A-current and type I/type II transition determine collective spiking from common input

Andrea K. Barreiro, Evan L. Thilo, and Eric Shea-Brown
Department of Applied Mathematics and Program in Neurobiology and Behavior, University of Washington, Seattle, Washington

Submitted 17 October 2011; accepted in final form 5 June 2012

Barreiro AK, Thilo EL, Shea-Brown E. A-current and type I/type II transition determine collective spiking from common input. J Neurophysiol 108: 1631–1645, 2012. First published June 6, 2012; doi:10.1152/jn.00928.2011.—The mechanisms and impact of correlated, or synchronous, firing among pairs and populations of neurons are under intense investigation throughout the nervous system. A ubiquitous circuit feature that can give rise to such correlations consists of overlapping, or common, inputs to pairs and populations of cells, leading to common spike train responses. Here, we use computational tools to study how the transfer of common input currents into common spike outputs is modulated by the physiology of the recipient cells. We focus on a key conductance, \( g_A \), for the A-type potassium current, which drives neurons between “type II” excitability (low \( g_A \)) and “type I” excitability (high \( g_A \)). Regardless of \( g_A \), cells transform common input fluctuations into a tendency to spike nearly simultaneously. However, this process is more pronounced at low \( g_A \) values. Thus, for a given level of common input, type II neurons produce spikes that are relatively more correlated over short time scales. Over long time scales, the trend reverses, with type II neurons producing relatively less correlated spike trains. This is because these cells’ increased tendency for simultaneous spiking is balanced by an anti-correlation of spikes at larger time lags. These findings extend and interpret prior findings for phase oscillators to conductance-based neuron models (Salinas and Sejnowski 2000). We demonstrate a novel implication for neural signal processing: downstream cells with long time constants are selectively driven by type I cell populations upstream and those with short time constants by type II cell populations. Our results are established via high-throughput numerical simulations and explained via the cells’ filtering properties and nonlinear dynamics.

spike time correlations: synchrony; linear response; spike-triggered average; Connor-Stevens model

NEURONS THROUGHOUT THE NERVOUS system, from the retina (Shlens et al. 2008), thalamus (e.g., Alonso et al. 1996), and cortex (e.g., Zohary et al. 1994) to motoneurons (Binder and Powers 2001), show temporal correlation between the discharge times of their spikes. This correlated spiking can impact sensory discrimination (Averbeck et al. 2006) and signal propagation (Salinas and Sejnowski 2000).

How do these correlations arise? We study a simple mechanism in which the inputs to a pair or population of neurons has a common component that is shared across multiple cells (Fig. 1). On an anatomical level, the large number of divergent connections that span layers and areas makes shared afferents to pairs of nearby cells unavoidable (Shadlen and Newsome 1998). Correlated spiking in areas upstream from the target cells can add to this anatomical factor. In fact, for some neural circuits, shared inputs are themselves the dominant source of correlated spiking (Trong and Rieke 2008). In general, correlating effects of shared input interact with effects of recurrent coupling (cf. Ostojic et al. 2009).

What makes shared input circuitry especially interesting is the pivotal role of spike-generating dynamics. For a given fraction of shared input, these dynamics control the fraction of shared output; that is, the fraction of spikes that will be shared across the two cells. This correlation transfer depends on two factors. The first is the mechanism of spike generation. The second is the operating point of the neurons (i.e., their rate and variability of firing or the strength of DC and time-varying inputs that they receive; Binder and Powers 2001; de la Rocha et al. 2007). Excepting Hong and De Schutter (2008), studies of correlation transfer have mostly focused on simplified neuron models, such as integrate-and-fire, phase, or threshold crossing systems, leaving open allied questions for models with more complex subthreshold and after-spike dynamics.

