nu}{\partial\sigma} = \frac{\sigma}{\varepsilon^2} \cdot \frac{\partial^2\nu}{\partial\mu^2}, \quad (27)$$

where

$$\bar{\varepsilon} = \int \varepsilon(t) dt \quad (28)$$

Therefore, when $\bar{\varepsilon} \neq 0$, as in the pure integrator, the contribution of the peak to the correlation is suppressed compared with the first-order prediction, Equation 30. In the pure coincidence detector, on the other hand, $\bar{\varepsilon} = 0$, but firing rate ν is also independent of μ (see Fig. 6B), which can still make $\partial\nu/\partial\sigma$ finite. Therefore, the predicted correlation in the pure coincidence detector [based on Eqs. 29, 30, and $\nu(\mu) = \text{constant}$] (Barreiro et al., 2010) will be profoundly modified by the quadratic-order approximation (see Results, Simulations in a phenomenological coincidence detector model).

Results

When two or more neurons receive correlated (i.e., shared) input, their output spike trains should exhibit some correlation despite the effects of independent input (Fig. 1A). Spike train covariance C and the correlation coefficient ρ (i.e., C normalized by spike train variance) should, therefore, carry information about the input correlation c . However, de la Rocha et al. (2007) showed that the relationship between input and output correlation (denoted correlation susceptibility $S = \rho/c$) depends on the mean μ and variance σ^2 of the input (Fig. 1B). This is important because stimulus-dependent changes in S prevent one-to-one mapping between input correlation and output correlation. Correlation-based coding is straightforward if S is independent of μ and σ^2 (Fig. 1C, left), but it is compromised or requires a

more complicated decoding scheme if S varies with μ and σ^2 (Fig. 1C, right).

One way of understanding why this occurs is that, for a given input correlation, output correlation will vary depending on the sensitivity (i.e., gain) of the firing rate ν of each neuron with respect to the stimulus mean μ : if stimulus fluctuations occur within a steep region of the $\nu-\mu$ curve, rate will fluctuate widely in each cell and the pair will exhibit large comodulated rate fluctuations, whereas stimulus fluctuations within a shallow region of the $\nu-\mu$ curve will logically drive smaller comodulated rate fluctuations (Fig. 1D). Consequently, ρ “inherits” the same tuning as $\nu(\mu)$, even when c is fixed; under these conditions, c cannot be unambiguously decoded from ρ without knowledge of μ . This line of reasoning triggered three concerns: (1) it applies to rate comodulation but not necessarily to spike-time synchronization, and thus it neglects one component of output correlation; (2) it implicitly assumes that input fluctuations are “noise” rather than “signal”; and (3) output rate and synchronization (i.e., correlation based on spike-time synchronization as opposed to rate comodulation) are liable to be tuned to different stimulus properties. This led us to our overall hypothesis that some cell types may encode signal-dependent fluctuations with precise spike-time synchronization and can do so independently of rate-based coding of other stimulus features.

The dependence of output rate ν on input parameters μ and σ differs fundamentally between cell types, as shown in Figure 2 for our conductance-based models: ν is principally sensitive to μ in the case of integrators, whereas it is also very sensitive to σ in the case of coincidence detectors (Higgs et al., 2006; Arsiero et al., 2007; Lundstrom et al., 2008). For simulations reported here, input was treated as a continuous stream rather than as discrete synaptic inputs (Destexhe et al., 2001); nonetheless, σ reflects coordinated fluctuations in presynaptic activity (Fellous et al., 2003), the temporal structure of which is reflected in the autocorrelation time τ (shorter τ implies more precise synchrony), while c specifies the proportion of inputs shared between two postsynaptic neurons. The first two parameters, σ and τ , affect the temporal precision of spiking in each postsynaptic neuron, while c affects correlation across neurons—all three parameters ultimately affect output synchrony.

The greater sensitivity of coincidence detectors to input synchrony relative to integrators (Fig. 2A, B) is a direct consequence of active membrane properties: activation of outward current (or inactivation of inward current) at perithreshold potentials helps ensure spike generation selectively in response to fast stimulus fluctuations (i.e., synchronous inputs), whereas perithreshold-activating inward current (or inactivating outward current) encourages repetitive spiking in response to constant or slow-changing input (Fourcaud-Trocmé et al., 2003; Svirskis et al., 2004; Higgs et al., 2006; Arsiero et al., 2007; Lundstrom et al., 2008; Prescott et al., 2008a). Differential sensitivity to input synchrony can be demonstrated most succinctly by contrasting which stimulus features preferentially elicit spikes in each cell type. We estimated the preferred spike-elicting stimulus feature as the spike-triggered-averaged stimulus (STA) of the response of each cell to noisy input (see examples in Fig. 2B). The integrator exhibits a relatively broad, monophasic STA (Fig. 2C, left), whereas the coincidence detector exhibits a biphasic STA with a positive phase that is remarkably narrow (Fig. 2C, right).

The shape of the STA should, in theory, relate directly to the cross-correlation of output spiking given that the CCG is the overlap integral of the STA of each neuron, according to the first-order ap-

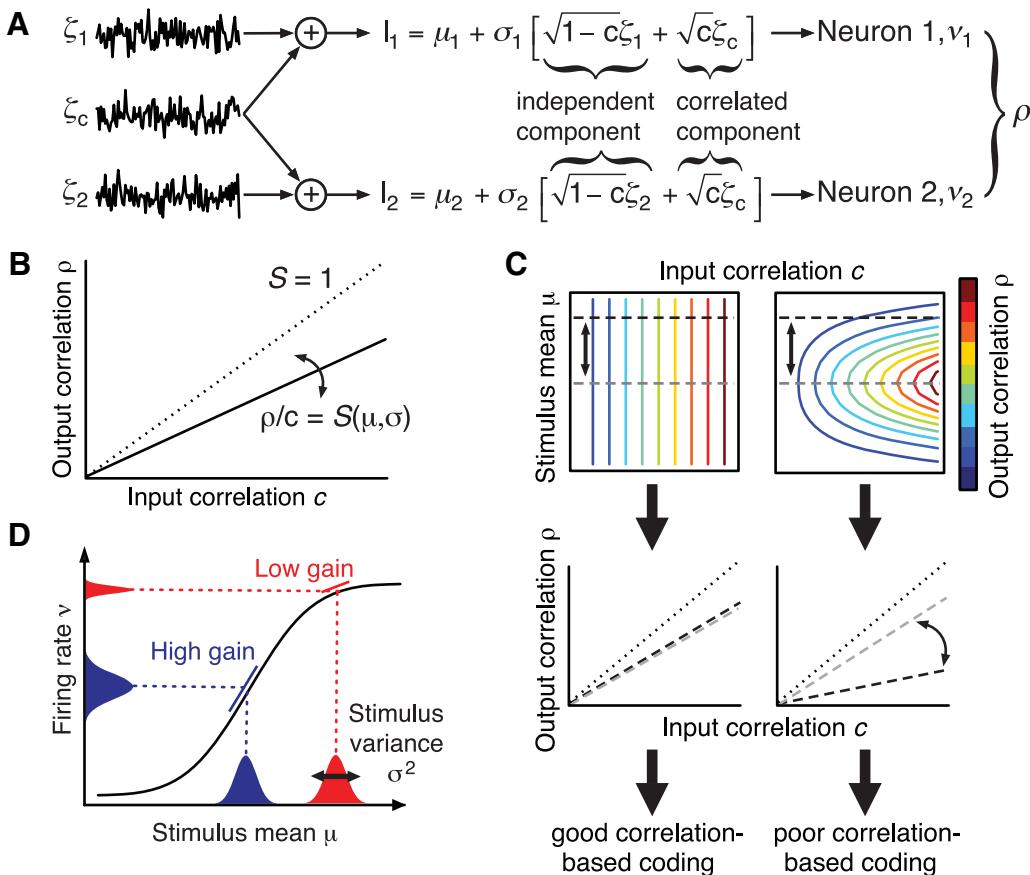


Figure 1. Relationship between input and output correlation. **A**, Stimulation paradigm in which neurons 1 and 2 receive fluctuating input I_1, I_2 , with mean μ_1, μ_2 , and variance σ_1^2, σ_2^2 . Fluctuating input was modeled as an Ornstein–Uhlenbeck process with $\tau = 5$ ms. Some fraction of that input is shared, or correlated, as defined by the input correlation c . Output firing rate ν_1, ν_2 , and the output correlation coefficient ρ (=spike train covariance C normalized by variance) were measured. **B**, Plotting output correlation ρ against input correlation c shows how much correlation is transferred by the pair of neurons. The slope of that curve, denoted correlation susceptibility S , is ≤ 1 but has been shown to depend on input parameters μ and σ (de la Rocha et al., 2007). **C**, Input correlation c can only be unambiguously decoded from ρ (without knowledge of other input parameters) if S does not vary with other input parameters. The dashed curves on the bottom plots show horizontal cross-sections through 3-D plots (top) at different μ . An invariant ρ - c relationship (left) is conducive to good correlation-based coding, whereas a variable relationship (right) is not unless a more complicated decoding scheme is invoked. **D**, If ν is tuned to μ , then fluctuations around μ will produce fluctuations in ν whose magnitude depends on $\partial\nu/\partial\mu$. If neurons 1 and 2 receive input with correlated fluctuations, ν_1 and ν_2 will be comodulated. Amplitude of ν comodulation naturally depends on $\partial\nu/\partial\mu$, rendering ρ and ν cotuned to μ . In that scenario, rate comodulation will not provide information about μ beyond that already provided by rates ν_1 and ν_2 , but this does not rule out spike-time synchronization providing information about σ if input fluctuations are considered signal rather than noise.

proximation in c (see Materials and Methods) (de la Rocha et al., 2007), as follows:

$$\text{CCG}_{\text{first order}}(t) \approx c \frac{\nu_1 \nu_2}{\sigma_1 \sigma_2} \int \text{STA}_1(\tau) \cdot \text{STA}_2(t + \tau) d\tau. \quad (29)$$

Hence, the broad monophasic integrator STA predicts a broad monophasic CCG for pairs of integrators (Fig. 2D, left), whereas the narrow biphasic coincidence detector STA predicts a narrow biphasic CCG for pairs of coincidence detectors (Fig. 2D, right). Using the relationship between the STA and firing rate, the covariance C of the output spike count is as follows (see Materials and Methods):

$$C = \int \text{CCG}(t) dt \approx c \sigma_1 \sigma_2 \frac{\partial \nu_1}{\partial \mu_1} \frac{\partial \nu_2}{\partial \mu_2}, \quad (30)$$

which is consistent with the study by de la Rocha et al. (2007) because sensitivity (or gain) of ν with respect to μ is the dominant factor determining the degree of correlated spiking given a cer-

tain degree of input correlation (Fig. 1D). Strong dependence of the relationship between ρ and c (i.e., correlation susceptibility S) on μ is not conducive to correlation-based coding (Fig. 1C, right). However, the STA does not always accurately represent the preferred spike-eliciting stimulus feature (this occurs when spike generation is sensitive to higher-order stimulus statistics), which invalidates predictions based on Equation 29 in certain cases (see below).

We therefore set out to identify (1) whether and how correlation susceptibility $S(\mu, \sigma)$ differs between integrators and coincidence detectors, (2) what the consequences of such differences are for correlation-based coding, and (3) precisely why the stimulus dependence of S differs between cell types.

Simulations in conductance-based integrator and coincidence detector models

To compare the correlation susceptibility S ($=\rho/c$) of integrators and coincidence detectors, we conducted a series of numerical simulations using pairs of model neurons receiving a mix of correlated and independent input (Fig. 1A). For each neuron type, we varied mean μ under high- or low-variance σ^2 conditions, and evaluated the output rate ν and correlation ρ (Fig. 3). As

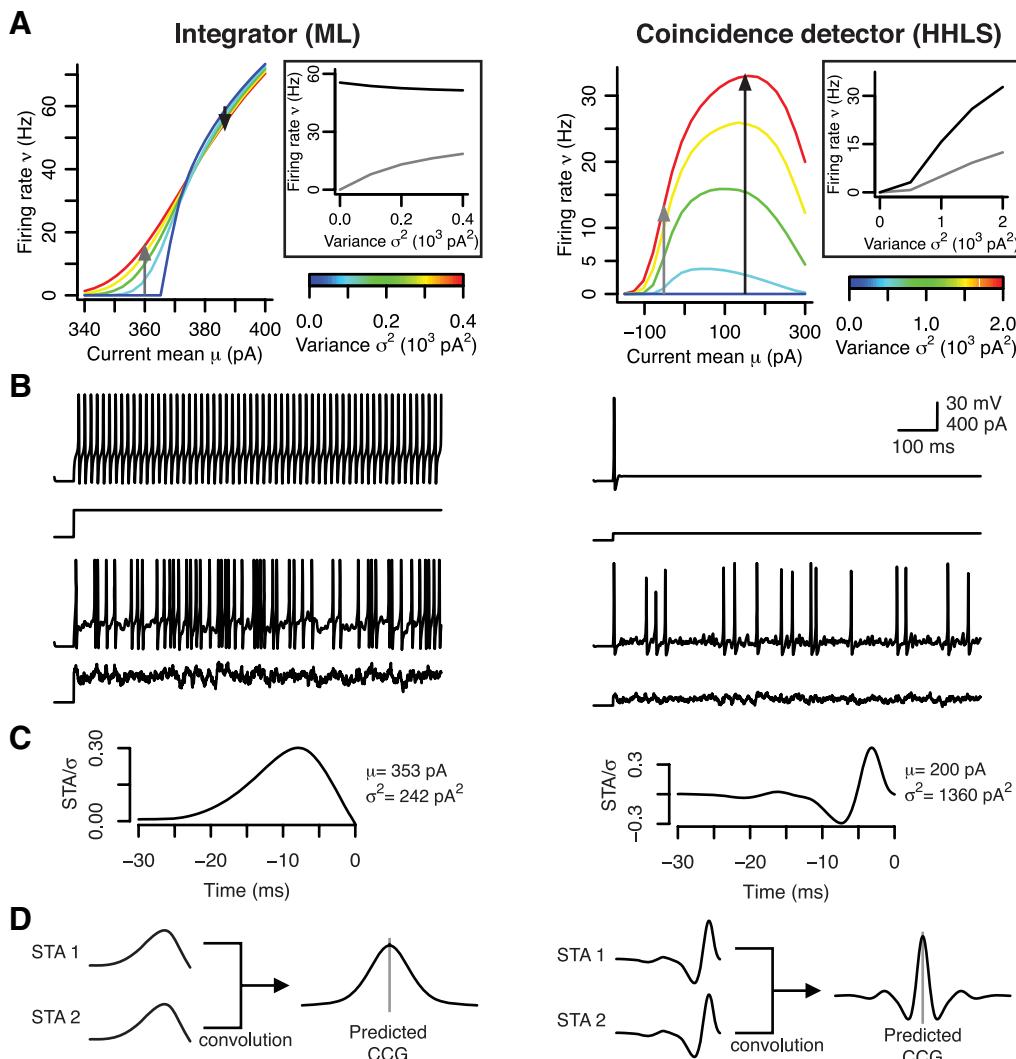


Figure 2. Integrators and coincidence detectors are sensitive to different stimulus statistics. Data here are based on simulations in conductance-based models (see Materials and Methods). **A**, In the integrator (left), ν was sensitive to μ but was relatively insensitive to σ , whereas ν was sensitive to both μ and σ in the coincidence detector (right). The insets highlight the differential ability of each cell type to encode σ ; black and gray curves correspond to arrows on ν - μ plots. Notably, firing rate variation with σ may reflect the rate of brief, suprathreshold input events rather than the (rate-encoded) magnitude of those events. **B**, Sample traces show differential responsiveness of each cell type to constant and fluctuating input. The coincidence detector responds preferentially to fast stimulus transients because its voltage-dependent currents implement a high-pass filter. The integrator also responds to fast stimulus fluctuations, but its voltage-dependent currents encourage repetitive spiking even when input is constant. **C**, The differential requirements for spike generation are evident from the spike-triggered-averaged response (STA) to fluctuating input. **D**, The CCG corresponds to the overlap integral of the STA from each neuron within a pair. Therefore, pairs of integrators are predicted to exhibit broad, monophasic CCGs (left), whereas pairs of coincidence detectors are predicted to exhibit narrow, biphasic CCGs based on the typical shape of their STA (right).

illustrated in Figure 2A, maximal ν differs between cell types and between σ conditions, which leads to unavoidable differences in the range of ν on each panel of Figure 3. Nevertheless, the $\nu(\mu)$ curves (gray) are similarly shaped in all four conditions. If ρ simply inherited the same tuning as ν with respect to μ , then $\rho(\mu)$ curves should also be similarly shaped—clearly, they are not.

As expected for integrators, ρ/c increased to a maximum <1 as μ was increased in both the high and low σ cases (Fig. 3A, top). In coincidence detectors, however, a similar pattern was observed in the low σ case, but not in the high σ case, in which ρ/c decreased after reaching a peak (Fig. 3A, bottom). Figure 3B shows the same data replotted with ν on the x -axis. For coincidence detectors receiving high σ input, the correlation–rate relationship was strongly negative at high rates. The same coincidence detectors receiving low σ input exhibited little change in correlation across most of the (albeit narrow) range of firing rates, which implies that S is not significantly dependent on μ or ν in this cell type

under these conditions. This property should be conducive to good correlation-based coding (see below).

To investigate how the stimulus dependence of S affects correlation-based coding by integrators and coincidence detectors, we measured ρ in response to different combinations of c and μ (Fig. 4A). As predicted, S was strongly dependent on μ in the case of integrators, which caused encoding of c by ρ to be ambiguous; in contrast, S was relatively unaffected by μ in the case of coincidence detectors, consistent with good correlation-based coding (compare Fig. 1C). Notably, Equation 29 failed to accurately predict ρ in the case of coincidence detectors (Fig. 4A, inset) despite having worked in the case of integrators (see below).

Small values of ρ observed among coincidence detectors stem from the shape of coincidence detector CCGs, which are narrow and biphasic, unlike integrator CCGs, which are broad and monophasic (Fig. 4B; compare prediction in Fig. 2D). Small val-

ues of ρ do not imply that synchrony transfer will fail, because downstream coincidence detectors prefer tall, narrow CCGs among their upstream neurons rather than short, broad CCGs, regardless of the absolute magnitude of ρ (S. Hong, S. A. Prescott, and E. De Schutter, unpublished observations). To summarize, correlation-based coding is viable despite small ρ as long as the relationship between ρ and c is insensitive to other stimulus features like μ . These data demonstrate the feasibility of correlation-based coding among coincidence detectors.

Next, we asked why $S(\mu, \sigma)$ differs between integrators and coincidence detectors. In all panels of Figure 3, ρ/c values predicted from Equation 29 were plotted for comparison with measured ρ/c values. The derivation of Equation 29 is based on linear response theory and requires several assumptions that are not strictly met in our simulations such as very small c . However, despite comparable stimulus parameters and identical methods used to calculate output correlation, Equation 29 predicted most of the output correlation observed for integrators but only a fraction of that observed for coincidence detectors. In the latter case, the degree of inaccuracy appeared to depend on σ . We reasoned that identifying why Equation 29 fails to accurately predict output correlation in coincidence detectors, especially under certain stimulus conditions, would help identify how correlated spiking in coincidence detectors differs from that in integrators.