Here, we study correlation transfer for a family of conductance-based neuron models that spans a wide range of excitability properties. This is the spectrum from type I excitability, in which firing can occur at arbitrarily low rates in response to a DC current (as for cortical pyramidal cells), to type II excitability, in which firing occurs at a non-zero “onset” rate (as for fast-spiking interneurons or the Hodgkin-Huxley model) (Rinzel and Ermentrout 1998; Hodgkin 1948; Izhikevich 2007). We use the Connor-Stevens model (Connor and Stevens 1971), which transitions between type I and type II as \( g_A \), the maximal conductance of the A-type potassium current, is varied (see Fig. 2). Beyond firing rates, type I vs. type II neurons with different levels of \( g_A \) differ in single-cell computation (Ermentrout et al. 2007) and synchronization under reciprocal coupling (see e.g., Rinzel and Ermentrout 1998; Ermentrout and Terman 2010 and references therein). While we use computational models here, we note that the dynamical response properties of type I and type II cells that will be central to our analysis here have been observed in in vitro studies (for example, see Reyes and Fetz 1993; Galán et al. 2005; Tateno and Robinson 2007; Netoff et al. 2005).

Here, we ask: how does \( g_A \) affect correlated spiking driven by common inputs? In superthreshold firing regimes (i.e., with DC inputs that drive periodic spiking), our findings for Connor-Stevens neurons make contact with earlier results for “normal form” phase oscillator models (Barreiro et al. 2010; Marella and Ermentrout 2008; Galán et al. 2007). We extend our investigation to subthreshold regimes and demonstrate novel trends in how correlation transfer depends on the oper-
Fig. 1. Shared input microcircuit, in which two neurons receive input currents with a common component that represents correlated activity or shared afferents upstream. Each neuron is a single-compartment Connor-Stevens model (see methods), with a maximal A-current conductance $g_A$ that we vary, eliciting a full range of type I to type II spiking dynamics. Shared input currents lead to correlated spikes, which are quantified as shown via spike counts $n_1, n_2$ over sliding time windows of length $T$. The input currents received by each cell have mean $\mu$ and fluctuate with total variance $\sigma^2$; the common noise is chosen with variance $\sigma^2$ and independent noise terms with variance $\sigma^2(1 - c)$.

Atting point of the neurons. We explain and interpret our findings via the common-input filtering properties of individual neurons in the presence of independent “background” noise. Finally, we demonstrate how the distinct features of correlated spiking in type I vs. type II neurons manifest in signal transmission in a simple feedforward circuit. Preliminary versions of some findings have appeared in abstract form (Barreiro 2009; Shea-Brown et al. 2009).

METHODS

Circuit setup. We primarily consider the feedforward circuit of Fig. 1. Here, each of two neurons receives two sources of fluctuating current: a common, or “shared” source $I_c = \sigma \sqrt{c} \xi(t)$, and a “private” source $I_i = \sigma \sqrt{1-c} \xi(t)$, or $I_i = \sigma \sqrt{1-c} \xi(t)$. Each of these inputs is chosen to be a scaled statistically independent, Gaussian white noise process (uncorrelated in time); that is, $(\xi(t) \xi(t + \tau)) = \delta(\tau)$ for $i = 1, 2, c$. This is for simplicity and agreement with prior studies of correlated spiking (Lindner et al. 2005; de la Rocha et al. 2007; Shea-Brown et al. 2008; Marella and Ermentrout 2008; Vilela and Lindner 2009; Barreiro et al. 2010). The common current $I_c$ has variance $\sigma^2 c$; each private current has zero mean and variance $\sigma^2(1 - c)$. Note that these scalings are chosen so that the total variance of current injected into each cell is always $\sigma^2$; while the parameter $c$ gives the fraction of this variance that arises from common input sources. For example, when $c = 0.5$, 50% of each neuron’s presynaptic inputs come from the shared and 50% from the independent input. Finally, the mean of the total current received by each cell is given by $\mu$. This term represents the total bias toward negative or positive currents from all sources; in Fig. 1, it is illustrated as part of the common input for simplicity.

The combined currents, $I_{app,i}(t) = \mu + I_c(t) + I_i(t)$ ($i = 1, 2$) are injected into identical single-compartment, conductance-based membrane models (see Neuron model); spike times are identified from the resulting voltage trace.

There are a number of ways in which overlapping and correlated presynaptic cells can provide a pair of neurons with input currents that have a given correlation coefficient $c$ (Renart et al. 2010; Rosenbaum et al. 2010). We specify a simple but illustrative example to offer one interpretation of this value. Consider a case in which a fraction $p$ of the cells presynaptic to each of our two neurons is drawn from a “shared,” correlated pool of neurons. Let the remaining inputs to each cell be drawn from a similar pool of cells, such that cells within the pool for each cell are correlated with one another but are independent of cells in the other pools. Finally, for simplicity, let within-pool correlation coefficients be the same in each pool. Then, for large pool sizes, the total inputs to each cell have correlation (cf. Rosenbaum et al. 2010)

$$c = \frac{p^2}{p^2 + (1 - p)^2}.$$ 

In particular, the value $c = 0.1$ used repeatedly in our study results from a shared fraction $p = 1/4$.

Neuron model. We investigate correlation transfer in the Connor-Stevens model, which was designed to capture the low firing rates of a crab motor axon (Connor and Stevens 1971; Connor et al. 1977). This model adds a transient potassium current, or A-current, to sodium and delayed rectifier potassium currents of Hodgkin-Huxley type. The A-type channel provides extended after-spike hyperpolarization currents, which lead to arbitrarily low firing rates and hence type I excitability (see Introduction).

The voltage equation is

$$\frac{dV}{dt} = -g_L(V - E_L) - g_NAP(V - E_Na) - g_Kg_A(V - E_K) - g_Ag_A(B(V - E_K)) + I_{app}. \quad (1)$$

The gating variables $m, n, h, A,$ and $B$ each evolve according to the standard voltage-gated kinetics; e.g., for $m$:

$$\frac{dm}{dt} = \frac{m_s(V) - m}{\tau_m(V)} \quad \text{(2)}$$

where $m_s(V)$ is the steady-state value and $\tau_m(V)$ is the (voltage-dependent) time constant. All equations and parameters are exactly as specified in Connor et al. (1977), with the exception that we vary the maximal A-current conductance, $g_A$, over the range of values reported below. As $g_A$ is decreased from the value set by Connor et al. (1977), the neuron transitions from type I to type II excitability.

Measuring spike train correlation. We represent the output spike trains as sequences of impulses $y_i(t) = \sum_k \delta(t - t_k^i)$, where $t_k^i$ is the time of the $k$th spike of the $i$th neuron. The firing rate of the $i$th cell, $(\langle y_i(t) \rangle)$, is denoted $\nu_i$. To quantify correlation over a given time scale $T$,
we compute the Pearson’s correlation coefficient of spike counts over a time window of length $T$ (as in, e.g., Zohary et al. 1994; Bair et al. 2001):

$$
\rho_T = \frac{\text{Cov}(n_1, n_2)}{\sqrt{\text{Var}(n_1)} \sqrt{\text{Var}(n_2)}}
$$

where $n_1$, $n_2$ are the numbers of spikes simultaneously output by neurons 1 and 2 respectively, in a time window of length $T$; i.e.,

$$
n_i(t) = \int_0^T g_A(y_1, y_2) \, ds.
$$

If $\rho_T$ is measured at values of $T$ that are less than a typical interspike interval, we are essentially measuring the degree of synchrony between individual spikes. For larger $T$ values, $\rho_T$ assesses total correlation between numbers of spikes emitted by each cell.