To test the accuracy of the prediction, we plotted predicted covariance against measured covariance in Figure 5. If prediction by Equation 29 were consistently accurate (i.e., for all stimulus conditions), all data points would lie along the diagonal line. Figure 5A illustrates the consistent accuracy of the prediction for pairs of integrators versus its inconsistency for pairs of coincidence detectors. These data represent responses to a broad range of different μ and σ . The important observation is that the first-order prediction was reasonably accurate for all stimulus conditions in the case of integrators, whereas it was grossly inaccurate for many stimulus conditions in the case of coincidence detectors. Whereas Figure 3 illustrates stimulus conditions for which prediction by Equation 29 is good or bad, Figure 5 focuses on why Equation 29 sometimes fails to predict output correlation among coincidence detectors. By plotting covariance C rather than the correlation coefficient ρ , Figure 5A confirms that the prediction error seen for coincidence detectors is not attributable to autovariances in the spike train. Sample CCGs (Fig. 5A, right) reveal that the predicted CCG differs most from the measured CCG at the central peak; this is true for both integrators and coincidence detectors,

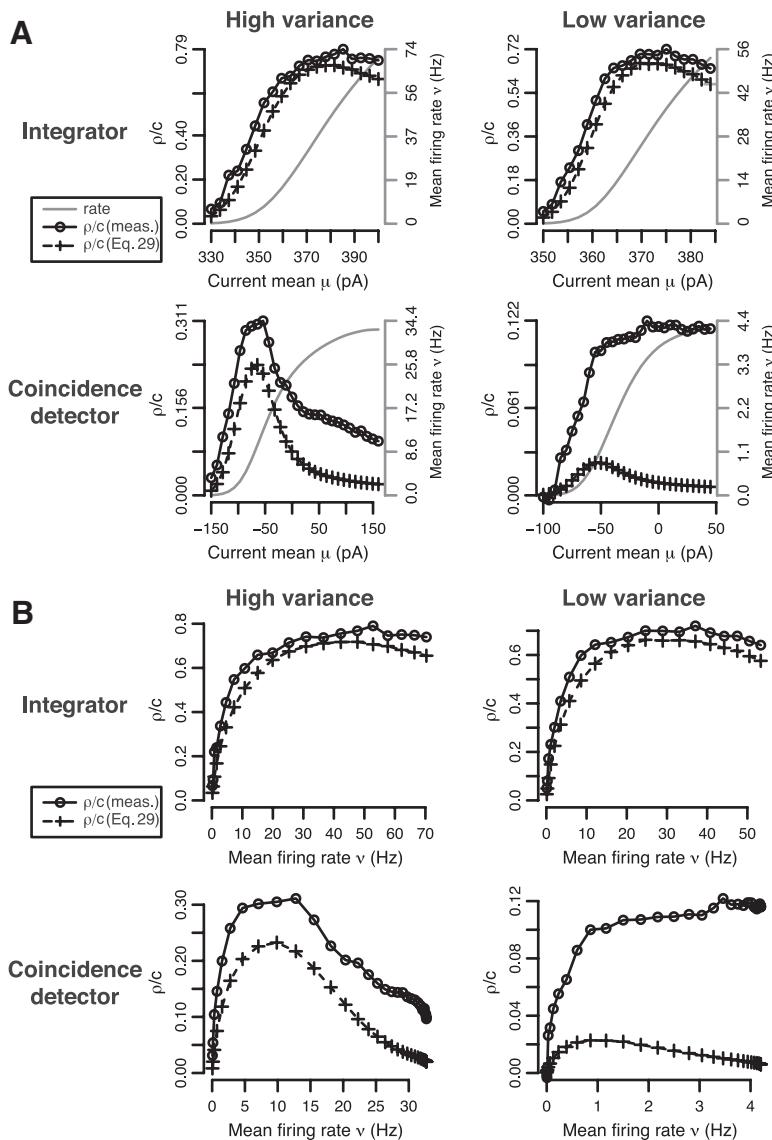


Figure 3. Output correlation is sensitive to stimulus variance in coincidence detectors, unlike in integrators, and has a variable relationship with output rate. **A**, Using the same conductance-based models as in Figure 2, the integrator (top) and coincidence detector (bottom) were stimulated with high- and low-variance input: $\sigma_{\text{high}}^2 = 400 \text{ pA}^2$ and $\sigma_{\text{low}}^2 = 100 \text{ pA}^2$ for the integrator; $\sigma_{\text{high}}^2 = 2000 \text{ pA}^2$ and $\sigma_{\text{low}}^2 = 520 \text{ pA}^2$ for the coincidence detector. Output rate ν and correlation ρ measured from simulations (\circ) together with ρ predicted from Equation 29 (+) were plotted against input mean μ . In **B**, ρ was replotted against ν to visualize the correlation–rate relationship. The range of ν (gray) varies between panels because of differences in maximal firing rate across cell types and stimulus conditions (Fig. 2A). For the integrator, ρ increased with ν and was quite accurately predicted by Equation 29 in both the high- and low-variance conditions. For the coincidence detector, ρ decreased for $\nu > 10 \text{ Hz}$ in the high-variance condition, which was only qualitatively predicted from Equation 29, and the prediction was even less accurate in the low-variance condition. These results demonstrate that output correlation in coincidence detectors is higher than predicted on the basis of rate comodulation, and is accentuated under stimulus conditions in which $\partial \nu / \partial \mu$ is low. Given that rate comodulation should be minimized under those conditions (Fig. 1D), correlation in excess of the prediction might be attributable to some mechanism other than rate comodulation. In every case, input correlation was $c = 0.3$.

but it translates into a larger discrepancy in predicting overall output correlation as the CCG gets narrower. Thus, these data qualitatively confirmed our starting predictions and identified a quantitative shortfall in the ability of Equation 29 to predict output correlation, especially in coincidence detectors.

In addition to reflecting the stimulus features to which each cell type is most sensitive, we reasoned that the width of the CCG also reflects the autocorrelation time of the correlated signal (which can be taken to reflect the temporal precision of spiking in presynaptic neurons) (see above). Therefore, we lengthened τ

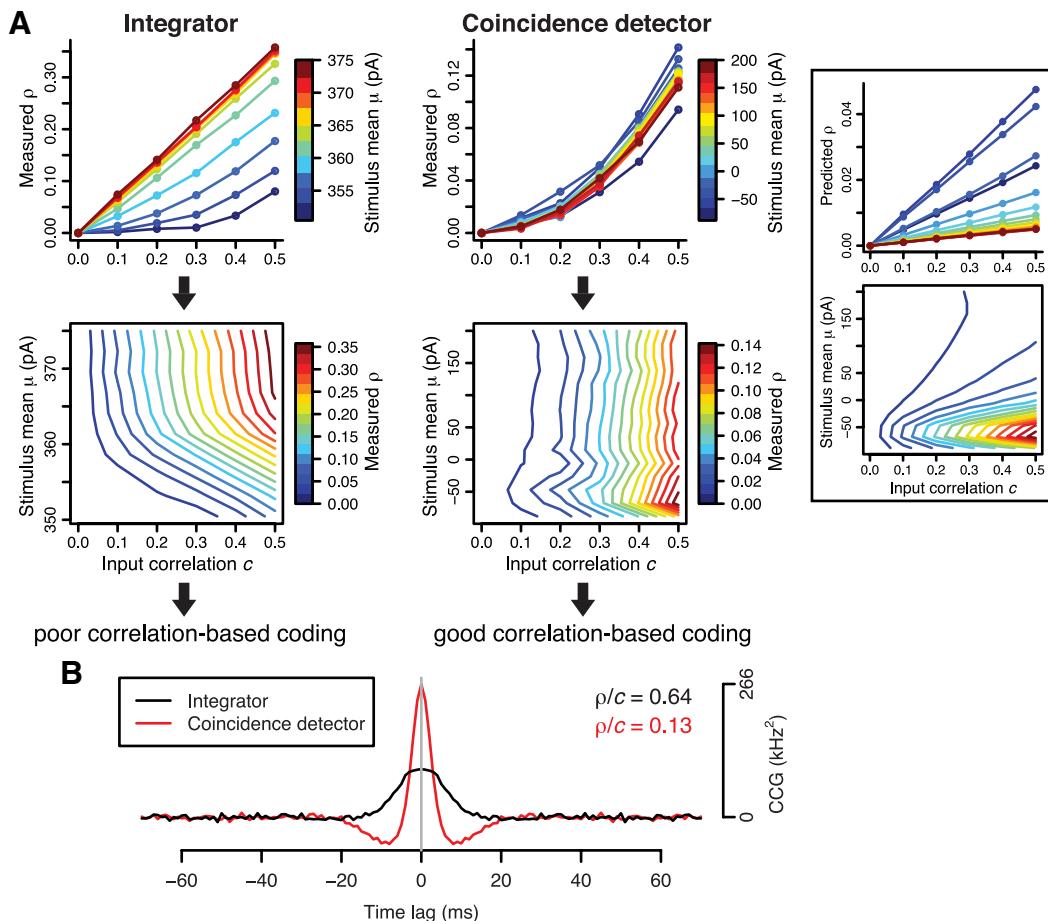


Figure 4. Correlation-based coding by integrators and coincidence detectors. **A**, We measured output correlation ρ in response to different combinations of mean μ and input correlation c (top). Curves on 2-D plots correspond to horizontal cross-sections, at different μ , through 3-D plots (bottom; compare Fig. 1C). Tight clustering of ρ – c curves despite differences in μ is conducive to good correlation-based coding by coincidence detectors (right), which contrasts to the broad distribution of those curves for integrators (left). Equivalent plots for the coincidence detector based on prediction by Equation 29 (inset) shows that the first-order prediction clearly fails to account for output correlation among coincidence detectors. Noticeably, ρ is smaller among coincidence detectors than among integrators. This stems from differences in the CCG, typical examples of which are shown in **B** based on measurement from pairs of each cell type with the same mean firing rate of 12 Hz; $c = 0.3$. These measured CCGs confirmed CCG shapes predicted in Figure 2*D*. Small values of ρ do not necessarily compromise synchrony transfer and correlation-based coding, whereas stimulus-dependent variability in S does (see text). In parts **A** and **B**, $\sigma^2 = 100 \text{ pA}^2$ for the integrator and $\sigma^2 = 1000 \text{ pA}^2$ for the coincidence detector.

from 5 to 50 ms for simulations in Figure 5*B* based on the hypothesis that widening the CCG might reduce the prediction error. As expected, lengthening τ improved the prediction for both integrators and coincidence detectors but, whereas the CCG for the integrators was dramatically widened, the CCG for coincidence detectors remained narrow and the prediction still exhibited a sizeable error near the peak of the CCG. To determine whether prediction error near the peak of the CCG was sufficient to explain the discrepancy observed on the measured versus predicted C plots, we removed data within ± 2 ms of the peaks of all coincidence detector CCGs and replotted the data for $\tau = 5$ ms—this amounts to preferentially removing precisely synchronized spikes from the calculation of total output correlation. The result was a near-perfect prediction by Equation 29 (Fig. 5*C*), meaning the prediction error near the peak of narrow CCGs was sufficient to explain the failure of the prediction in coincidence detectors for all input conditions for which the original prediction was poor.

Next, we investigated (1) whether and how the original prediction might be improved so that the predicted CCG would more accurately match the measured CCG for coincidence detectors at the central peak, and (2) whether this would improve the measured versus predicted C plots (as expected given the results

in Fig. 5*C*). We reasoned that if coincidence detectors are sensitive to stimulus variance (Fig. 2), then the prediction (Eq. 29) should take into account how stimulus variance affects spike generation. Therefore, we incorporated a second-order term into the prediction based on the following:

$$\delta\text{CCG}_{\text{second order}}(t)$$

$$= \frac{c^2 \nu_1 \nu_2}{2 \sigma_1^2 \sigma_2^2} \int \int Q_1(\tau_1, \tau_2) Q_2(\tau_1 + t, \tau_2 + t) d\tau_1 d\tau_2, \quad (31)$$

where $Q(t, t')$ is the spike-triggered correlation of the stimuli (STCor) (see Materials and Methods for derivation). Total output correlation should thus be predicted by the combination of Equations 29 and 31. For conditions in which input statistics and preferred spike-elicting stimulus feature are equal, Equation 31 predicts the additional peak height as follows:

$$\delta\text{CCG}_{\text{second order}}(0) = \frac{c^2}{2} \sigma^2 \left(\frac{\partial \nu}{\partial \sigma} \right)^2. \quad (32)$$

Using Equation 31, we find that the coincidence detector CCG was much better predicted at its central peak (Fig. 5*D*, right) and

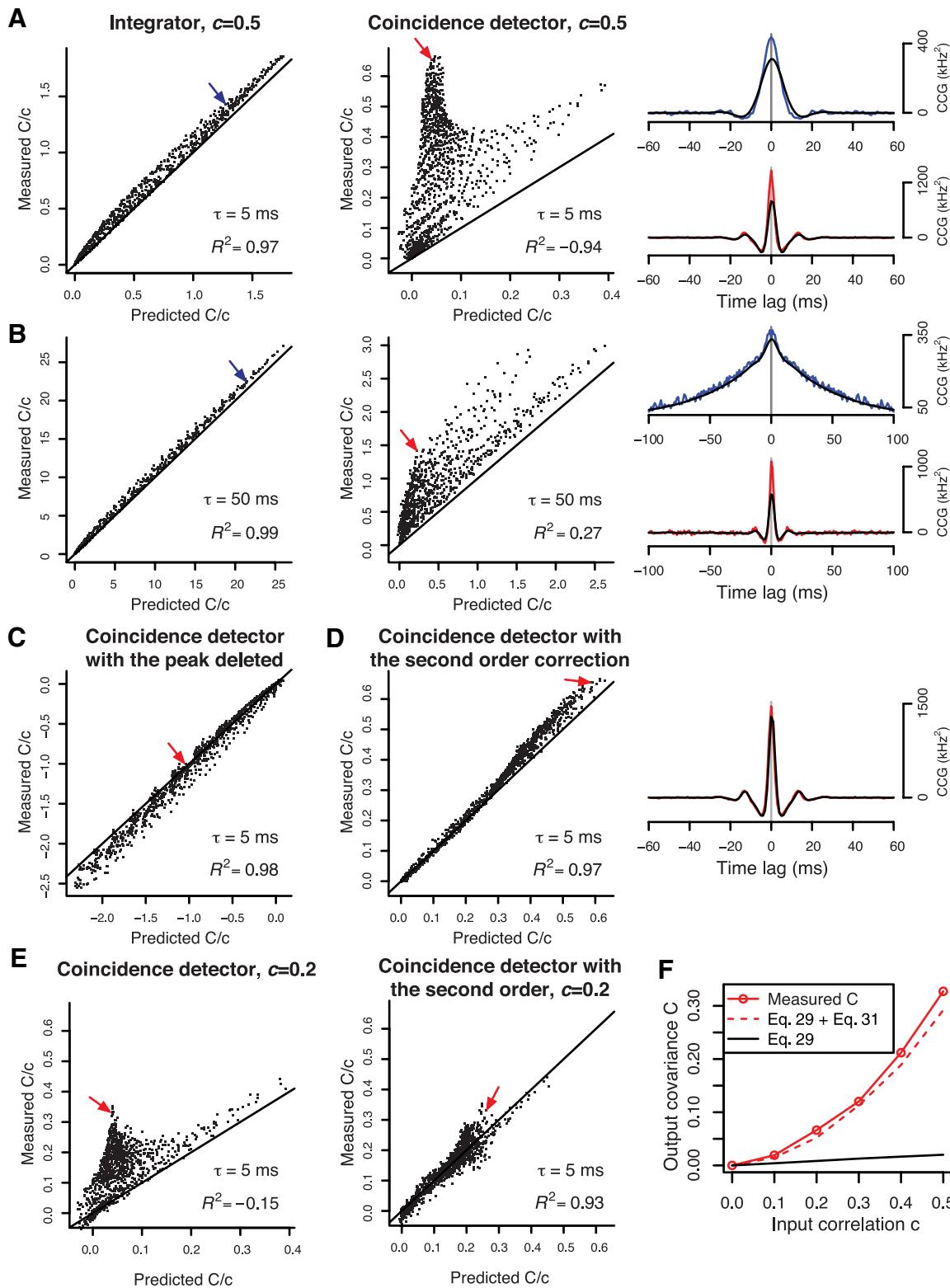


Figure 5. Sensitivity to second-order stimulus statistics explains why the first-order prediction cannot fully account for output correlation in coincidence detectors. **A**, Measured output covariance normalized by input correlation (C/c) was plotted against predicted C/c for conductance-based integrator and coincidence detectors models receiving fluctuating input ($\tau = 5 \text{ ms}$) for a broad range of μ and σ values. Prediction is based on Equation 29. The arrows point to data points for which sample CCGs are shown (right). Consistently accurate prediction (i.e., for all stimulus conditions) would be evident from points clustering around the line, which is seen for the integrator ($R^2 = 0.97$) but not for the coincidence detector ($R^2 = -0.94$). Predicted CCGs (black) deviated from measured CCGs (color) mostly around the central peak. This prediction error impacts the total output correlation for the coincidence detector more significantly than for the integrator because the CCG of the former is biphasic, which causes the predicted C to almost vanish. **B**, Same plots as in **A** but with slowly fluctuating input ($\tau = 50 \text{ ms}$). The integrator CCG broadened and resulted in improved prediction, but the coincidence detector CCG remained narrow and with still a sizable deviation from prediction. **C**, Given that the first-order prediction deviated from measured correlation primarily near the central peak of the CCG, we removed data $\pm 2 \text{ ms}$ around that peak and replotted measured versus predicted C/c based on data for $\tau = 5 \text{ ms}$ input. After removing the “excess” synchronization near the CCG peak, the first-order prediction was very accurate for coincidence detectors ($R^2 = 0.98$), meaning that Equation 29 fails specifically to explain (Figure legend continues.)

measured C was accurately predicted for a broad range of input conditions (Fig. 5D, left). Successfully modifying our quantitative prediction by incorporating Equation 31 argues that the difference in threshold mechanism is critical insofar as spike generation in integrators is preferentially sensitive to first-order stimulus statistics, whereas spike generation in coincidence detectors is also sensitive to second-order stimulus statistics (Fig. 2). On a more practical note, our ability to modify our quantitative prediction so that it works consistently (i.e., for all stimulus conditions) for integrators as well as for coincidence detectors confirms that we did not violate any fundamental assumption to the point of invalidating the prediction as a whole.

In principle, the second-order contribution (Eq. 31) can decrease much more quickly than the first-order prediction (Eq. 29) as the input correlation c becomes smaller. However, we found that the first-order prediction remained inconsistent for coincidence detectors even at relatively small values of c (Fig. 5E, left) and that including second-order terms significantly improved that prediction (Fig. 5E, right). In fact, measured C was closely matched by the second-order prediction across a broad range of c , and both were much larger than the first-order prediction except for very small c (Fig. 5F), proving that our result is not an artifact of large input correlations.

The nonlinear contribution to precise spike-time synchrony is not necessarily limited to second-order terms, and indeed the remaining systematic deviations from our second-order prediction in Figure 5, D and E, suggest that even higher-order terms contribute to synchrony, although that contribution is evidently quite small. Quantifying the n th order nonlinearity becomes more difficult for higher n , as the dimensionality of the n th moment of the spike-triggering stimuli rapidly increases, and was considered beyond the scope of the present study.

Simulations in a phenomenological coincidence detector model

For the conductance-based models described above, the difference in threshold mechanism between integrators and coincidence detectors derives from distinct spike initiation dynamics (Prescott et al., 2008a). Based on results from phenomenological models without “dynamics” (de la Rocha et al., 2007), one might suspect that the specific dynamics of the cell model are not vital for the correlation–rate relationship. The type of threshold and the dynamical mechanism responsible for “thresholding” are inextricably linked, but, nonetheless, to test whether the type of threshold is critical rather than the dynamics per se, we constructed a phenomenological coincidence detector model comprising a filter and threshold in which the filter is biphasic, like the STA in our conductance-based coincidence detector model (Fig.

6A; compare Fig. 2C). In this dynamics-free model, ν was completely independent of μ but varied with σ (Fig. 6B; compare Fig. 2A, right)—in effect, this model constitutes an “ideal” coincidence detector. Despite the first-order prediction (i.e., Eq. 29) that there should be no output correlation (given that $\partial\nu/\partial\mu = 0$) (Barreiro et al., 2010), pairs of ideal coincidence detectors receiving correlated input did indeed exhibit correlated spiking (Fig. 6C, left). Moreover, sample CCGs (Fig. 6C, right) were comparable in shape to those of the conductance-based coincidence detector model (compare Figs. 4, 5). Furthermore, the same prediction error was observed near the peak of the narrow CCGs when applying Equation 29, but our quantitative prediction was near-perfect when second-order terms (Eq. 31) were included. Overall, these results argue that the type of threshold used by a pair of neurons receiving correlated input will impact the correlation in their output.

To summarize results up to this point, ideal coincidence detectors (exemplified by our FT model) exhibit correlations arising solely from spike-time synchronization. In contrast, more realistic coincidence detectors (exemplified by our conductance-based HHLS model) exhibit a mixture of spike-time synchronization and rate comodulation, whereas realistic integrators (exemplified by our conductance-based ML model) exhibit mostly rate comodulation. The timescales of these two types of output correlation differ but nonetheless overlap. More importantly, the two types of output correlation exhibit fundamentally different sensitivities to firing rate: rate comodulation varies with firing rate, whereas spike-time synchronization does not. We predicted that real neurons should exhibit rate comodulation and spike-time synchronization if those neurons operate at least partially as coincidence detectors and, moreover, that the predominance of each type of output correlation will depend on the balance of operating modes.

Experiments in CA1 pyramidal neurons made to behave preferentially as integrators or coincidence detectors via manipulation of threshold mechanism by dynamic clamp

Integrators and coincidence detectors are found throughout the nervous system but tend to exhibit specializations beyond threshold mechanism. Because those additional specializations could confound our comparison of correlation susceptibility, we chose to compare correlation susceptibility in a single type of neuron whose threshold mechanism was experimentally manipulated such that the neuron behaves preferentially as an integrator or coincidence detector.

Pyramidal neurons, including those in the CA1 region of hippocampus, display the hallmarks of integrators when recorded in acute brain slices but behave more like coincidence detectors upon introduction of a virtual leak conductance by dynamic clamp (Fig. 7A), consistent with a predicted switch in threshold mechanism (Prescott et al., 2006, 2008b). Thus, by manipulating the membrane properties of CA1 pyramidal neurons, we were able to compare correlation susceptibility in a single cell type operating preferentially in one or the other mode—the shift in operating mode is quantitative, not absolute. In the interests of comparing experimental and simulation data [and to compare with past studies (de la Rocha et al. (2007))], we applied the same stimulation paradigm used in simulations (i.e., noisy current injection) together with the simplest dynamic-clamp manipulation capable of switching the threshold mechanism (i.e., constant leak conductance).

Compared with STAs in the integrator-mode (Fig. 7A, left), STAs in the coincidence detector-mode (Fig. 7A, right) had a

←

(Figure legend continued.) precisely synchronized spikes. **D**, Equation 31 describes a second-order correction based on c^2 . Prediction based on first- and second-order statistics (i.e., Eqs. 29, 31) gave dramatically improved accuracy near the peak of the coincidence detector CCG. Similar albeit smaller improvement was observed for the integrator (data not shown). According to the measured versus predicted C/c plot, inclusion of second-order terms could largely account for output correlation in the coincidence detector where the first-order prediction had failed. In **A–D**, input correlation $c = 0.5$. **E**, Same plots as in **A** and **D** for the coincidence detector but with input correlation $c = 0.2$. The red arrows indicate the same data marked in **A** and **D**. **F**, Success of the second-order prediction in accounting for measured output covariance across a broad range of input correlations, compared with failure of the first-order prediction, argues that output covariance among coincidence detectors is dominated by second-order terms except when input correlation c is extremely small ($\ll 0.1$). These show that the results in **A–E** qualitatively hold for a wide range of input correlation.

higher peak height, and were therefore steeper than integrator STAs, but the negative phase was far less prominent than in STAs from our coincidence detectors models (compare Fig. 2C). This is likely due to spike initiation in real neurons being influenced by more currents with gating variables spanning a broader range of timescales than were included in our minimal conductance-based models. In any case, as we would predict, the steep weakly biphasic STAs are consistent with the narrow weakly biphasic CCGs observed for pyramidal neurons in the coincidence detector-mode (see below). Most importantly for our purposes, firing rate gain with respect to μ , which is proportional to how biphasic the STA is (Eq. 20), was significantly lower in the coincidence detector-mode compared with the default integrator-mode (0.14 ± 0.041 vs 0.05 ± 0.016 Hz/pA; $p < 0.01$, t test), whereas firing rate gain with respect to σ was similar between modes (0.034 ± 0.022 vs 0.022 ± 0.011 Hz/pA), which means $\partial v / \partial \sigma$ was, relative to $\partial v / \partial \mu$, higher in the coincidence detector-mode (0.25 ± 0.16 vs 0.50 ± 0.29 ; $p < 0.01$, t test). This is consistent with the differential response properties reported for the conductance-based models in Figure 2A. Like for simulation data in Figure 5, we plotted our experimental data as measured versus predicted C based on the first-order prediction described in Equation 29 (Fig. 7B, top). As expected, the prediction error was much greater for the coincidence detector-mode than for the integrator-mode, which was also evident in the sample CCGs (Fig. 7B, bottom). Consistent with simulations (compare Figs. 4, 5, CCGs), the CCG for the coincidence detector-mode was much narrower than for the integrator-mode, and the first-order prediction deviated from the measured CCG primarily at its peak (Fig. 7B, bottom). As for simulation data, we hypothesized that removing a narrow region around the peak of the CCG (where the actual and predicted CCGs differ most) would dramatically improve the prediction as visualized on the measured versus predicted C plots, which indeed it did (Fig. 7C). Similarly, including the second-order terms described by Equation 31 improved our prediction, especially for the coincidence detector-mode (Fig. 7D). These experiments demonstrate that pyramidal neurons receiving shared input can exhibit output correlation in excess of that predicted from firing rate comodulation, and that this spike-time synchronization is greater when the neurons behave more like coincidence detectors.

Discussion

Through simulations and experiments, we have shown that pairs of neurons receiving correlated input can exhibit spiking that is

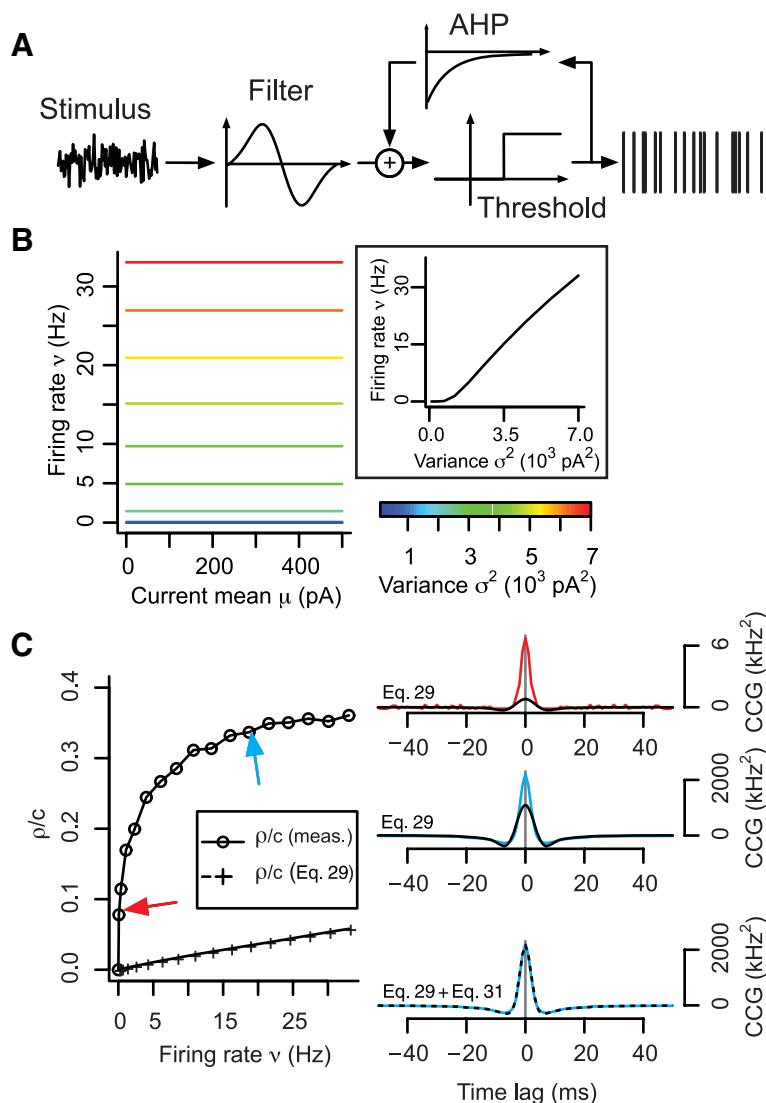


Figure 6. Phenomenological model sensitive only to stimulus variance exhibits correlated spiking due uniquely to spike-time synchronization. **A**, FT model was constructed by coupling a biphasic filter (reminiscent of the coincidence detector STA; Fig. 2C) and a step-function representing threshold. **B**, Because of its filter properties, the FT model is uniquely sensitive to stimulus fluctuations. Firing rate gain with respect to μ is zero but is positive with respect to σ . **C**, Using the stimulation paradigm shown in Figure 1A, a pair of FT models were given correlated input with equal μ and σ ; $c = 0.5$; $\tau = 5$ ms. Measured p/c values increased with v , despite the first-order prediction (Eq. 29) giving a much smaller correlation, which vanishes as the time window size T increases (data not shown) on the basis of $\partial v / \partial \mu$ equaling 0. The arrows point to data for which sample CCGs are shown. Predicted CCGs (black) deviated from measured CCGs (color) in the same manner as when the first-order prediction (Eq. 29) was applied to the conductance-based coincidence detector model in Figure 5A. When second-order terms were included (Eq. 31), the predicted CCG was much more accurate (bottom).

correlated in different ways, and that the overall contribution of each type of output correlation depends on intrinsic cellular properties. Ideal integrators exhibit output correlations comprised entirely of rate comodulation. This makes sense given that individual integrators use rate encoding in which spiking is dictated by the mean (i.e., a first-order statistic) of the input; therefore, input can be reconstructed entirely by applying a first-order filter (i.e., the STA) to the spike train. Conversely, ideal coincidence detectors exhibit output correlations comprised entirely of spike-time synchronization (Fig. 6). However, realistic coincidence detectors exhibit output correlations comprised partly of spike-time synchronization and partly of rate comodulation (Fig. 5). This too makes sense given that individual coincidence detectors use temporal encoding in which spiking is sensitive to the

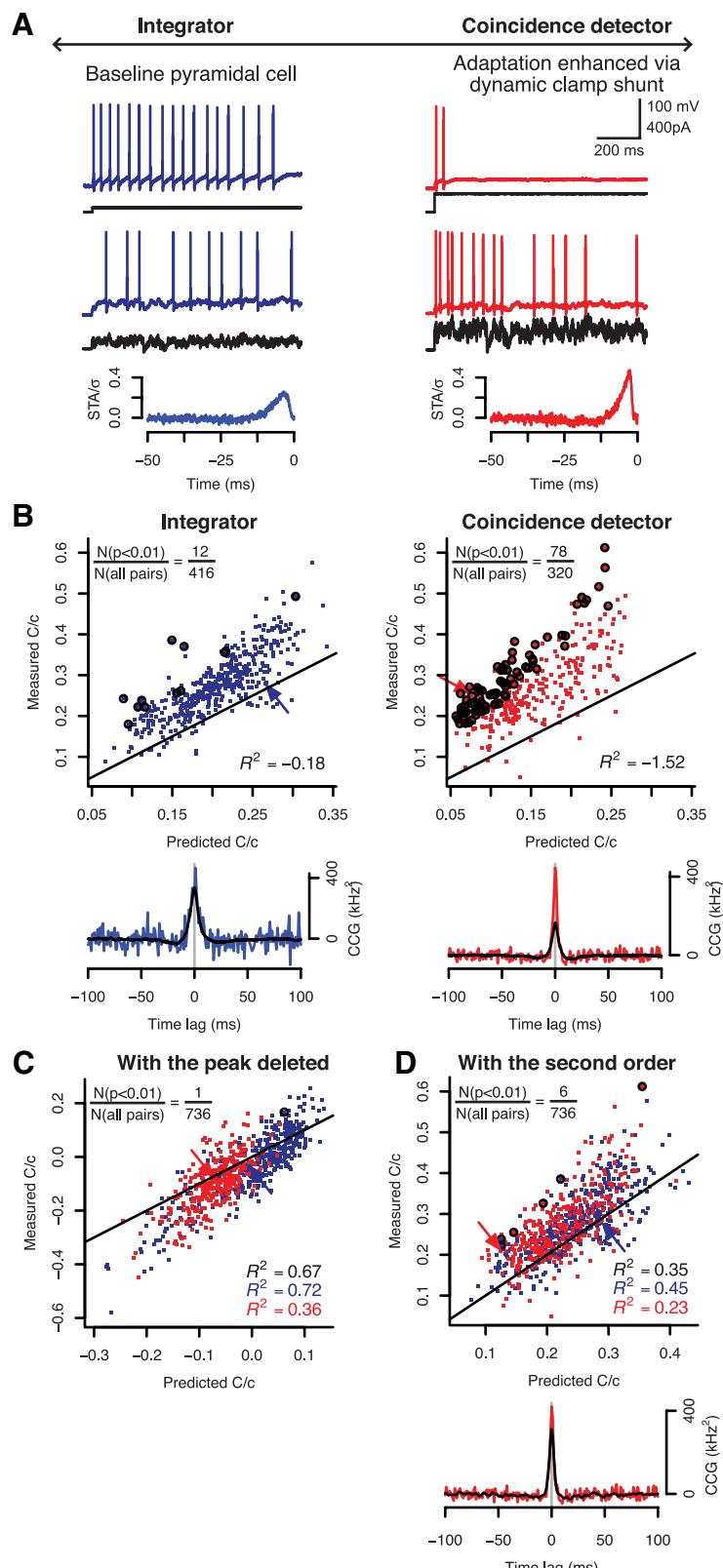


Figure 7. CA1 pyramidal neurons operating preferentially in a coincidence detector-mode exhibit greater spike-time synchronization than when operating in an integrator-mode. **A**, Responses from a typical regular spiking pyramidal neuron to constant and fluctuating ($\tau = 5$ ms) input are shown without any manipulation (left) and with a virtual shunt (10 nS, -70 mV reversal potential) inserted by dynamic clamp (right). We used the shunt to enhance adaptation by depolarizing threshold, thereby shifting the neuron from acting preferentially as an integrator to acting preferentially as a coincidence detector, as shown previously (Prescott et al., 2006). Trials with and without dynamic clamp were interleaved to assess each neuron in both operating modes. The bottom panels show representative STA from each mode from one of the neurons comprising the pair used to sample CCGs (as indicated by arrows) in **B–D**. **B**, Following the same approach used for conductance-based models (Fig. 5), we plotted measured and predicted C/c for each “preferred” operating mode. The prediction here is based on Equation 29. The circles identify those points that deviated significantly from prediction ($p < 0.01$, t test). In the integrator-mode, only 12 of 416 points showed significant deviation of the measured C/c from prediction, while significantly more points (78 of 320) showed significant deviation in the coincidence detector-mode ($p < 0.001$, χ^2 test); in other words, the first-order prediction was less consistent for the coincidence detector-mode. The arrows point to data for which sample CCGs are shown. Predicted CCGs (black) deviated from measured CCGs (color) the same way as in Figure 5A. Data are from a total of 12 cells. **C**, Measured and predicted C/c values for both operating modes were replotted after removing data ± 2 ms around the peak in the CCGs. As in Figure 5C, excluding excess synchrony not predicted on the basis of rate comodulation gave a closer match between measured and predicted C/c values, leaving only one data point (circled) that deviated significantly from prediction ($p < 0.01$, t test). **D**, Likewise, inclusion of second-order terms (i.e., Eq. 31) in our prediction gave a dramatically improved match, leaving only three data points (circled) for each operating mode with significant deviation from prediction ($p < 0.01$, t test). The improved prediction of spike-time synchrony afforded by inclusion of the second-order correction is also evident from the sample CCG (compare Fig. 5D).

variance (i.e., a second-order statistic) of the input; therefore, reconstructing the input from the spike train requires inclusion of first- and second-order filters (Theunissen and Miller, 1995). Real neurons and realistic conductance-based models are never pure integrators or coincidence detectors. We have juxtaposed the two operating modes for didactic purposes, but, ultimately, operating mode is a continuum representing the interplay between stimulus kinetics and neural dynamics. Neural dynamics—most notably spike generation—differ between neurons, can be modulated within a given neuron, and directly impact how a neuronal ensemble encodes and transmits information.

Some cell types are optimized for integration or coincidence detection; for example, many neurons early in the auditory pathway (e.g., cochlear nucleus and superior olive) are exquisite coincidence detectors (Manis and Marx, 1991; Cao et al., 2007; Mathews et al., 2010), whereas certain neurons in the entorhinal cortex are near-perfect integrators (Egorov et al., 2002). In addition to morphological specializations (Agmon-Snir et al., 1998), a broad range of ion currents can influence operating mode, the common feature being that such currents activate or inactivate at voltages near threshold, or that they impact the gating of other peri-threshold currents, for example, by shifting voltage threshold (Prescott et al., 2006,

indicated by arrows) in **B–D**. **B**, Following the same approach used for conductance-based models (Fig. 5), we plotted measured and predicted C/c for each “preferred” operating mode. The prediction here is based on Equation 29. The circles identify those points that deviated significantly from prediction ($p < 0.01$, t test). In the integrator-mode, only 12 of 416 points showed significant deviation of the measured C/c from prediction, while significantly more points (78 of 320) showed significant deviation in the coincidence detector-mode ($p < 0.001$, χ^2 test); in other words, the first-order prediction was less consistent for the coincidence detector-mode. The arrows point to data for which sample CCGs are shown. Predicted CCGs (black) deviated from measured CCGs (color) the same way as in Figure 5A. Data are from a total of 12 cells. **C**, Measured and predicted C/c values for both operating modes were replotted after removing data ± 2 ms around the peak in the CCGs. As in Figure 5C, excluding excess synchrony not predicted on the basis of rate comodulation gave a closer match between measured and predicted C/c values, leaving only one data point (circled) that deviated significantly from prediction ($p < 0.01$, t test). **D**, Likewise, inclusion of second-order terms (i.e., Eq. 31) in our prediction gave a dramatically improved match, leaving only three data points (circled) for each operating mode with significant deviation from prediction ($p < 0.01$, t test). The improved prediction of spike-time synchrony afforded by inclusion of the second-order correction is also evident from the sample CCG (compare Fig. 5D).

2008a). Whether pyramidal neurons in the neocortex and hippocampus function as integrators or coincidence detectors has been controversial (Abeles, 1982; Softky and Koch, 1993; Shadlen and Newsome, 1998). The answer depends not only on intrinsic cell properties but also on stimulus conditions (Rudolph and Destexhe, 2003) and other external factors like background synaptic activity that influence membrane conductance (Destexhe et al., 2003). Shunting and adaptation, especially in combination, can encourage coincidence detection in pyramidal neurons that might otherwise behave preferentially as integrators (Prescott et al., 2006, 2008b). The adaptation current I_M has been shown to encourage coincidence detection in hippocampal and neocortical pyramidal neurons (Hu et al., 2007; Guan et al., 2011) and is itself subject to endogenous modulators and to several drugs (for review, see Brown and Passmore, 2009). This supports the view that pyramidal neurons (and presumably other cell types through adaptation by I_M and via other mechanisms) can shift between operating modes. By not operating at one or the other extreme, pyramidal neurons can flexibly use rate- and/or synchrony-based coding (see below).

The coexistence of rate- and synchrony-based coding is consistent with recent modeling work showing that differently encoded information can be simultaneously propagated through feedforward networks depending on network properties (Kremkow et al., 2010; Kumar et al., 2010). Our results emphasize the importance of cellular properties for exactly the same issues. Network and cellular properties can be intimately related (see above regarding synaptic input and membrane conductance) and may interact nonlinearly such that forms of modulation, which individually have weak effects, combine to produce powerful gating mechanisms that switch a network between propagating synchronous or asynchronous activity (see below). But it is not necessarily the case that synchronous and asynchronous spikes act independently; for example, in spiny stellate cells receiving thalamocortical input via a limited number of weak synapses, asynchronous background cortical input may be crucial for setting the membrane potential of stellate cells close enough to spike threshold that synchronous thalamic inputs can elicit spikes (Douglas and Martin, 2007; da Costa and Martin, 2011). In that example, asynchronous input may facilitate the propagation of synchronous spiking and, furthermore, might act as a continuously variable gain-control mechanism rather than as a simple on–off switch.

An important related issue is correlation between excitatory and inhibitory input. This has been proposed as a mechanism to regulate information transmission (Salinas and Sejnowski, 2001; Kremkow et al., 2010), amplify external inputs (Murphy and Miller, 2009), and enhance response fidelity (Wehr and Zador, 2003; Cafaro and Rieke, 2010). Such correlations may exist with delays of only a few milliseconds, likely because of feedforward inhibition ensuring a narrow integration window (Wehr and Zador, 2003; Higley and Contreras, 2006). Under those conditions, presynaptic spikes synchronized with millisecond precision are required to reliably evoke responses in postsynaptic neurons.

Information encoded by rate or synchrony must be reliably transmitted (when allowed by gating mechanisms) and eventually decoded. Unambiguous decoding requires independent rate- and correlation-based coding (Fig. 1). Recent findings (de la Rocha et al., 2007) have cast doubt on this independence. We concur that a correlation–rate relationship compromises correlation-based coding based on rate comodulation, but we demonstrate here that output correlation and rate can be correlated without interfering with synchrony-based coding. Count-

ing spikes neglects the information carried by spike timing. The absolute timing of spikes in a single presynaptic cell has little impact on postsynaptic activation, but the relative timing of spikes across multiple presynaptic cells is crucial if the postsynaptic cell operates as a coincidence detector; by extension, the absolute timing of synchronized volleys likely carries information about the original stimulus.

It follows that integrators exhibit the correlation–rate relationship described by de la Rocha et al. (2007), but that same relationship, although it can be observed in coincidence detectors (Barreiro et al., 2010), applies to only a fraction (i.e., to only one component) of their output correlation. Notably, de la Rocha et al. used computer models in which spike generation depended on mean input (i.e., integration) and the 1-s-long stimuli used in their experiments were sufficiently short that slow adaptation was minimal, thus favoring spike generation based on mean input (Prescott and Sejnowski, 2008). In addition to testing “integrator” models, we tested conductance-based and phenomenological models in which spike generation depends on input variance (i.e., coincidence detector models). Our experiments were conducted with 300-s-long stimuli to favor adaptation, our rationale being that neurons *in vivo* exist in a chronically depolarized and shunted state because of synaptic bombardment (Destexhe et al., 2003). By testing a single cell type under different “virtual network conditions” rather than testing different cell types representing each extreme of the continuum between integration and coincidence detection, our results show that the neuronal operating mode can be shifted along that continuum and that this shift in operating mode adjusts the balance of rate- and synchrony-based coding.

These results demonstrate the importance of cell-level properties for network-level coding. There has been a longstanding bias in the network modeling community to focus on synaptic weights and network architecture, with far less emphasis put on cellular properties. Here, we have shown what happens when neurons do not “integrate and fire.” Similarly, Burak et al. (2009) and Barreiro et al. (2010) recently showed differences in output correlation depending on the type of model used. Effects of synaptic kinetics, background noise, and input spike train statistics on output correlation have also been documented (Maex et al., 2000; Tetzlaff et al., 2008; Ostojic et al., 2009). Tchumatchenko et al. (2010) have shown specifically that output correlations can be rate dependent or independent according to input conditions. Our results emphasize how both types of correlations can coexist and that this depends on input conditions and (how that input is encoded given) single cell firing properties. Indeed, whereas simplified models tend to favor pure integration or pure coincidence detection, real neurons and realistic conductance-based models exhibit a context-dependent mixture of operating modes. Neglecting the richness of cell-level coding properties will surely translate into underestimation of network-level coding possibilities.

In conclusion, different types of neurons or even the same neuron operating under different conditions can be differentially sensitive to first- and second-order stimulus statistics, namely mean and variance. If we do not assume that fluctuations are simply noise, and if a neuron is sensitive to those fluctuations (which is true of coincidence detectors), then output spiking should carry information about the signal variance. Two such neurons receiving correlated input may carry information about signal variance in their precisely synchronized spiking. Even if correlated with firing rate (indeed, such output correlations might arise from correlations between features of the original input), spike-time synchronization is separate from rate co-

modulation, which, by definition, is directly linked to the firing rate that is tuned to the signal mean. Both forms of correlation may coexist and operate independently. Thus, the richness of single-neuron coding abilities translates into even richer multi-neuron coding possibilities.

References

- Abeles M (1982) Role of the cortical neuron: integrator or coincidence detector? *Isr J Med Sci* 18:83–92.
- Abeles M (1991) *Corticonics: neural circuits of the cerebral cortex*. Cambridge, MA: Cambridge UP.
- Aertsen A, Diesmann M, Gewaltig MO (1996) Propagation of synchronous spiking activity in feedforward neural networks. *J Physiol Paris* 90:243–247.
- Agmon-Snir H, Carr CE, Rinzel J (1998) The role of dendrites in auditory coincidence detection. *Nature* 393:268–272.
- Arsiero M, Lüscher HR, Lundstrom BN, Giugliano M (2007) The impact of input fluctuations on the frequency–current relationships of layer 5 pyramidal neurons in the rat medial prefrontal cortex. *J Neurosci* 27:3274–3284.
- Averbeck BB, Latham PE, Pouget A (2006) Neural correlations, population coding and computation. *Nat Rev Neurosci* 7:358–366.
- Bair W, Zohary E, Newsome WT (2001) Correlated firing in macaque visual area MT: time scales and relationship to behavior. *J Neurosci* 21:1676–1697.
- Barreiro AK, Shea-Brown E, Thilo EL (2010) Time scales of spike-train correlation for neural oscillators with common drive. *Phys Rev E Stat Nonlin Soft Matter Phys* 81:011916.
- Bialek W, de Ruyter Van Steveninck RR (1988) Real-time performance of a movement-sensitive neuron in the blowfly visual system: coding and information transfer in short spike sequences. *Proc R Soc Lond B Biol Sci* 234:379–414.
- Brown DA, Passmore GM (2009) Neural KCNQ (Kv7) channels. *Br J Pharmacol* 156:1185–1195.
- Burak Y, Lewallen S, Sompolinsky H (2009) Stimulus-dependent correlations in threshold-crossing spiking neurons. *Neural Comput* 21:2269–2308.
- Cafaro J, Rieke F (2010) Noise correlations improve response fidelity and stimulus encoding. *Nature* 468:964–967.
- Cao XJ, Shatadal S, Oertel D (2007) Voltage-sensitive conductances of bushy cells of the mammalian ventral cochlear nucleus. *J Neurophysiol* 97:3961–3975.
- Chialvo DR, Longtin A, Müller-Gerking J (1997) Stochastic resonance in models of neuronal ensembles. *Phys Rev E Stat Nonlin Soft Matter Phys* 55:1798–1808.
- da Costa NM, Martin KA (2011) How thalamus connects to spiny stellate cells in the cat's visual cortex. *J Neurosci* 31:2925–2937.
- de la Rocha J, Doiron B, Shea-Brown E, Josić K, Reyes A (2007) Correlation between neural spike trains increases with firing rate. *Nature* 448:802–806.
- Destexhe A, Rudolph M, Fellous JM, Sejnowski TJ (2001) Fluctuating synaptic conductances recreate in vivo-like activity in neocortical neurons. *Neuroscience* 107:13–24.
- Destexhe A, Rudolph M, Paré D (2003) The high-conductance state of neocortical neurons in vivo. *Nat Rev Neurosci* 4:739–751.
- Douglas RJ, Martin KA (2007) Mapping the matrix: the ways of neocortex. *Neuron* 56:226–238.
- Egorov AV, Hamam BN, Fransén E, Hasselmo ME, Alonso AA (2002) Graded persistent activity in entorhinal cortex neurons. *Nature* 420:173–178.
- Engel AK, Roelfsema PR, Fries P, Brecht M, Singer W (1997) Role of the temporal domain for response selection and perceptual binding. *Cereb Cortex* 7:571–582.
- Fellous JM, Rudolph M, Destexhe A, Sejnowski TJ (2003) Synaptic background noise controls the input/output characteristics of single cells in an *in vitro* model of *in vivo* activity. *Neuroscience* 122:811–829.
- Fourcaud-Trocmé N, Hansel D, van Vreeswijk C, Brunel N (2003) How spike generation mechanisms determine the neuronal response to fluctuating inputs. *J Neurosci* 23:11628–11640.
- Gerstner W, Kreiter AK, Markram H, Herz AV (1997) Neural codes: firing rates and beyond. *Proc Natl Acad Sci U S A* 94:12740–12741.
- Guan D, Higgs MH, Horton LR, Spain WJ, Foehring RC (2011) Contributions of Kv7-mediated potassium current to sub- and suprathreshold responses of rat layer II/III neocortical pyramidal neurons. *J Neurophysiol* 106:1722–1733.
- Higgs MH, Slee SJ, Spain WJ (2006) Diversity of gain modulation by noise in neocortical neurons: regulation by the slow afterhyperpolarization conductance. *J Neurosci* 26:8787–8799.
- Higley MJ, Contreras D (2006) Balanced excitation and inhibition determine spike timing during frequency adaptation. *J Neurosci* 26:448–457.
- Hines ML, Carnevale NT (1997) The NEURON simulation environment. *Neural Comput* 9:1179–1209.
- Hong S, Lundstrom BN, Fairhall AL (2008) Intrinsic gain modulation and adaptive neural coding. *PLoS Comput Biol* 4:e1000119.
- Hu H, Vervaeke K, Storm JF (2007) M-channels (Kv7/KCNQ channels) that regulate synaptic integration, excitability, and spike pattern of CA1 pyramidal cells are located in the perisomatic region. *J Neurosci* 27:1853–1867.
- König P, Engel AK, Singer W (1996) Integrator or coincidence detector? The role of the cortical neuron revisited. *Trends Neurosci* 19:130–137.
- Kremkow J, Aertsen A, Kumar A (2010) Gating of signal propagation in spiking neural networks by balanced and correlated excitation and inhibition. *J Neurosci* 30:15760–15768.
- Kumar A, Rotter S, Aertsen A (2010) Spiking activity propagation in neuronal networks: reconciling different perspectives on neural coding. *Nat Rev Neurosci* 11:615–627.
- Lundstrom BN, Hong S, Higgs MH, Fairhall AL (2008) Two computational regimes of a single-compartment neuron separated by a planar boundary in conductance space. *Neural Comput* 20:1239–1260.
- Maex R, Vos BP, De Schutter E (2000) Weak common parallel fibre synapses explain the loose synchrony observed between rat cerebellar Golgi cells. *J Physiol* 523:175–192.
- Mainen ZF, Sejnowski TJ (1995) Reliability of spike timing in neocortical neurons. *Science* 268:1503–1506.
- Manis PB, Marx SO (1991) Outward currents in isolated ventral cochlear nucleus neurons. *J Neurosci* 11:2865–2880.
- Mathews PJ, Jercog PE, Rinzel J, Scott LL, Golding NL (2010) Control of submillisecond synaptic timing in binaural coincidence detectors by K_v1 channels. *Nat Neurosci* 13:601–609.
- Meister M, Berry MJ (1999) The neural code of the retina. *Neuron* 22:435–450.
- Murphy BK, Miller KD (2009) Balanced amplification: a new mechanism of selective amplification of neural activity patterns. *Neuron* 61:635–648.
- Ostojic S, Brunel N, Hakim V (2009) How connectivity, background activity, and synaptic properties shape the cross-correlation between spike trains. *J Neurosci* 29:10234–10253.
- Palanca BJ, DeAngelis GC (2005) Does neuronal synchrony underlie visual feature grouping? *Neuron* 46:333–346.
- Perkel DH, Gerstein GL, Moore GP (1967) Neuronal spike trains and stochastic point processes. II. Simultaneous spike trains. *Biophys J* 7:419–440.
- Pinto RD, Elson RC, Szücs A, Rabinovich MI, Selverston AI, Abarbanel HD (2001) Extended dynamic clamp: controlling up to four neurons using a single desktop computer and interface. *J Neurosci Methods* 108:39–48.
- Prescott SA, De Koninck Y (2009) Impact of background synaptic activity on neuronal response properties revealed by stepwise replication of *in vivo*-like conditions *in vitro*. In: *The dynamic clamp: from principles to applications* (Destexhe A, Bal T, eds), pp 89–114. New York: Springer.
- Prescott SA, Sejnowski TJ (2008) Spike-rate coding and spike-time coding are affected oppositely by different adaptation mechanisms. *J Neurosci* 28:13649–13661.
- Prescott SA, Ratté S, De Koninck Y, Sejnowski TJ (2006) Nonlinear interaction between shunting and adaptation controls a switch between integration and coincidence detection in pyramidal neurons. *J Neurosci* 26:9084–9097.
- Prescott SA, De Koninck Y, Sejnowski TJ (2008a) Biophysical basis for three distinct dynamical mechanisms of action potential initiation. *PLoS Comput Biol* 4:e1000198.
- Prescott SA, Ratté S, De Koninck Y, Sejnowski TJ (2008b) Pyramidal neurons switch from integrators *in vitro* to resonators *under in vivo*-like conditions. *J Neurophysiol* 100:3030–3042.
- Reyes AD (2003) Synchrony-dependent propagation of firing rate in iteratively constructed networks *in vitro*. *Nat Neurosci* 6:593–599.

- Rieke F, Warland D, Bialek W (1999) Spikes: exploring the neural code. Cambridge, MA: MIT.
- Rudolph M, Destexhe A (2003) Tuning neocortical pyramidal neurons between integrators and coincidence detectors. *J Comput Neurosci* 14:239–251.
- Salinas E, Sejnowski TJ (2001) Correlated neuronal activity and the flow of neural information. *Nat Rev Neurosci* 2:539–550.
- Schneidman E, Berry MJ 2nd, Segev R, Bialek W (2006) Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature* 440:1007–1012.
- Schreiber S, Fellous JM, Tiesinga P, Sejnowski TJ (2004) Influence of ionic conductances on spike timing reliability of cortical neurons for suprathreshold rhythmic inputs. *J Neurophysiol* 91:194–205.
- Shadlen MN, Movshon JA (1999) Synchrony unbound: a critical evaluation of the temporal binding hypothesis. *Neuron* 24:67–77, 111–125.
- Shadlen MN, Newsome WT (1998) The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J Neurosci* 18:3870–3896.
- Shea-Brown E, Josić K, de la Rocha J, Doiron B (2008) Correlation and synchrony transfer in integrate-and-fire neurons: basic properties and consequences for coding. *Phys Rev Lett* 100:108102.
- Singer W (1993) Synchronization of cortical activity and its putative role in information processing and learning. *Annu Rev Physiol* 55:349–374.
- Softky WR, Koch C (1993) The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *J Neurosci* 13:334–350.
- Svirskis G, Kotak V, Sanes DH, Rinzel J (2004) Sodium along with low-threshold potassium currents enhance coincidence detection of sub-threshold noisy signals in MSO neurons. *J Neurophysiol* 91:2465–2473.
- Tchumatchenko T, Malyshev A, Geisel T, Volgshev M, Wolf F (2010) Correlations and synchrony in threshold neuron models. *Phys Rev Lett* 104:058102.
- Tetzlaff T, Rotter S, Stark E, Abeles M, Aertsen A, Diesmann M (2008) Dependence of neuronal correlations on filter characteristics and marginal spike train statistics. *Neural Comput* 20:2133–2184.
- Theunissen F, Miller JP (1995) Temporal encoding in nervous systems: a rigorous definition. *J Comput Neurosci* 2:149–162.
- Tiesinga P, Fellous JM, Sejnowski TJ (2008) Regulation of spike timing in visual cortical circuits. *Nat Rev Neurosci* 9:97–107.
- Treisman A (1999) Solutions to the binding problem: progress through controversy and convergence. *Neuron* 24:105–110, 111–125.
- Uhlenbeck GE, Ornstein LS (1930) On the theory of Brownian motion. *Phys Rev* 36:823–841.
- Victor J, Shapley R (1980) A method of nonlinear analysis in the frequency domain. *Biophys J* 29:459–483.
- Wehr M, Zador AM (2003) Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. *Nature* 426:442–446.
- Wolfe J, Houweling AR, Brecht M (2010) Sparse and powerful cortical spikes. *Curr Opin Neurobiol* 20:306–312.
- Zohary E, Shadlen MN, Newsome WT (1994) Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature* 370:140–143.

Efficient estimation of phase-response curves via compressive sensing

Sungho Hong,¹ Quinten Robberechts,² and Erik De Schutter^{1,2}

¹Computational Neuroscience Unit, Okinawa Institute of Science and Technology, Onna, Onna-son, Okinawa, Japan;
and ²Theoretical Neurobiology, University of Antwerp, Antwerp, Belgium

Submitted 11 October 2011; accepted in final form 19 June 2012

Hong S, Robberechts Q, De Schutter E. Efficient estimation of phase-response curves via compressive sensing. *J Neurophysiol* 108: 2069–2081, 2012. First published June 20, 2012; doi:10.1152/jn.00919.2011.—The phase-response curve (PRC), relating the phase shift of an oscillator to external perturbation, is an important tool to study neurons and their population behavior. It can be experimentally estimated by measuring the phase changes caused by probe stimuli. These stimuli, usually short pulses or continuous noise, have a much wider frequency spectrum than that of neuronal dynamics. This makes the experimental data high dimensional while the number of data samples tends to be small. Current PRC estimation methods have not been optimized for efficiently discovering the relevant degrees of freedom from such data. We propose a systematic and efficient approach based on a recently developed signal processing theory called compressive sensing (CS). CS is a framework for recovering sparsely constructed signals from undersampled data and is suitable for extracting information about the PRC from finite but high-dimensional experimental measurements. We illustrate how the CS algorithm can be translated into an estimation scheme and demonstrate that our CS method can produce good estimates of the PRCs with simulated and experimental data, especially when the data size is so small that simple approaches such as naive averaging fail. The tradeoffs between degrees of freedom vs. goodness-of-fit were systematically analyzed, which help us to understand better what part of the data has the most predictive power. Our results illustrate that finite sizes of neuroscientific data in general compounded by large dimensionality can hamper studies of the neural code and suggest that CS is a good tool for overcoming this challenge.

phase-response curve; compressive sensing; neural coding; small data size; cerebellar Golgi cell

THE PHASE-RESPONSE CURVE CHARACTERIZES how an oscillation is influenced by external input delivered at a particular phase and is one of the most important tools for analyzing biological oscillators, such as neurons and neural ensembles (Achuthan et al. 2011; Netoff et al. 2005; Oprisan et al. 2004; Perkel et al. 1964; Pinsker 1977; Preyer and Butera 2005; Tateno and Robinson 2007; Winfree 2001). The key part of this characterization is the infinitesimal phase-response curve (IPRC or simply PRC), which is the linear kernel relating the infinitesimal and instantaneous stimulus with phase (Achuthan et al. 2011; Preyer and Butera 2005). At the level of a single neuron, the PRC characterizes the intrinsic properties of the neuron and determines how its spiking pattern will change with input. For example, whether the PRC is monophasic or biphasic indicates whether the neuron is an integrator or resonator based on its excitability class (Ermentrout 1996; Gutkin et al. 2005; Hansel et al. 1995; Izhikevich 2010; Moehlis et al. 2006; Phoka et al. 2010; Stiefel et al. 2008). Beyond the single cell level, the PRC

Address for reprint requests and other correspondence: S. Hong, Computational Neuroscience Unit, Okinawa Institute of Science and Technology, 7542 Onna, Onna-son, Okinawa 904-0411, Japan (e-mail: shhong@oist.jp).

has been useful for predicting population behavior such as synchrony. In a weakly coupled neural network, the PRCs of the individual neurons determine the interactions and many aspects of the network dynamics can thereby be inferred (Ermentrout et al. 2001; Hansel et al. 1995; Oprisan et al. 2004; Smeal et al. 2010; Teramae and Fukai 2008). The PRC has also been used to show how synchrony is generated within a population of uncoupled oscillators receiving common noise input (Abouzeid and Ermentrout 2009; Goldobin and Pikovsky 2005; Marella and Ermentrout 2008; Nakao et al. 2005; Teramae and Tanaka 2004).

Therefore, it is important to determine the PRCs of neurons, particularly using experimental data, to predict their population/network behavior. This can be done by measuring the phase shifts with respect to probe stimuli, and several estimation methods have been proposed (reviewed in Torben-Nielsen et al. 2010). Because the PRC is a function of phase and therefore also of time, its determination with sufficient temporal resolution requires the stimuli to contain high frequency components as explained by the standard time/frequency uncertainty principle (Pinsky 2009). Indeed, in all of the existing methods, the probe stimuli are given as short pulses or continuous noise, which have strong power at high frequencies. However, the time scales of the PRC are caused by neuronal dynamics that reside in a lower frequency range. Therefore, a priori, the degrees of freedom, or dimensionality, of the data (probe stimuli) are much larger than those actually constituting the PRC, and this becomes one of the main challenges in PRC estimation. Additional problems arise in practice when the firing is not perfectly regular and the period has to be estimated, rather than directly measured, together with the phase shifts (Phoka et al. 2010).

In principle, if one has infinitely large data, the extra irrelevant degrees of freedom can be simply averaged out. However, the required number of samples rapidly grows with the data dimensionality, and in practice this scheme is not feasible in most experimental settings. A possible solution is to describe the candidate PRC in terms of only arbitrarily selected low frequency component, which forces reduced data dimensionality by hand, and find the PRC by minimizing an error function (Achuthan et al. 2011; Galán et al. 2005; Izhikevich 2010; Preyer and Butera 2005; Torben-Nielsen et al. 2010). Also, as this problem is a typical case of the curse of dimensionality (Bishop 2007), a Bayesian estimation method assuming a smooth PRC distribution as prior has been proposed (Nakae et al. 2010).

In this study, we propose a more systematic and efficient approach based on recent progress in signal processing theory called compressive sensing (CS; Baraniuk 2007; Candès and Tao 2005, 2006; Donoho 2006) (see also Ganguly and Som-

polinsky 2012). CS provides near-optimal recovery algorithms for signals from a small number of samplings when the signal is sparsely constructed, i.e., composed of much less degrees of freedom than the measured samples. Because our main goal is to efficiently discover a small number of frequency components relevant to the PRC out of the much wider stimulus spectrum, we demonstrate that our problem can be translated into the sparse signal recovery theory covered by the CS framework. We also show that our CS-based estimation

method works successfully with simulated data from two different conductance-based models and experimental data recorded from cerebellar Golgi neurons (GoCs).

MATERIALS AND METHODS

Neuron Models and Simulations Procedure

The first neuron model is based on Morris-Lecar (ML) with the following kinetics (Morris and Lecar 1981)

$$C \frac{dV}{dt} = -g_{Na}m_\infty(V - E_{Na}) - g_Kw(V - E_K) - g_{Leak}(V - E_{Leak}) + (I_0 + I_{ext})/A,$$

$$\tau_w(V) \frac{dw}{dt} = w - w_\infty(V), \quad \tau_w(V) = \frac{\varphi}{\cosh\left(\frac{V - \beta_w}{2\gamma_w}\right)},$$

$$m_\infty(V) = \{1 + \tanh[(V - \beta_m)/\gamma_m]\}/2, \quad w_\infty(V) = \{1 + \tanh[(V - \beta_w)/\gamma_w]\}/2,$$

We used the parameters in Prescott et al. (2008), $g_{Leak} = 2 \text{ mS/cm}^2$, $g_{Na} = 20 \text{ mS/cm}^2$, $g_K = 20 \text{ mS/cm}^2$, $C = 2 \mu\text{F/cm}^2$, $E_K = -100 \text{ mV}$, $E_{Leak} = -70 \text{ mV}$, $E_{Na} = 50 \text{ mV}$, $\phi = 0.15$, $\beta_m = -1.2 \text{ mV}$, $\gamma_m = 18 \text{ mV}$, $\beta_w = 0 \text{ mV}$, and $\gamma_w = 10 \text{ mV}$. The cell surface area is $A =$

$100 \mu\text{m}^2$, and we set $I_0 = 369 \text{ pA}$ so that the mean interspike interval (ISI) size is $T_0 = 50.2 \text{ ms}$.

The second one is the standard Hodgkin-Huxley (HH) model (Hodgkin and Huxley 1952)

$$C \frac{dV}{dt} = -g_{Na}m^3h(V - E_{Na}) - g_Kn^4(V - E_K) - g_{Leak}(V - E_{Leak}) + (I_0 + I_{ext})/A,$$

$$\tau_z(V) \frac{dz}{dt} = z - z_\infty(V), \quad \tau_z = (\alpha_z + \beta_z), \quad z_\infty = \alpha_z/\tau_z, \quad z = m, n, h.$$

$$\alpha_n(V) = \frac{0.01(V + 55)}{1 - \exp[-0.1(V + 55)]}, \quad \beta_n(V) = 0.125 \exp[-(V + 65)/80],$$

$$\alpha_m(V) = \frac{0.01(V + 40)}{1 - \exp[-0.1(V + 40)]}, \quad \beta_m(V) = 4 \exp[-(V + 65)/18],$$

$$\alpha_h(V) = 0.07 \exp[-(V + 65)/20], \quad \beta_h(V) = \frac{1}{1 + \exp[-0.1(V + 35)]},$$

where $g_{Leak} = 0.3 \text{ mS/cm}^2$, $g_{Na} = 120 \text{ mS/cm}^2$, $g_K = 36 \text{ mS/cm}^2$, $C = 1 \mu\text{F/cm}^2$, $E_K = -100 \text{ mV}$, $E_{Leak} = -70 \text{ mV}$, and $E_{Na} = 50 \text{ mV}$. The surface area is again $A = 100 \mu\text{m}^2$, and $I_0 = 73 \text{ pA}$, which made the mean ISI size $T_0 = 16.8 \text{ ms}$.

The external input was given by $I_{ext}(t) = \sigma\xi$ where ξ is Gaussian white noise with zero mean and unit variance. The $\sigma = 1.6$ and 7 pA , and ξ was updated with 10 and 20 kHz for the ML and HH model, respectively.

All the simulations were done using the NEURON simulation platform (Hines and Carnevale 1997).

Experimental Procedures

All experiments were performed according to animal procedures, which were approved by the Ethical Committee of the University of Antwerp in accordance with the Federal Laws of Belgium and the European Committee. Parasagittal cerebellar slices were prepared from Wistar rats (aged 15 to 22 days old) according to the procedures described in Robberechts et al. (2010). During recording, the slices were held between nylon nets in a submerged chamber on an upright microscope (Leica DM LFS) and continuously perfused at near physiological temperature ($32\text{--}34^\circ\text{C}$) with artificial cerebrospinal fluid at 2–3 ml per minute using a Minipulse peristaltic pump

(Gilson). Synaptic transmission was blocked by addition of 10 μM SR95531 [6-imino-3-(4-methoxyphenyl)-1(6H)-pyridazinebutanoic acid hydrobromide] (gabazine), a GABA_A receptor antagonist; 1 μM strychnine, a glycine receptor antagonist, 10 μM 2,3-dihydroxy-6-nitro-7-sulfamoyl-benzo(f)quinoxaline, a AMPA receptor antagonist; and 50 μM APV, an N-methyl-D-aspartate receptor antagonist, in the artificial cerebral spinal fluid.

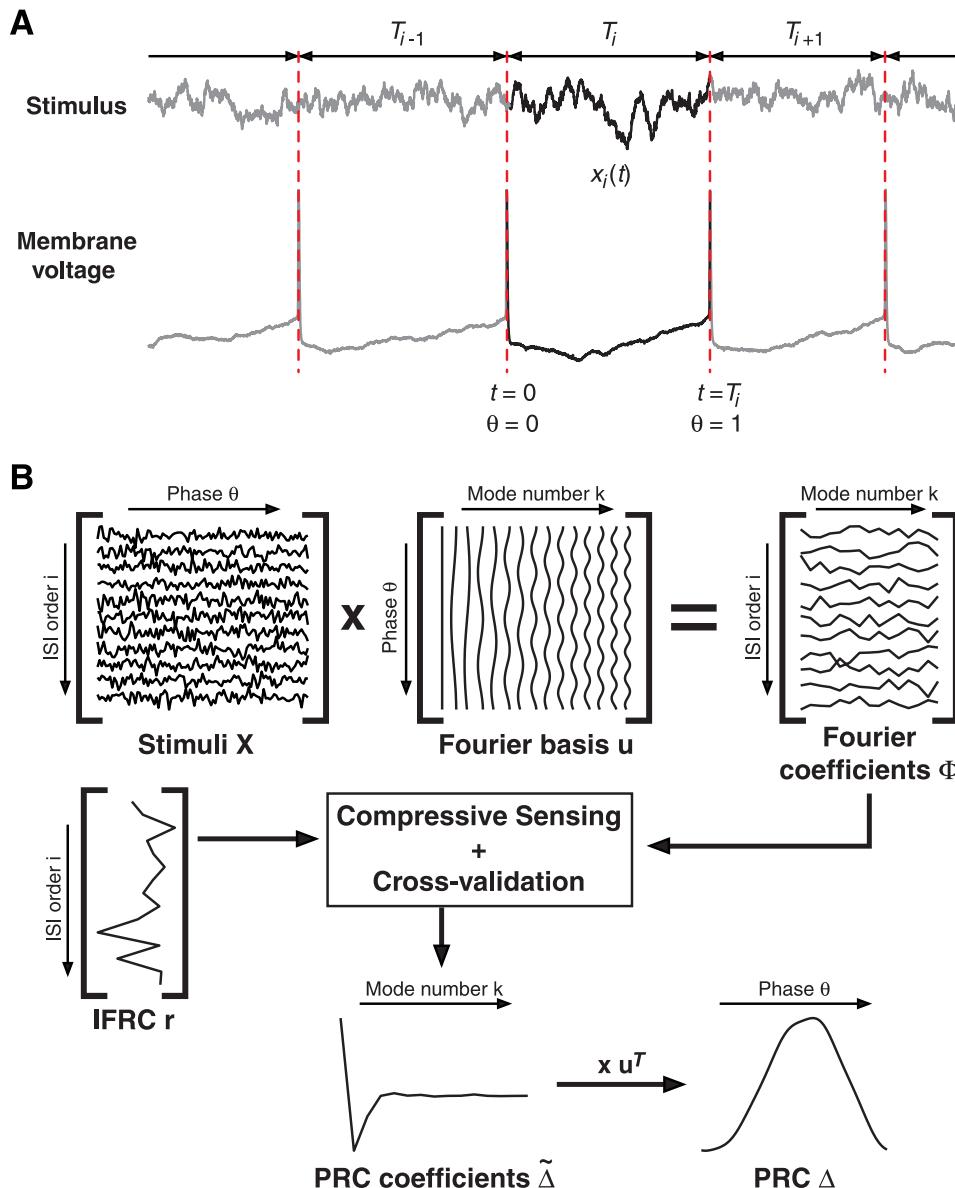
Whole cell patch-clamp recordings were obtained using glass pipettes of 2–3 $M\Omega$ resistance. The recordings were performed with a potassium gluconate-based intracellular solution containing the following (in mM): 135 potassium gluconate, 10 KCl, 10 HEPES, 0.1 EGTA, 4 Mg-ATP, 0.4 Na₃GTP, and 14 Na₂-phosphocreatine, with pH 7.25–7.3 titrated with KOH and supplemented with 0.2% biocytin. The liquid junction potential was 10 mV (Neher 1992) but was not corrected for. The pipette capacitance was compensated for once the gigaseal was obtained, just before entry to the cell, in the cell-attached configuration. Seal resistance was $\geq 2 \text{ G}\Omega$.

Then, whole cell voltage-clamp recordings were performed with a Heka (Lambrecht, Germany) EPC 10 amplifier. Initially the GoCs were voltage clamped at -70 mV to have the cell stabilize for 5–10 min before switching to current clamp. In current clamp, no current was injected and on regular time intervals passive cell parameters like

membrane potential (V_m), input resistance (R_{input}), and recordings were discontinued if significant changes in these parameters occurred. Active cell parameters were also routinely tested like the FI-curve, amount of adaptation, and the reliability of the cell response to frozen noise injection. The continuous noise stimulus was created via a Ornstein-Uhlenbeck process

$$\frac{dx}{dt} = -\frac{x}{\tau} + \xi,$$

where ξ is Gaussian white noise with zero mean and variance σ^2 . The stimulus was injected somatically into the GoCs typically during periods of 40 up to 200 s depending on the spiking frequency. When injecting small currents the noise injection was continued longer to have ≥ 500 spikes for the PRC analysis. In between each noise stimulation protocol, the cell was left quiet for ≥ 2 min. In general, the amplitude of current injection was varied to have 5–10 Hz firing responses and on top of this the variance σ^2 was varied. The time constant $\tau = 4$ ms in the range of the time constants of synaptic currents. Signals were captured at 20 kHz and low-pass filtered at 2.5 kHz.



PRC Estimation Procedures

Estimation procedure. The first step of the estimation procedure was to collect our relevant variables, all the ISI sizes, $\{T_i\}$, and stimulus segments within the corresponding ISIs, $\{\mathbf{x}_i\}$ (Fig. 1, A and B). Then, these variables were transformed to the instantaneous firing rate change (IFRC) $\{r_i\}$, and K coefficients in the Fourier series expansion of the stimuli, arranged in a measurement matrix Φ as (Fig. 1B)

$$(\mathbf{x}_i)_j \approx \Phi_{i0}u_{0,j} + \Phi_{i1}u_{1,j} + \Phi_{i2}u_{2,j} + \dots + \Phi_{i,K-1}u_{K-1,j} = (\Phi\mathbf{u}^T)_{ij}, \quad (1)$$

where

$$u_{0,j} = \sqrt{\frac{1}{L}}, \quad u_{2m-1,j} = \sqrt{\frac{2}{L}} \sin\left(\frac{\pi m j}{L}\right), \quad u_{2m,j} = \sqrt{\frac{2}{L}} \cos\left(\frac{\pi m j}{L}\right), \quad (2)$$

and L is the length of \mathbf{x}_i 's.

First of all, we computed the IFRC for each i as

Fig. 1. Phase-response curve (PRC) estimation from the phase responses of a neuron to the incoming input. A: membrane voltage oscillation (top) of a cerebellar Golgi neuron fluctuating in response to the current input (bottom). Due to the external stimulus $x_i(t)$, the phase θ , ranging from 0 to 1 for each i -th interspike interval (ISI), has a nontrivial relationship with the time parameter t ranging from 0 to T_i , given by Eq. 5. B: simple diagram for our PRC estimation procedure. We construct a stimulus matrix $[\mathbf{x}_1, \mathbf{x}_0, \dots]$ where \mathbf{x}_i is the stimulus during the i -th ISI. Then, we transform it to the Fourier coefficient matrix Φ by multiplication with the Fourier basis matrix. This Φ and instantaneous firing rate changes (IFRC) $r_i = (T_0 - T_i)/T_i$ are plugged into the compressive sensing (CS) algorithms, Eqs. 14 or 15 with cross-validation, and we obtain the Fourier series coefficients of the estimated PRC, which can be converted back to the time (phase) domain representation.

$$r_i = (T_0 - T_i)/T_i,$$

where we chose T_0 as the mean ISI width, $T_0 = \frac{1}{N} \sum_{i=1}^N T_i$ (Ota et al. 2009). In principle, T_0 should be the ISI size without any perturbation. However, in the continuous stimulus paradigm that we used many slow dynamical components can be affected and significantly change membrane dynamics across multiple ISIs. Therefore, using T_0 from the unperturbed data led to consistently and significantly worse performance except for the simulations. In one test case, we also set T_0 as a free parameter and tried to find the optimal T_0 , but the result was not significantly different from using the mean ISI size.

To compute Φ , we resampled \mathbf{x}_i for each i by using the discrete Fourier transformation (DFT) so that all of them have the same length $L = 401$: if the length of \mathbf{x}_i is $L_i > L$, only the first L DFT coefficients were taken, and, in the other case, $L - L_i$ DFT coefficients were set to zero. Then, each resampled \mathbf{x}_i was rescaled by $\sqrt{L_i}/L$. This guaranteed that at least the first L Fourier series coefficients do not change after resampling despite the change in length. Then, these resampled \mathbf{x}_i were arranged into a matrix $X_{ij} = (\mathbf{x}_i)_j$, and finally we obtained the measurement matrix by $\Phi = \mathbf{X}\mathbf{u}^T$. In principle, we need $M = L$ for \mathbf{u} to form a basis that can span all the possible stimuli, but this is not really necessary since we are only interested in the low frequency part, and therefore we used $K = 201$ (401) for the simulated (experimental) data.

Finally, we estimated the Fourier series coefficients of the PRC, $\tilde{\Delta}$, from \mathbf{r} and Φ by using the optimization packages that provide the routines for all the CS schemes that we used. For the basis pursuit (BP) algorithm (*Eq. 14*) that we adapted to the noiseless ML model, we used ℓ_1 -MAGIC (<http://users.ece.gatech.edu/~justin/l1magic/>). For the ML model with background noise, HH model, and Golgi cell, we used the Dantzig selector algorithm (*Eq. 15*) via the ℓ_1 -Homotopy package (Salman and Romberg 2010). The PRC in the time-domain representation could be obtained by $\Delta = \mathbf{u}^T \tilde{\Delta}$. Our estimation code is available for public download at <http://groups.oist.jp/cnu/software>.

Cross-validation. If our CS scheme is the Dantzig selector (*Eq. 15*), the error bound η should be specified. The η controls sparseness of the estimated PRC and consequently also determines how much the estimation overfits to the data noise. Therefore, η should be small enough to fit the given data but also large enough to avoid poor fitness when the estimation is tested with a new data set. To find the optimal η , we employed the k -fold cross-validation method (Hastie et al. 2009): the whole data set of N ISIs were randomly segmented into k blocks each with $[N/k]$ ISIs. We obtained the estimated PRC $\tilde{\Delta}(i, \eta)$ with certain η and a training data set formed by leaving only the i -th block out from the whole set. Then, we obtained the cross-validation error $\chi_i(\eta)$, which is the sum of squared errors (SSE; *Eq. 9*) of the actual IFRC \mathbf{r} and predictions by $\tilde{\Delta}(i, \eta)$ for the i -th block data. After we did this with sufficiently large number of the blocks k' , we could obtain the average cross-validation error and the optimal $\eta = \eta^*$ as

$$\chi(\eta) = \frac{1}{k'} \sum_{i=1}^{k'} \chi_i(\eta), \quad \eta^* = \arg \min_{\eta} \chi(\eta). \quad (3)$$

We used $k = k' = 100$, which yielded a consistent η^* value in each case. In finding the minimum of $\chi(\eta)$, we employed a simple brute force approach that we evaluate $\chi(\eta)$ within an interval $[\eta_{\min}, \eta_{\max}]$ with 40 sample points and selected η with the smallest value of $\chi(\eta)$. The $\eta_{\min, \max}$ were determined by a few test runs for each data set, and the typical values were $\eta_{\min} \approx 5$ and $\eta_{\max} \approx 40$.

Inclusion of uncontrolled noise and CS-total least-squares estimation. To investigate the effect of uncontrolled noisy fluctuation of the ISI, we simulated the ML model in the same way described above except that we injected another Gaussian white noise current with zero mean. The standard deviation was varied from 0 to 1.6 pA by a step of 0.32

pA. When the noise level is very high, the CS estimation showed systematic deviation from the no-noise case. See *Effects of the uncontrolled noise* for our discussion on this point.

To resolve this problem, we augmented our CS-based estimation procedure by a pruning step followed by the total least-squares (TLS) estimation: first, we computed $\tilde{\Delta}$ by using the Dantzig selector with a particular η and found a set of significant modes $S(\eta) = \{j | \tilde{\Delta}_j \text{ is significant}\}$ where their coefficients have significantly larger amplitudes than the rest by maximizing $\min_{j \in S(\eta)} |\tilde{\Delta}_j| - \max_{j \notin S(\eta)} |\tilde{\Delta}_j|$. Then, we constructed the “reduced stimuli” Φ' from Φ_{ij} only for $j \in S(\eta)$ and carried out the TLS estimation by using singular value decomposition (Markovsky and Hufel 2007). The cross-validation was performed in the same way except that $\chi_i(\eta)$ is given by the Rayleigh quotient,

$$E_{RQ} = \frac{\|\mathbf{r} - \Phi \tilde{\Delta}\|_2^2}{1 + \|\tilde{\Delta}\|_2^2} \quad (4)$$

Phoka et al. (2010) recently discussed an interesting phenomenon that the noisy in the spike time leads to a systematic bias in the estimated PRC particularly in the later phase. We could not observe similar bias in our simulated or experimental data and did not try to prevent it. However, in principle, we can address this problem in the continuous stimulus paradigm in a similar way as Phoka et al. (2010) by using corrections from multiple ISIs: we first make the stimulus-IFRC pair from every two consecutive ISI, such as $\{T_i + T_{i+1}\}$ and $\{[\mathbf{x}_i, \mathbf{x}_{i+1}]\}$, and estimate the PRC for these two ISIs as $[\Delta; \Delta]$. In this way, we can eliminate the effect of the spike time uncertainty in the spike between the i -th and $(i+1)$ -th ISI. We can continue this step with three and more consecutive ISIs, and the averages of all estimated PRCs is our final estimation.

Computation of the theoretical PRC. The theoretical PRC can be computed by solving the adjoints of the original dynamical equations (Izhikevich 2010). We computed the theoretical PRCs, in Figs. 2B and 3B, using the XPPAUT software (Ermentrout 2002), which implements a numerical algorithm to calculate the adjoint.

Statistical Analysis

We computed the predictive power of each estimation method by evaluating the coefficient of determination,

$$R^2 = 1 - \frac{\sum_{i=1}^N (r_i - s_i)^2}{\sum_{i=1}^N (r_i - \bar{r})^2}, \quad \bar{r} = \frac{1}{N} \sum_{i=1}^N r_i,$$

where r_i and s_i are the measured and predicted IFRC for each ISI. If s_i is obtained by proper regression ensuring $\bar{s} \sim \bar{r}$, R^2 is simply a squared correlation coefficient between $\{r_i\}$ and $\{s_i\}$. However, if the prediction gets as bad as $\bar{s} \neq \bar{r}$, R^2 can be negative despite its notation.

If the estimated PRC captures most of the linear relationship between stimuli and measured IFRC \mathbf{r}_{meas} , the predictions from the PRC \mathbf{r}_{PRC} should not be significantly correlated with the residuals $\delta\mathbf{r} = \mathbf{r}_{\text{meas}} - \mathbf{r}_{\text{PRC}}$. We evaluated the significance of the correlation ρ between \mathbf{r}_{PRC} and $\delta\mathbf{r}$ by a bootstrap test: we first constructed resampled samples of $\{\mathbf{r}_{\text{PRC}}^{\text{resampled}}, \delta\mathbf{r}^{\text{resampled}}\}_i$ ($i = 1, \dots, 1,000$) and computed the correlations ρ_i . Then, a bootstrap t -score was evaluated by $t = \langle \rho_i \rangle / \sqrt{\text{Var}[\rho_i]}$ and gave the corresponding P value.

All the preprocessing of the data, PRC estimation, cross-validation, and statistical analysis were done in MATLAB (Mathworks, MA).

RESULTS

Estimation of the PRC as a Linear Regression Problem

We consider a neuron receiving a time-dependent stimulus $x(t)$, which makes the ISI T deviate from the original ISI T_0 . The

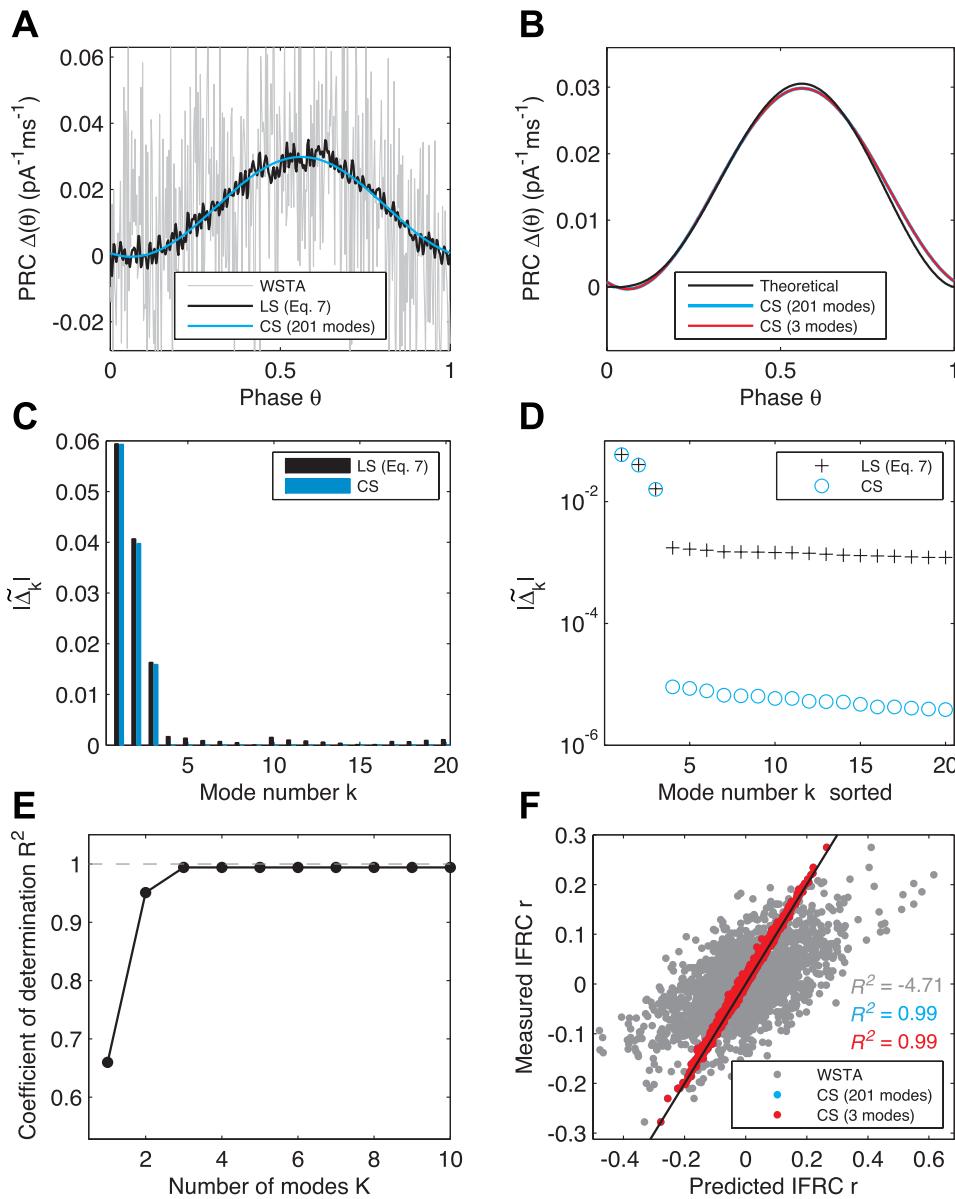


Fig. 2. Estimated PRC of the Morris-Lecar model. *A*: estimated PRC from the simulated data of the Morris-Lecar model via weighted spike-triggered average (WSTA; gray), least square (LS, Eq. 10; black), and CS method (BP, Eq. 14; blue). The number of ISIs is $N = 250$, the mean ISI size is $T_0 = 50.2$ ms, and the coefficient of variation (CV) for the ISI sizes is $CV = 0.073$. Computing the CS estimate took ~ 1.6 s on a desktop computer with a single 2.9 GHz Core 2 Duo processor. *B*: theoretical PRC computed from the model equation (black) is compared with the estimated PRC via CS method (blue). PRC constructed only with the top three PRC coefficients in magnitude is also shown (red; see also *D*). *C* and *D*: sizes of the PRC coefficients $\tilde{\Delta}_k$ in the LS (black) and CS estimate (blue). In *D*, the coefficients are sorted by their magnitudes, $|\tilde{\Delta}_k|$. *E*: number of the included modes K vs. goodness-of-fit, R^2 . We computed R^2 of the measured and predicted IFRCs using a prediction based on the PRC constructed with the top K PRC coefficients in magnitude (as in *D*). R^2 rapidly improves up to $K = 3$ and saturates thereafter, showing that the PRC computed with the three modes (red in *B*) is a sufficiently good estimation. *F*: predicted IFRC, r , of the estimated PRCs vs. measured IFRC in each method. Black line represents the perfect match ($R^2 = 1$). While we use the PRCs from *A* and *B*, the predicted and measured r are computed with a different larger data set with 2000 ISIs to illustrate the predictive power of each estimation. Note that the red dots (CS, 3 modes) almost completely overlap with the blue ones (CS, 201 modes) as the additional modes do not contribute significantly (*E*). See MATERIALS AND METHODS for calculation of R^2 and how to interpret negative R^2 .

oscillation period begins at $t = 0$ and finishes at $t = T$ (Fig. 1A), and the phase of this oscillation $\theta(t)$ should increase monotonically from $\theta(t) = 0$ to $\theta(t) = 1$. Then, in the limit that $x(t)$ is weak, $\theta(t)$ and $x(t)$ are related by the infinitesimal PRC $\Delta(\theta)$ as

$$\frac{d\theta}{dt} = \frac{1}{T_0} + \Delta(\theta)x(t). \quad (5)$$

Our goal is to find out $\Delta(\theta)$ based on our data $\{T_i\}$ and $\{x_i(t)\}$ that we collected from each i -th ISI.

A key step to achieve this is that *Eq. 5* can be approximately recasted into a linear regression problem since we can assume that $x(t)$ and the induced change in $\theta(t)$ are sufficiently small so that we can take a linear approximation of *Eq. 5* in which the higher order fluctuations in $\Delta(\theta)$ due to the first order perturbations in θ are ignored. In practice, we use discretized representations such as $x_i(t) \rightarrow \mathbf{x}_i = [x_i(0), x_i(\delta\tau), x_i(2\delta\tau), \dots]$ and similarly $\Delta(\theta) \rightarrow \Delta$. Then, when the IFRC for each ISI is $r_i = (T_0 - T_i)/T_i$, we can show that linear approximation leads to (see APPENDIX for derivation)

$$\mathbf{r} = \mathbf{X}\Delta, \quad X_{ij} = (\mathbf{x}_i)_j. \quad (6)$$

Note that the IFRC is approximately proportional to the ISI fluctuation as $r_i = (T_0 - T_i)/T_i \approx (T_0 - T_i)/T_0 + O[(T_0 - T_i)^2]$ (Achuthan et al. 2011).

The second key observation is that we can take alternative representations for Δ and \mathbf{x}_i in *Eq. 6* by a simple basis change. In this study, we will focus on representing the PRC in terms of the finite Fourier series,

$$\Delta(\theta) = \tilde{\Delta}_0 + \tilde{\Delta}_1 \sqrt{2} \sin(\pi\theta) + \tilde{\Delta}_2 \sqrt{2} \cos(\pi\theta) + \dots + \tilde{\Delta}_{K-1} \sqrt{2} \sin(K\pi\theta/2) + \tilde{\Delta}_K \sqrt{2} \cos(K\pi\theta/2). \quad (7)$$

From *Eq. 6*, the equation for $\tilde{\Delta}$ is simply

$$\mathbf{r} = \Phi \tilde{\Delta}, \quad \Phi = \mathbf{X} \mathbf{u}^T. \quad (8)$$

where \mathbf{u} is a matrix composed of Fourier modes given by *Eq. 2*. Here Δ and $\tilde{\Delta}$ essentially refer to the same object, but, to

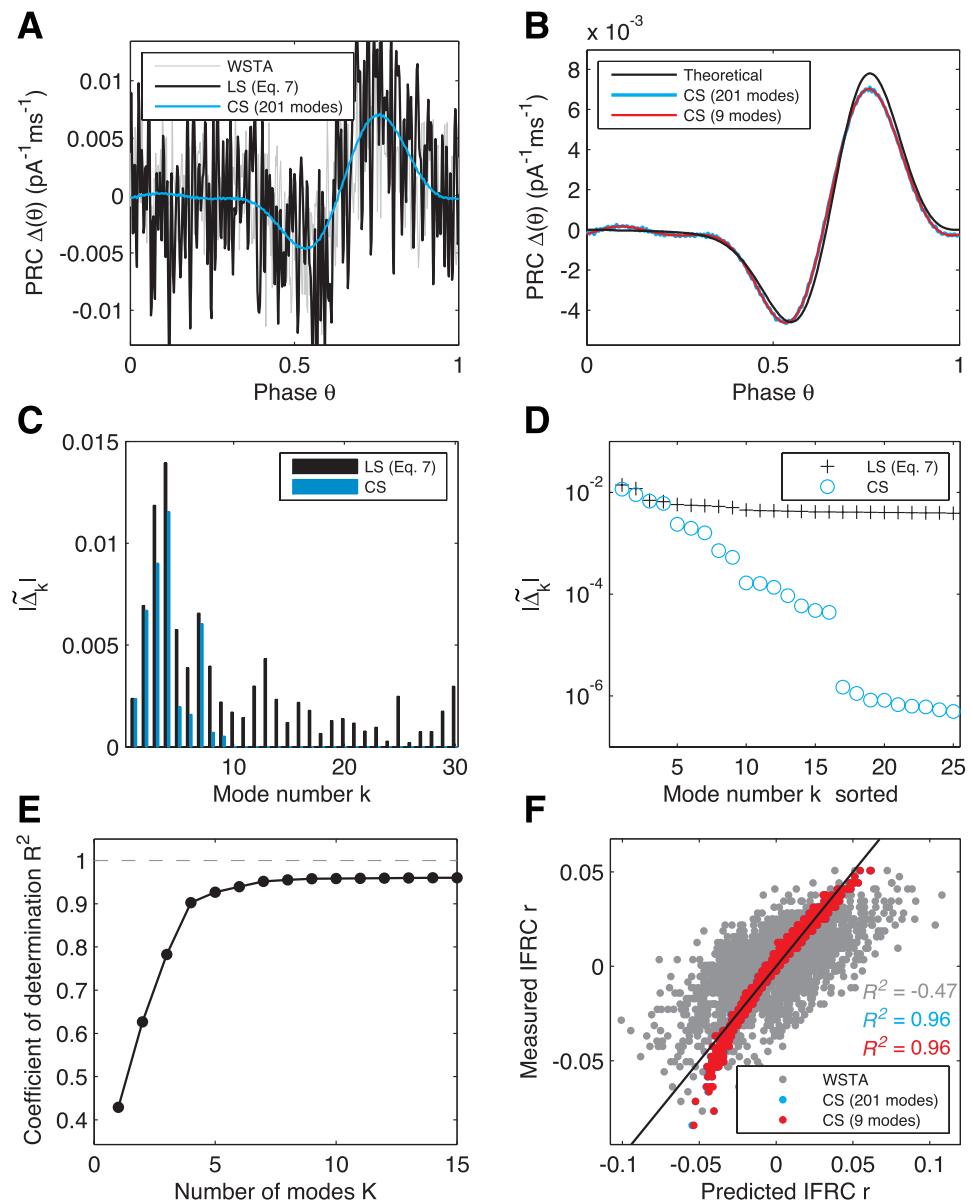


Fig. 3. Estimated PRC of the Hodgkin-Huxley model. *A*: estimated PRC from the simulated data of the Hodgkin-Huxley model via WSTA (gray), LS (black), and CS method (Dantzig selector, Eq. 15; blue). $N = 200$, $T_0 = 16.8$ ms, and $\text{CV} = 0.019$. Computing the CS estimate including the cross-validation step (see MATERIALS AND METHODS) took 60.5 s on the same hardware as Fig. 2. *B*: theoretical PRC computed from the model equation (black) is compared with the estimated PRC via CS method (blue). Estimated PRC constructed only with the top nine PRC coefficients in magnitude is also shown (red). Small deviation is due to small data size and vanishes as the data size increases (see Fig. 4*B*). *C* and *D*: sizes of the PRC coefficients $\tilde{\Delta}_k$ in the LS (black) and CS estimate (blue). In *D*, the coefficients are sorted by their magnitudes, $|\tilde{\Delta}_k|$. Compared with the ML case (Fig. 2, *C* and *D*), the LS estimate has many large PRC coefficients and does not lead to a sparsely constructed PRC while only ~ 9 modes make dominant contribution to the CS estimate. *E*: Number of the included modes K vs. goodness-of-fit, R^2 . As expected from *B-D*, R^2 almost saturates around $K = 9$. *F*: predicted IFRC of the estimated PRCs vs. measured IFRC in each method. The black line again represents the perfect match ($R^2 = 1$). Predicted and measured r are computed with a larger data set with 2,000 ISIs. Again, the red dots (CS, 9 modes) almost completely overlap with the blue ones (CS, 201 modes) as expected from *E*.

prevent confusion, we will call $\tilde{\Delta}$ a PRC coefficient vector or PRC coefficients while PRC will refer only to its time-domain representation Δ .

Relation to some existing methods. A direct approach for solving a problem like Eq. 8 is to minimize the sum of squared errors (SSE),

$$E_{\ell_2} = \|\mathbf{r} - \Phi \tilde{\Delta}\|_2^2 = \sum_i [r_i - (\Phi \tilde{\Delta})_i]^2, \quad (9)$$

and a well-known solution for this least square (LS) problem is (Bishop 2007)

$$\tilde{\Delta} = (\Phi^T \Phi)^{-1} \Phi^T \mathbf{r}. \quad (10)$$

Note that $\Phi^T \mathbf{r}$ is the sum of the stimuli each weighted by r_i and $\Phi^T \Phi$ corresponds to the stimulus-stimulus correlation across the ISIs. This means that the best fitting estimate of the PRC coefficients are always expressed in the form of the weighted average of the stimuli normalized by the stimulus correlation factor.

Therefore, if there is no prior inter-ISI stimulus correlation and the number of the ISIs N is so large that chance correlations (due to random fluctuations) are suppressed, $\Phi^T \Phi$ provides trivial normalization and the PRC estimation can be expressed in a simple form. For example, let's consider PRC estimation in the time domain (i.e., $\Phi = \mathbf{X}$ and $\tilde{\Delta} = \Delta$) when the stimuli \mathbf{x}_i are Gaussian white noise with zero mean and variance σ^2 . When the number of ISIs N is very large, we find $\Phi^T \Phi \approx N\sigma^2 \mathbf{1}$. In this limit, the LS estimate of Δ via Eq. 10 is

$$\Delta = \frac{1}{N\sigma^2} \sum_{i=1}^N r_i \mathbf{x}_i. \quad (11)$$

This is the weighted spike-triggered average (WSTA) estimate of the PRC, found in Ota et al. (2009). We conclude that the WSTA estimate is one special example of the general LS estimate, Eq. 10.

In another example, Galán et al. (2005) proposed a method based on pulse stimuli and the Fourier series representation as

in Eq. 7. Here the pulse stimuli are delivered randomly, say at τ_i for each ISI, and the ISI size T_i is predicted from a candidate PRC, which is presumably a linear combination of only a few Fourier basis vectors with low frequencies. The PRC coefficients are then estimated by minimizing the sum of the squared errors between the predicted and measured T_i . This method is effectively equivalent to minimizing E_{ℓ_2} when $\Phi = \mathbf{X}\mathbf{u}^T$ is constructed with the pulse stimuli $x_i(\tau) = \sigma\delta(\tau - \tau_i)$, and the selected Fourier mode vectors $\{\mathbf{u}_k\}$ ($k = 1, \dots, K$) (from Eq. 2). Then, when N is large and τ_i is uniformly distributed within each ISI, we find $\Phi^T\Phi \approx N\rho^2\sigma^2\mathbf{1}$, where ρ is the pulse density within a time bin $[\tau, \tau + \delta\tau]$, and therefore

$$\tilde{\Delta}_k = \frac{1}{N\rho^2\sigma^2} \sum_{i=1}^N r_i u_{k,j_i}, \quad (12)$$

where j_i is the index corresponding to τ_i , $j_i = [\tau_i/\delta\tau]$. Again, this is nothing but the weighted average of Fourier mode evaluated at the pulse times.

Izhikevich (2010) proposed using a similar scheme as the Galán et al. method for continuously fluctuating noise stimulus. This method is equivalent to minimizing E_{ℓ_2} , where Φ is constructed with noisy stimuli and K randomly selected Fourier basis vectors. A least square optimization in this case seems to converge with much more difficulty than the Galán et al. method (Torben-Nielsen et al. 2010), but nevertheless a direct solution can be obtained via Eq. 10. When the stimuli are Gaussian white noise with zero mean and variance σ^2 and N is large, we again find $\Phi^T\Phi \approx N\sigma^2\mathbf{1}$. This leads us to

$$\tilde{\Delta}_k = \frac{1}{\sigma^2 N} \sum_{i=1}^N r_i \mathbf{x}_i \cdot \mathbf{u}_k. \quad (13)$$

Comparing this to Eq. 11, we can see that the result of the Izhikevich method is nothing but a WSTA projected onto a space spanned by the Fourier mode vectors $\{\mathbf{u}_k\}$. Therefore, the Izhikevich method is simply equivalent to the WSTA estimate in this limit.

In the comparison study by Torben-Nielsen et al. (2010), these methods were classified according to whether the PRC is directly evaluated, such as WSTA, or inferred via minimizing the SSE, such as the Galán et al. method. However, we showed that this distinction is not necessary since SSE minimization can be solved directly by the “weighted average” solution, Eq. 10. Therefore, the direct estimation and optimization methods naturally share most of their pros and cons.

Instead, a more important difference between the direct estimation and optimization methods is whether the PRC is densely or sparsely constructed: for example, in the WSTA method, Δ_k for every time step $k = 1, 2, \dots, L$ is calculated from the data by averaging. Conversely, the Galán et al. method assumes that the PRC can be represented by a Fourier series with $K \ll L$. Therefore, this PRC contains only a small number of nonzero PRC coefficients and is sparsely constructed. As we discussed in the Introduction, this is a reasonable assumption considering that the time scale/frequency information of $\tilde{\Delta}$ is concentrated in a much lower range compared with the wide spectrum of Φ .

Searching specifically for a sparsely constructed PRC is important because it resolves many issues in estimation due to the small number of samples (ISIs) N . The first problem is

feasibility; if data size N is small, the data cannot give us the full information necessary to estimate the PRC correctly. For example, if $N < L$, the number of unknowns exceeds the number of equations and the estimation is simply infeasible. However, even if $N \geq L$, the large number of parameters can introduce artifacts in the estimated PRC. In the LS solution, Eq. 10, has an additional normalization by the stimulus autocorrelation $\Phi^T\Phi$ and, as we have seen above, this is usually assumed to be trivial particularly in the limit when N is very large and the estimated PRC becomes a weighted average of stimuli. However, if an autocorrelation is given, the normalization would be already nontrivial even with the large N , and furthermore, in the case when N is small, $\Phi^T\Phi$ contains many (growing as $\sim L^2$) nonzero stimulus correlations due to random fluctuations (chance correlations). In both cases, the simple normalization scheme fails and this causes artifacts in the PRC, such as a wrong amplitude.

This problem has been partially addressed by using the total stimulus autocorrelation, not stimulus variance, for normalizing the WSTA estimate (Ota et al. 2009). However, this ignores the off-diagonal correlations and can distort the estimated PRC as we will see in the examples. A more careful approach is to explicitly deconvolve the full stimulus correlation out from the candidate PRC (Achuthan et al. 2011; Preyer and Butera 2005) as our LS solution does by the factor of $(\Phi^T\Phi)^{-1}$. However, particularly when the data are small, there can be significant off-diagonal chance correlations, and trying to deconvolve these out can mostly contribute to fitting the PRC to the data noise rather than to the data signal, which is a general problem in any regression with a large number of parameters and small data size (Bishop 2007; Hastie et al. 2009).

Sparse estimation relies on a much smaller number of variables ($K \ll L$) and may not suffer from the same problems. However, in this case, we have a new challenge because we must determine how many and which basis vectors are to be used. In the Galán et al. and Izhikevich methods, this decision is somewhat arbitrarily based on random truncation or selection of the modes. In this study, we describe a much more systematic approach to solve this problem.

CS Method

Once we assume the sparseness of the PRC, we note that its estimation can be translated into a problem in signal processing, called sparse signal recovery. Sparse signal recovery aims to reconstruct an original signal from the data made by a smaller number of measurements, possibly contaminated with noise, by taking advantage of the prior knowledge that the signal itself can be constructed by a smaller number of basis vectors than the apparent dimension of the signal. This problem has exactly the same form as our PRC estimation equation Eq. 8 where $\tilde{\Delta}$ is the signal to be recovered, Φ is the measurement matrix, and the result of measurement is \mathbf{r} . Therefore, algorithms that can solve this problem can be also used for PRC estimation.

One strategy for the sparse signal recovery is to test how well each basis vector can predict \mathbf{r} from Φ one by one and exclude those whose contributions are insignificant (below a threshold) from our set of basis vectors. However, this matching pursuit (MP) strategy (Mallat and Zhang 1993) is known to be compu-

tationally demanding and inefficient (Natarajan 1995). A much more efficient alternative is BP (Chen et al. 1998)

$$\min \|\tilde{\Delta}\|_1 \text{ subject to } \mathbf{r} = \Phi \tilde{\Delta}, \quad (14)$$

where $\|\mathbf{x}\|_1 = \sum_i |x_i|$. One interpretation is that the BP provides “regularization,” such that the data can be explained with minimal number of parameters as minimizing $\|\tilde{\Delta}\|_1$ will make many $\tilde{\Delta}_i$ zero (Bishop 2007; Hastie et al. 2009). Recently, it has been proven that, if some conditions such as sparseness of the signal are met, the BP algorithm can achieve exact signal recovery even when the number of samples is as small as $N \sim K \log L$, and this strategy has been named CS (Candès and Tao 2006; Donoho 2006).

CS is currently an actively developing field and many interesting variants of the original breakthrough have been proposed (Baraniuk 2007). In this study, we focus on two basic CS algorithms: in addition to the exact BP strategy (*Eq. 14*), we also used the Dantzig selector (Candès and Tao 2007),

$$\min \|\tilde{\Delta}\|_1 \text{ subject to } \|\Phi^T(\mathbf{r} - \Phi \tilde{\Delta})\|_\infty \leq \eta, \quad (15)$$

where $\|\mathbf{x}\|_\infty = \max |x_i|$. In the case of PRC estimation, this algorithm tries to minimize the distance between the WSTA and the estimated PRC convolved with the stimulus correlation but also minimizes the number of nonzero PRC coefficients at the same time. Compared with the exact BP algorithm, the Dantzig selector performs better when the data are contaminated with Gaussian noise. In such cases, the exact BP algorithm, *a priori* assuming that the recovered signal predicts the measurement results exactly, may fail to converge. Furthermore, the Dantzig selector also applies to cases where the signal is approximately sparse such that many coefficients actually remain nonzero, but their magnitudes decay at least faster than a power law, $|\tilde{\Delta}_k| \sim 1/k^n$ for a positive integer n . This is important since a PRC can have this property: for example, it has been reported that cerebellar Purkinje neuron can have a square wave-like PRC, indicating neuronal near-perfect integration (Phoka et al. 2010), and in this case many Fourier coefficients will be nonzero but decay as $|\tilde{\Delta}_k| \sim 1/k$.

In our case, the exact BP successfully converged in only a few examples of the simulated neurons, and we used the Dantzig selector for the rest. We also tried another algorithm for noisy data called BP denoising (Candès et al. 2006b) for those examples, but we could not find any significant improvements over the Dantzig selector (data not shown).

Fig. 4. Comparison of the estimated PRCs with very large data. Same as in Fig. 2, A and B (A), and 3, A and B (B) with the larger data sets. Number of ISIs are $N = 71,200$ (A) and $N = 26,000$ (B), respectively. WSTA, LS, CS estimate, and theoretical PRC converge as the data size grows.

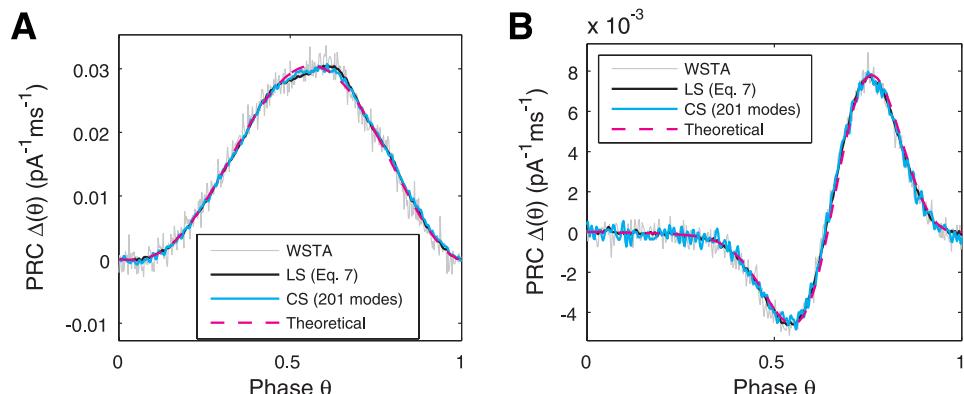
Not any measurement Φ can be used for CS recovery, and the precise conditions for good signal recovery seem rather complex (Donoho and Tanner 2005a,b). However, surprisingly, it has been proven that Φ satisfies the sufficiency condition if it is given randomly such as by Gaussian noise (Baraniuk et al. 2008; Candès and Tao 2005, 2006). This can be achieved by using continuous Gaussian noise stimuli, and therefore we restrict our study to this case.

Applications to Neuronal PRCs

Simulated data: ML and HH neurons. We first show two examples of estimating the PRCs from the simulated data of two different model neurons, each representing a distinct type of neural excitability (Hodgkin 1948; Izhikevich 2010). The first one, the ML model (Morris and Lecar 1981) has type I excitability, and the other model, the HH model (Hodgkin and Huxley 1952) is type II. Type I and II neurons undergo saddle-node or Hopf bifurcations, respectively, and this difference is reflected in the shape of their PRCs. A type I neuron has a positive and monophasic PRC such that a depolarizing stimulus will always increase the instantaneous firing rate, while the PRC of the type II neuron is biphasic, characterized by a significant negative part (Ermentrout 1996). We simulated these models with continuously fluctuating current stimuli, collected the spiking data, and estimated the PRC. For basis vectors, we used the Fourier mode vectors from *Eq. 2*.

Figure 2 shows the results from the simulated data of the ML model with 300 ISIs. In Fig. 2A, we can see that the WSTA estimate has a large signal-to-noise ratio, which is even worse than the LS estimation via *Eq. 10*. As we already discussed, these two estimations are essentially the same except that, in the WSTA, the normalization in *Eq. 10* is replaced by the large N approximation $\Phi^T \Phi \approx N \sigma^2 \mathbf{I}$, which is causing a problem. On the other hand, the CS estimate via the BP algorithm (*Eq. 14*) shows very good signal-to-noise ratio. For comparison, these three estimations indeed get close to each other when the data becomes much larger (Fig. 4A).

If the stimuli are assumed to be infinitesimally small, the PRC can be exactly computed from the dynamical equations of the system (Ermentrout 2002; Izhikevich 2010). This theoretical PRC matches well not only with the full CS estimate but also the PRC with only the three largest PRC coefficients (Fig. 2B). Indeed, both in the LS and CS estimate, most power is in those three modes (Fig. 2, C and D), which seem to be sufficient to predict the actual IFRC with good accuracy (Fig. 2, E and F). This confirms that our CS-based estimation



method, based on the sparseness assumption of the PRC, is an effective strategy.

In the HH case where the data consisted of only 200 ISIs, we obtained more interesting results. Here the signal-to-noise ratio of the LS estimate is only slightly better than the WSTA estimate while the CS method (via the Dantzig selector) can estimate a well-denoised PRC (Fig. 3A). Again, not only the full estimate but also the PRC only with nine largest coefficients, matches well with the theoretical PRC (Fig. 3B). Again, small differences among the estimates and theoretical PRC vanish at very large data sizes (Fig. 4B).

As we expect from these, the LS estimate is not sparse: while most of the power of the CS estimate is concentrated in the first $\sim 8\text{--}9$ modes, the LS estimate has many other coefficients as large as those (Fig. 3, C and D). Figure 3, E and F, also shows that the CS estimates, even with a few PRC coefficients, can be a good predictor for the actual IFRCs in the data. Also note that R^2 still keeps improving significantly even beyond $K = 4$ (Fig. 2E), while $K = 3\text{--}5$ is often used when K is chosen by hand (for example, Torben-Nielsen et al. 2010).

In all the cases, the LS estimate should be the best predictor ($R^2 = 1$) for the particular data that the estimation is based on. However, this is simply due to fitting to the data noise as we discussed: with a different data set with 2,000 ISIs as in Figs. 2F and 3F, the LS estimates actually perform worse ($R^2 = -1.40$ in the HH case, in the ML case $R^2 = 0.97$; see MATERIALS AND METHODS for when R^2 becomes negative and how we interpret it). On the other hand, the CS estimate can avoid this overfitting due to the imposed sparseness constraint.

Effects of the uncontrolled noise. Real neurons are almost always under the influence of many noise sources such as channel stochasticity, background synaptic activity, etc., and therefore it is important to characterize how our PRC estimation performs at the different noise levels.

For this purpose, we simulated the ML model with additional uncontrolled noise in the input and estimated the PRC. Figure 5A shows how the estimated PRC deviates from the theoretical PRC depending on the noise level and number of ISIs. The CS method already performs much better compared with the WSTA estimation method (dotted line), but we often found that the estimated PRC has a significantly and system-

atically reduced amplitude compared with the theoretical one (Fig. 5B), particularly when the noise level is high. This seems to be due to the fact that the algorithm minimizes $\|\tilde{\Delta}\|_1$ and therefore, in the high noise regime, it also tries to minimize nonzero and significantly contributing PRC coefficients.

To solve this problem, we added a “pruning” step after the CS estimation where we identified which Fourier modes make significant contributions and reestimated the PRC by using only those modes without minimizing $\|\tilde{\Delta}\|_1$. For reestimation, we can simply use the LS estimation, but instead we used the TLS method (Markovsky and Huffel 2007), which minimizes not the SSE but the Rayleigh quotient (RQ),

$$E_{RQ} = \frac{\|\mathbf{r} - \Phi\tilde{\Delta}\|}{1 + \|\tilde{\Delta}\|_2^2},$$

(see MATERIALS AND METHODS for more details). Our motivation for this is that our noisy data is formed by injecting additional input noise that models intrinsic noises due to background synaptic inputs, etc. In our PRC equation, this can be described as

$$\mathbf{r} = (\Phi + \tilde{\mathbf{n}}_{\text{input}})\tilde{\Delta}, \quad \tilde{\mathbf{n}}_{\text{input}} = \mathbf{n}_{\text{input}}\mathbf{u}^T, \quad (16)$$

where $\mathbf{n}_{\text{input}}$ represents the uncontrolled input noise, and this is a typical problem where the TLS method applies. Figure 5C shows that our CS-TLS hybrid method gives us much better estimations, without the amplitude problems and the deviations diminish nicely with larger number of ISIs even at higher noise levels. In general, we observed that this method also can correct well the systematic deviation in the CS estimation for cases where the noise is included in different ways, such as simply adding noise to the ISIs (data not shown).

Note that, for example in Fig. 5B, it would be very difficult to figure out which estimation is better based on the measures such as SSE, without knowing the theoretical PRC or actual noise level: $R^2 \approx 0.5$ in both cases, and actually it would not make sense to look for an estimation with higher R^2 since the amplitudes of the input noise probe stimuli are the same in this case. Both in the CS and CS-TLS, what prevented those

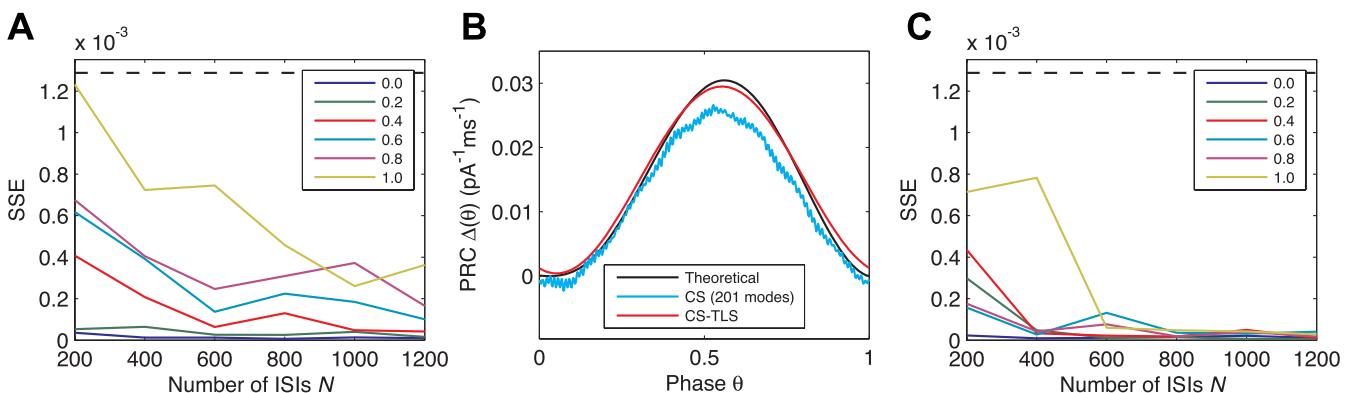


Fig. 5. Influence of uncontrolled noise in PRC estimation. A: sum of squared error (SSE) between the estimated and theoretical PRC, $\|\Delta_{\text{estimated}} - \Delta_{\text{theoretical}}\|_2^2$. The dotted black line represents the WSTA with no uncontrolled noise and $N = 5,000$ for comparison. Each line represents a different noise level, $\sigma_{\text{noise}}/\sigma_{\text{stim}}$. B: examples of estimated PRCs via the CS-based and CS-TLS method. The simulation ran with exactly the same condition as Fig. 2, except that we added Gaussian noise on top of the stimuli with the same magnitude ($\sigma_{\text{noise}} = 1.6 \text{ pA}$), which resulted in $CV = 0.1$ and $T_0 = 50.4 \text{ ms}$. Number of ISIs was $N = 800$ in both cases. We can see that the CS estimation has a significantly smaller amplitude than the theoretical PRC, even though the SSE from the theoretical PRC is much smaller than WSTA estimation with many more ISIs (see A). On the other hand, the CS-TLS estimation is free of this problem. C: same as A but using the CS-TLS method. Estimated PRC is much more similar to the theoretical PRC, particularly as N grows.

overfitting estimations was cross-validation, which emphasizes the importance of this step. Also interestingly, we empirically observed that the cross-validation in the CS case can behave relatively worse at the higher noise level and pick out a nonsparsified PRC estimation. On the other hand, the CS-TLS case consistently preferred the sparse estimations. However, the CS-TLS cross-validation was sometimes more unstable and we had to repeat the step several times with differently shuffled data. One possible reason is that, during the CS-TLS estimation, if a certain $\tilde{\Delta}_i$ becomes small after the CS step but still passes our significance criterion, $\tilde{\Delta}_i$ can grow back after the pruning step and contribute to generating an overfitting estimation. Therefore, future improvements can be made by putting a more careful criterion about whether a particular $\tilde{\Delta}_i$ is significant or not.

Experimental data: cerebellar Golgi neuron. We also tested our CS method with experimental data obtained from patch-clamp recordings of GoCs stimulated with a Gaussian noise current filtered with time constant $\tau = 4$ ms.

Figure 6A shows the estimated PRCs by our CS method (via the Dantzig selector) from the data sets recorded from one cell with different standard deviations σ of the stimulus. The estimated PRC can vary with σ due to many factors influencing neuronal dynamics and the limitation that the PRC is a linear approximation (Ermentrout et al. 2007; Torben-Nielsen et al. 2010). However, such differences are relatively small in Fig. 6A: the three estimated PRCs largely overlap with each other, showing consistent triangular shapes. Interestingly, PRCs of GoCs with a similar shape have been observed in a different context, where stimuli were delivered via gap junctions and firing of another coupled GoC (Vervaeke et al. 2010).

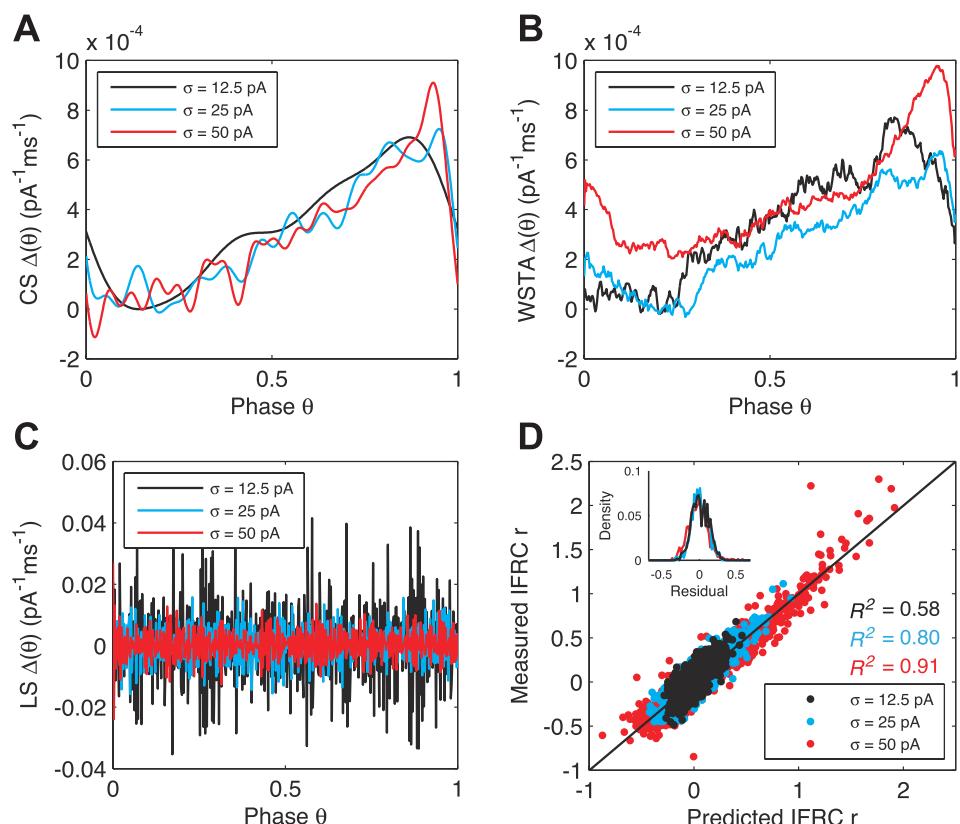
On the other hand, each WSTA estimate shows a very different shape from the others (Fig. 6B). Here the amplitudes of the PRCs are normalized by the total stimulus autocorrelation (Ota et al. 2009) because normalization by the stimulus variance, as Eq. 11, makes the PRCs larger by several orders of magnitude (up to $0.04 \text{ pA}^{-1} \cdot \text{ms}^{-1}$). Using the total autocorrelation for normalization is equivalent to an approximation $\Phi^T \Phi \approx N\sigma^2 \mathbf{1}$ where σ^2 is the average of the total autocorrelations in $\Phi^T \Phi$ matrix, i.e., $\sigma^2 = L^{-1} \sum_{i,j} (\Phi^T \Phi)_{ij}$. Therefore, this scheme fixes the issue with the amplitude of the PRC but still ignores the effects of the off-diagonal terms in $\Phi^T \Phi$, which leads to the artifacts in the shape.

On the other hand, the LS estimates suffer from a completely different problem: here the measured IFRCs can be perfectly matched with the predictions ($R^2 = 1$ in all cases), but the estimated PRCs are completely noisy and do not seem to reflect any intrinsic dynamics of the neuron at all (Fig. 6C). Again, this indicates that the LS estimate is fitting to the data noise.

The goodness-of-fits for the CS estimates are in a good range of $R^2 = 0.58–0.91$ (Fig. 6D) and the residuals (= [measured IFRC] – [predicted IFRC]) are stable for different stimulus σ (Fig. 6D, inset). In comparison, the goodness-of-fits for the WSTA estimates seem to be only slightly worse ($R^2 = 0.58–0.78$). However, a closer look reveals that the residuals from the WSTA estimates are significantly correlated with the predicted IFRC ($P < 0.001$, bootstrap test) except for $\sigma = 12.5$ pA, while those from the CS estimates are not. This indicates that the WSTA estimate can miss a significant part of the linear relationship between the IFRCs and stimuli.

The residuals for the CS estimates, which are almost invariant to the stimulus variance (Fig. 6D, inset), suggest that the

Fig. 6. Estimated PRC of the cerebellar Golgi neuron (GoC). **A:** PRCs estimated by the CS method (Dantzig selector, Eq. 15) from the experimental data obtained by injecting a noisy current stimulus into a GoC in vitro. Same cell was injected with stimuli with three different noise amplitudes $\sigma = 12.5, 25$, and 50 pA. $N = 714, 757, 842$, $T_0 = 91, 86, 77$ ms, and $CV = 0.18, 0.24$, and 0.46. Each estimation including the cross-validation step took about 184.6, 166.6, and 276.2 s, respectively. **B:** WSTA estimates from the same data as A. Shapes of estimated PRCs are unphysiologically distinct from each other for different σ . **C:** LS estimates from the same data as A. $R^2 = 1$ in all cases but the shapes of the estimations are almost pure noise since the estimates are now overfitting the data. **D:** predicted IFRC via the PRCs in A vs. measured IFRC. **Inset:** (density) histogram of the residuals, which are almost invariant with respect to σ .



uncaptured variance of the actual IFRC ($= 1 - R^2$) is mostly affected by intrinsic irregularity. Because the GoC fires somewhat irregularly due to noise in its excitability (Forti et al. 2006), the PRC cannot predict this part. As the stimulus amplitude increases, we force the cell more and more to fire coupled to the stimulus and the prediction can become more accurate (R^2 increases) until eventually the assumption of a linear perturbation stops holding.

In summary, these results demonstrate that our CS-based estimation strategy can indeed extract the biophysically relevant information from experimental data while avoiding dangerous overfitting.

DISCUSSION

In this study, we have introduced a novel and efficient method to estimate the PRC based on CS. CS is a recent development in signal processing theory that proposes strategies to recover a signal from highly undersampled data. We showed the similarity between the signal recovery problem and the estimation of the PRC from a limited amount of experimental data and illustrated how we can make use of the CS framework for PRC estimation. We finally demonstrated that the CS method can successfully extract the PRC from small sized data samples using data from simulated neurons and experimental data recorded *in vitro*.

One of the reasons that our CS method worked well is that it deals with the problems arising for a regression analysis with finite data in a correct way. To illustrate this, we discussed in detail how those issues emerge in PRC estimation, which is formulated as a linear regression problem. In particular, we focused on two problems: applying an approximation that is only good for very large data to a small data set and overfitting noise in the data. As we observed in the examples, those factors caused artifacts by entirely ignoring (i.e., only true at infinitely large data) or fully considering (i.e., overfitting to all the data noise) stimulus correlations, particularly chance correlations of data noise. The CS algorithms, designed to recover a signal from undersampled data, avoid these pitfalls and generate PRC estimations that are little affected by chance correlations. In fact, these problems are well known in statistical learning theory (Bishop 2007; Hastie et al. 2009) but have rarely been considered in PRC estimation except for a recent Bayesian estimation approach (Nakae et al. 2010).

The CS method is also efficient: it generated the estimation in each of our simulated and experimental data sets, including cross-validation, within at most a few minutes in the Matlab environment (Mathworks, MA) running on a desktop computer with a single 2.9 GHz Core 2 Duo processor (Apple, CA). Previously, the Izhikevich method (Izhikevich 2010) with continuous stimuli data has been observed to converge very slowly (Torben-Nielsen et al. 2010) because the relevant Fourier modes are determined by random testing whether each mode improves the fitness. This type of algorithm is a variant of matching pursuit (Mallat and Zhang 1993), which is computationally expensive (Natarajan 1995). Therefore, it is not surprising that the CS method, based on a much more efficient BP strategy, performs faster. Similarly, the Bayesian estimation algorithm also takes several hours to run on a cluster computer (Nakae et al. 2010). The most time-consuming part of our computation was the cross-validation, which used a grid-based simple brute force approach, repeated 4,000 times,

to find the optimal sparseness parameter. Using better optimization algorithms can definitely improve efficiency at this stage by reducing the number of repeated evaluations.

On the other hand, the CS method puts constraints on the probe stimulus. The nature of a boundary sharply separating good and bad signal recovery is complicated (Donoho and Tanner 2005a,b), but there exist proven sufficiency criteria (Baraniuk et al. 2008; Candès and Tao 2005; 2006), which are satisfied by the Gaussian random measurements. This corresponds to the continuous Gaussian noise current stimulus, which has been suggested to mimic integrated synaptic current *in vivo* (Mainen and Sejnowski 1995), and this is the scheme that we used in this study. Another popular choice for the probe stimuli is a set of pulses (for example, see Ermentrout et al. 2001; Galán et al. 2005; Gutkin et al. 2005; Phoka et al. 2010; Stiefel et al. 2008). Since the pulses are sharply localized in time while we estimate the PRC in the Fourier domain, this corresponds to one of the earliest examples of CS where the so-called robust uncertainty principle guarantees good signal recovery (Candès et al. 2006a). Therefore, two popular choices for the probe stimulus are in principle well-suited for applying the CS method.

However, in practice there are caveats. First of all, the stimuli often have temporal correlations to mimic those of synaptic inputs. In our case, the Gaussian noise stimuli for the GoC experiments had a relaxation time of $\tau = 4$ ms. A similar scheme has been used for the pulse inputs, too (Achuthan et al. 2011; Preyer and Butera 2005). In general, such correlations reduce the effective amount of independent information from the whole data set. As we have discussed, the CS method is able to deal with this to some degree since it looks for a sparse solution and therefore circumvents the problem of undersampling, but still coherence can be one of the main factors that limit good estimation (Candès and Romberg 2006). If significant effects of correlation are present, methods specialized for such cases should be used (Zhang and Rao 2011).

Secondly, we used the linear approximation of the phase-response equation, *Eq. 5*, which already linearly approximates the full phase response of the neuron. Therefore, in practice, all the nonlinearities that we have ignored can affect the estimations in many different ways, particularly depending on what kind of stimulus protocol is used. In the pulse stimulus protocol, the unperturbed period can be relatively well defined while there still can be effects of the intrinsic background noise (Phoka et al. 2010), and the effects of the stimulus on top of that can be better isolated. However, in the continuous stimulus protocol, the neuron is continuously driven without any time to relax down to the unperturbed state. Therefore, the nonlinearity can significantly affect the baseline of subthreshold dynamics (either in *Eq. 5* or the full dynamics) and we should be careful about interpreting the estimated PRC.

Our method relies on the central assumption that the PRC is (at least approximately) sparse and can be constructed from a much smaller number of (Fourier) basis modes than the stimulus dimensionality given by the number of sampling points in an average ISI. This also implies that neurons are sensitive specifically to a low dimensional subspace (which we might call the feature space) of the space of all possible stimuli with much larger dimensionality. Recent investigations of how such input specificity, or feature selectivity, can arise from single neuron dynamics suggest that the sensitivity mechanism can be

understood via linearization of the neuronal dynamics in the subthreshold regime (Famulare and Fairhall 2010; Hong et al. 2007; Ostoic and Brunel 2011; Paninski 2006). Therefore, the dimensionality of the feature space, which impacts not only computation at the single neuron level (Hong et al. 2008) but also at the population level (Hong et al. 2012), is naturally limited by the number of the active mechanisms in the neuron, each of which contributes a characteristic time scale/frequency (Hong et al. 2007). The PRC characterizes how inputs modulate spontaneous firing rather than how the stimuli evoke spikes, as in the studies cited, and therefore it has rarely been interpreted as an input/output kernel [with notable exceptions (Ermentrout et al. 2007; Moehlis et al. 2006; Phoka et al. 2010)]. However, the neuron spends most of the interspike period in the subthreshold regime where the dynamics evolve slowly, and linearization with limited dimensionality should be able to provide a good approximation of the full dynamics. Therefore, we suggest that our assumption of the sparseness of the PRC is biophysically plausible. Our results indeed confirm the effectiveness of the sparseness assumptions particularly for PRC estimation.

In many neural systems, high input feature specificity and sparseness have been suggested as an underlying principle that enables efficient coding (Olshausen and Field 2004). Based on our results, we further suggest that the CS algorithms can be great tools for discovering and characterizing the input/output function of a single neuron and also of neural ensembles, particularly when the experimentally obtainable data are small but of large dimensionality (Ganguly and Sompolinsky 2012).

APPENDIX

Derivation of Eq. 8

Here we describe the detailed derivation of the linear approximation, *Eq. 8*. We first begin with the definition of the PRC

$$\frac{d\theta}{dt} = \frac{1}{T_0} + \lambda \Delta(0)x(t),$$

where λ is a dimensionless expansion parameter that we will set $\lambda = 1$ in the end. By integrating this equation, we get

$$\begin{aligned} 1 &= \frac{T}{T_0} + \lambda \int_0^T \Delta(0)x(t)dt = \frac{T}{T_0}\lambda \left(\frac{T}{T_0}\right) \int_0^{T_0} \Delta(0)x(\tau)d\tau \\ &\approx \frac{T}{T_0} + \lambda \left(\frac{T}{T_0}\right) \int_0^{T_0} \Delta(\tau/T_0)x(\tau)d\tau + O(\lambda^2), \end{aligned}$$

where $\tau = t(T/T_0)$. We define the ISI deviation r and rescale $x(\tau)$ as $\lambda x(\tau) \rightarrow x(\tau)$, and we obtain a simple linear expression, which is correct up to the first order in the magnitude of $x(t)$ (Ota et al. 2009),

$$r \approx \int_0^{T_0} \Delta(\tau/T_0)x(\tau)d\tau, \quad r = (T_0 - T)/T. \quad (17)$$

In actual data, the time is discretized with some time step $\delta\tau = T_0/L$, and a function of time is represented by an L -dimensional vector as $f(t) \rightarrow \mathbf{f} = \{f(0), f(\delta\tau), \dots, f((L-1)\delta\tau)\}$. For each i , *Eq. 17* then becomes

$$r_i = \sum_{j=1}^{L-1} \Delta(j\delta\tau/T_0) \cdot x_i(j\delta\tau) = \mathbf{x}_i \cdot \Delta,$$

where we also scaled the PRC as $\Delta\delta\tau \rightarrow \Delta$ for convenience. When we collect each r_i into a vector as $\mathbf{r} = [r_1, r_2, \dots, r_N]$, we get a discrete form of *Eq. 17*

$$\mathbf{r} = \mathbf{X}\Delta, \quad (18)$$

where $X_{ij} = x_i(j\delta\tau)$. Therefore, the PRC estimation has become a linear regression problem to find a vector Δ that predicts \mathbf{r} given the matrix \mathbf{X} .

So far, we have discussed the PRC represented in the time domain, but this may not be the best representation for efficient estimation when Δ can be constructed much more simply in a different representation, such as in the frequency domain via Fourier transformation. In general, we can consider a suitable set of basis vectors $u_{kj} = (\mathbf{u}_k)_j$ to transform *Eq. 18*. If the PRC can be constructed with K such basis vectors as

$$\Delta = \sum_{k=1}^K \tilde{\Delta}_k \mathbf{u}_k = \mathbf{u}^T \tilde{\Delta},$$

then we can rewrite *Eq. 18* as

$$\mathbf{r} = \Phi \tilde{\Delta}, \quad \Phi = \mathbf{X} \mathbf{u}^T, \quad \tilde{\Delta} = \mathbf{u} \Delta,$$

which is *Eq. 8*.

ACKNOWLEDGMENTS

We thank Ben Torben-Nielsen, Timm Lochmann, Mario Negrello, and Joao Couto for helpful discussion.

GRANTS

S. Hong and E. De Schutter are supported by Okinawa Institute of Science and Technology.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

Author contributions: S.H. designed the method and analyzed the data; Q.R. designed and did the experiments; S.H., Q.R., and E.D.S. wrote the paper.

REFERENCES

- Abouzeid A, Ermentrout B.** Type-II phase resetting curve is optimal for stochastic synchrony. *Phys Rev E Stat Nonlin Soft Matter Phys* 80: 011911, 2009.
- Achuthan S, Butera RJ, Canavier CC.** Synaptic and intrinsic determinants of the phase resetting curve for weak coupling. *J Comput Neurosci* 30: 373–390, 2011.
- Baraniuk RG.** Compressive sensing. *IEEE Signal Proc Mag* 24: 118–124, 2007.
- Baraniuk RG, Davenport M, DeVore RA, Wakin M.** A simple proof of the restricted isometry property for random matrices. *Constr Approx* 28: 253–263, 2008.
- Bishop CM.** *Pattern Recognition and Machine Learning (Information Science and Statistics)* (2nd ed). New York, NY: Springer, 2007.
- Candés EJ, Romberg JK.** Sparsity and incoherence in compressive sampling. *Inverse Problems* 23: 969–985, 2006.
- Candés EJ, Romberg JK, Tao T.** Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans Inf Theory* 52: 489–509, 2006a.
- Candés EJ, Romberg JK, Tao T.** Stable signal recovery from incomplete and inaccurate measurements. *Comm Pure Appl Math* 59: 1207–1223, 2006b.
- Candés EJ, Tao T.** Decoding by linear programming. *IEEE Trans Inf Theory* 51: 4203–4215, 2005.
- Candés EJ, Tao T.** Near-optimal signal recovery from random projections: universal encoding strategies? *IEEE Trans Inf Theory* 52: 5406–5425, 2006.
- Candés EJ, Tao T.** Dantzig selector: statistical estimation when p is much larger than n . *Ann Stat* 35: 2313–2351, 2007.
- Chen SS, Donoho DL, Saunders MA.** Atomic decomposition by basis pursuit. *SIAM J Sci Comput* 20: 33–61, 1998.
- Donoho DL.** Compressed sensing. *IEEE Trans Inf Theory* 52: 1289–1306, 2006.

- Donoho DL, Tanner J.** Neighborhood of randomly projected simplices in high dimensions. *Proc Natl Acad Sci USA* 102: 9452–9457, 2005a.
- Donoho DL, Tanner J.** Sparse nonnegative solution of underdetermined linear equations by linear programming. *Proc Natl Acad Sci USA* 102: 9446–9451, 2005b.
- Ermentrout B, Pascal M, Gutkin B.** The effects of spike frequency adaptation and negative feedback on the synchronization of neural oscillators. *Neural Comput* 13: 1285–1310, 2001.
- Ermentrout GB.** Type I membranes, phase resetting curves and synchrony. *Neural Comput* 8: 979–1001, 1996.
- Ermentrout GB.** *Simulating, Analyzing, and Animating Dynamical Systems: A Guide to XPPAUT for Researchers and Students*. Philadelphia, PA: Soc Industr Appl Math, 2002.
- Ermentrout GB, Galán RF, Urban NN.** Relating neural dynamics to neural coding. *Phys Rev Lett* 99: 248103, 2007.
- Famolare M, Fairhall A.** Feature selection in simple neurons: how coding depends on spiking dynamics. *Neural Comput* 22: 581–598, 2010.
- Forti L, Cesare E, Mapelli J, D'Angelo E.** Ionic mechanisms of autorhythmic firing in rat cerebellar Golgi cells. *J Physiol* 574: 711–729, 2006.
- Galán RF, Ermentrout GB, Urban NN.** Efficient estimation of phase-resetting curves in real neurons and its significance for neural-network modeling. *Phys Rev Lett* 94: 158101, 2005.
- Ganguly S, Sompolinsky H.** Compressed sensing, sparsity, and dimensionality in neuronal information processing and data analysis. *Ann Rev Neurosci* 35: 485–508, 2012.
- Goldobin DS, Pikovsky A.** Synchronization and desynchronization of self-sustained oscillators by common noise. *Phys Rev E* 71: 045201, 2005.
- Gutkin BS, Ermentrout GB, Reyes AD.** Phase-response curves give the responses of neurons to transient inputs. *J Neurophysiol* 94: 1623–1635, 2005.
- Hansel D, Mato G, Meunier C.** Synchrony in excitatory neural networks. *Neural Comput* 7: 307–337, 1995.
- Hastie T, Tibshirani R, Friedman JH.** *The Elements of Statistical Learning: Data mining, Inference, and Prediction* (2nd ed). New York, NY: Springer, 2009.
- Hines ML, Carnevale NT.** NEURON simulation environment. *Neural Comput* 9: 1179–1209, 1997.
- Hodgkin AL.** Local electric changes associated with repetitive action in a non-medullated axon. *J Physiol* 107: 165–181, 1948.
- Hodgkin AL, Huxley AF.** A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol* 117: 500–544, 1952.
- Hong S, Agüera y Arcas B, Fairhall AL.** Single neuron computation: from dynamical system to feature detector. *Neural Comput* 19: 3133–3172, 2007.
- Hong S, Lundstrom L, Fairhall AL.** Intrinsic gain modulation and adaptive neural coding. *PLoS Comput Biol* 4: e1000119, 2008.
- Hong S, Ratté S, Prescott SA, De Schutter E.** Single neuron firing properties impact correlation-based population coding. *J Neurosci* 32: 1413–1428, 2012.
- Izhikevich EM.** *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting (Computational Neuroscience)*. Cambridge, MA: MIT Press, 2010.
- Mainen ZF, Sejnowski TJ.** Reliability of spike timing in neocortical neurons. *Science* 268: 1503–1506, 1995.
- Mallat S, Zhang Z.** Matching pursuits with time-frequency dictionaries. *IEEE Trans Signal Process* 41: 3397–3415, 1993.
- Marella S, Ermentrout GB.** Class-II neurons display a higher degree of stochastic synchronization than class-I neurons. *Phys Rev E* 77: 041918, 2008.
- Markovsky I, Huffel SV.** Overview of total least-squares methods. *Signal Processing* 87: 2283–2302, 2007.
- Moehlis J, Shea-Brown E, Rabitz H.** Optimal inputs for phase models of spiking neurons. *J Comp Nonlinear Dynam* 1: 358–367, 2006.
- Morris C, Lecar H.** Voltage oscillations in the barnacle giant muscle fiber. *Biophys J* 35: 193–213, 1981.
- Nakae K, Iba Y, Tsubo Y, Fukai T, Aoyagi T.** Bayesian estimation of phase response curves. *Neural Netw* 23: 752–763, 2010.
- Nakao H, Arai K, Nagai K, Tsubo Y, Kuramoto Y.** Synchrony of limit-cycle oscillators induced by random external impulses. *Phys Rev E* 72: 026220, 2005.
- Natarajan BK.** Sparse approximate solutions to linear systems. *SIAM J Comput* 24: 227, 1995.
- Neher E.** Correction for liquid junction potentials in patch clamp experiments. *Methods Enzymol* 207: 123–131, 1992.
- Netoff TI, Banks MI, Dorval AD, Acker CD, Haas JS, Kopell N, White JA.** Synchronization in hybrid neuronal networks of the hippocampal formation. *J Neurophysiol* 93: 1197–1208, 2005.
- Olshausen BA, Field DJ.** Sparse coding of sensory inputs. *Curr Opin Neurobiol* 14: 481–487, 2004.
- Oprisan SA, Prinz AA, Canavier CC.** Phase resetting and phase locking in hybrid circuits of one model and one biological neuron. *Biophys J* 87: 2283–2298, 2004.
- Ostojic S, Brunel N.** From spiking neuron models to linear-nonlinear models. *PLoS Comput Biol* 7: e1001056, 2011.
- Ota K, Nomura M, Aoyagi T.** Weighted spike-triggered average of a fluctuating stimulus yielding the phase response curve. *Phys Rev Lett* 103: 024101, 2009.
- Paninski L.** The most likely voltage path and large deviations approximations for integrate-and-fire neurons. *J Comput Neurosci* 21: 71–87, 2006.
- Perkel DH, Schulman JH, Bullock TH, Moore GP, Segundo JP.** Pacemaker neurons: effects of regularly spaced synaptic input. *Science* 145: 61–63, 1964.
- Phoka E, Cuntz H, Roth A, Häusser M.** A new approach for determining phase response curves reveals that Purkinje cells can act as perfect integrators. *PLoS Comput Biol* 6: e1000768, 2010.
- Pinsker HM.** Aplysia bursting neurons as endogenous oscillators. I. Phase-response curves for pulsed inhibitory synaptic input. *J Neurophysiol* 40: 527–543, 1977.
- Pinsky MA.** *Introduction to Fourier Analysis and Wavelets (Graduate Studies in Mathematics)*. Providence, IL: American Mathematical Society, 2009.
- Prescott SA, de Koninck Y, Sejnowski TJ.** Biophysical basis for three distinct dynamical mechanisms of action potential initiation. *PLoS Comp Biol* 4: e1000198, 2008.
- Preyer AJ, Butera RJ.** Neuronal oscillators in aplysia californica that demonstrate weak coupling in vitro. *Phys Rev Lett* 95: 138103, 2005.
- Robberechts Q, Wijnants M, Giugliano M, De Schutter E.** Long-term depression at parallel fiber to Golgi cell synapses. *J Neurophysiol* 104: 3413–3423, 2010.
- Salman MA, Romberg J.** Dynamic updating for minimization. *IEEE J Sel Topics Signal Process* 4: 421–434, 2010.
- Smeal RM, Ermentrout GB, White JA.** Phase-response curves and synchronized neural networks. *Proc Trans R Soc B* 365: 2407–2422, 2010.
- Stiefel KM, Gutkin BS, Sejnowski TJ.** Cholinergic neuromodulation changes phase response curve shape and type in cortical pyramidal neurons. *PLoS One* 3: e3947, 2008.
- Tateno T, Robinson HPC.** Phase resetting curves and oscillatory stability in interneurons of rat somatosensory cortex. *Biophys J* 92: 683–695, 2007.
- Teramae J, Tanaka D.** Robustness of the noise-induced phase synchronization in a general class of limit cycle oscillators. *Phys Rev Lett* 93: 204103, 2004.
- Teramae J, Fukai T.** Temporal precision of spike response to fluctuating input in pulse-coupled networks of oscillating neurons. *Phys Rev Lett* 101: 248105, 2008.
- Torben-Nielsen B, Usuari M, Stiefel KM.** A comparison of methods to determine neuronal phase-response curves. *Front Neuroinf* 4: 6, 2010.
- Vervaeke K, Lorincz A, Gleeson P, Farinella M, Nusser Z, Silver RA.** Rapid desynchronization of an electrically coupled interneuron network with sparse excitatory synaptic input. *Neuron* 67: 435–451, 2010.
- Winfree AT.** *Geometry of Biological Time* (2nd ed). New York, NY: Springer, 2001.
- Zhang Z, Rao BD.** Sparse signal recovery with temporally correlated source vectors using sparse bayesian learning. *IEEE J Sel Topics Signal Process* 5: 912–926, 2011.