A short calculation (cf. Bair et al. 2001; Cox and Lewis 1966) shows that Cov($n_1$, $n_2$) is

$$
\text{Cov}(n_1, n_2) = T \int_0^T C_{12}(t) T - \left( \frac{t}{T} \right)^2 dt
$$

where the spike train cross-covariance $C_{12}(\tau) = \langle y_1(t) y_2(t + \tau) \rangle - \langle y_1(t) \rangle \langle y_2(t + \tau) \rangle$ satisfies

$$
\text{Var}(n_1) = T \int_0^T A_1(t) T - \left( \frac{t}{T} \right)^2 dt,
$$

and similarly for neuron 2.

Linking linear response theory, spike-triggered averages, and spike count correlations. When the variance $c$ of the shared input is small (see Fig. 1), then we can treat the circuit with a shared input as a perturbation from two independently firing neurons. We describe this perturbation via linear response theory (Lindner et al. 2005; de la Rocha et al. 2007; Ostojic et al. 2009; Ostojic and Brunel 2011), which is related to classical Linear-Poisson models of neural spiking (Perkel et al. 1967).

That is, we make the assumption that the change in a neuron’s instantaneous firing rate $v_i(t)$ due to the shared input signal can be represented by linearly filtering the common (perturbing) input $I_c$:

$$
v_i(t) = \langle y_i(t) \rangle I_c = v_{0,i} + \int_0^t K(s) I_c(t - s) ds
$$

where the filter $K(t)$ is 0 for $t < 0$ (causality) and $v_{0,i}$ is the “background” average firing rate of the independently firing neuron (Ostojic et al. 2009).

Equation 4 is extremely useful, because it isolates the common component of the response of neurons $i = 1 \text{ and } i = 2$, which is an enhanced (or depressed) tendency to emit spikes, at a rate determined by the filtered, common input. As a result, the cross-covariance function $C_{12}(\tau) = \langle (v_1(t) - v_{0,1})(v_2(t + \tau) - v_{0,2}) \rangle$ becomes

$$
C_{12}(\tau) = \text{Var}(I_c) \int_0^\infty K(s) K(s + \tau) ds
$$

where $K(t)$ is precisely given by the spike-triggered average STA($\tau$) (Gabbiani and Koch 1998; Dayan and Abbott 2001), different from the classical setting, but as in Ostojic et al. (2009), it is only the common component of the input current that is averaged in this procedure. Specifically, if the neuron produces $N$ spikes at times $t_k$, then STA($\tau$) = \frac{1}{N} \sum_{k=1}^N I_c(t_k - \tau)$ and in the limit of large $N$,

Below, we use this expression to derive $K(\tau)$ from numerically computed STAs.

Numerical simulations and estimates of spike count statistics. To compute spike count correlations and other statistical quantities, we performed Monte Carlo simulations of the circuit in Fig. 1. We compute spike count correlations over a time window of length $T$, i.e., $n_i(t) = \int_0^T y_i(s) \, ds$. When $\rho_T$ is measured at values of $T$ that are less than a typical interspike interval, we are essentially measuring the degree of synchrony between individual spikes. For larger $T$ values, $\rho_T$ assesses total correlation between numbers of spikes emitted by each cell.

A short calculation (cf. Bair et al. 2001; Cox and Lewis 1966) shows that Cov($n_1$, $n_2$) is

$$
\text{Cov}(n_1, n_2) = T \int_0^T C_{12}(t) T - \left( \frac{t}{T} \right)^2 dt (3)
$$

where the spike train cross-covariance $C_{12}(\tau) = \langle y_1(t) y_2(t + \tau) \rangle - \langle y_1(t) \rangle \langle y_2(t + \tau) \rangle$ satisfies

$$
\text{Var}(n_1) = T \int_0^T A_1(t) T - \left( \frac{t}{T} \right)^2 dt,
$$

and similarly for neuron 2.

Linking linear response theory, spike-triggered averages, and spike count correlations. When the variance $c$ of the shared input is small (see Fig. 1), then we can treat the circuit with a shared input as a perturbation from two independently firing neurons. We describe this perturbation via linear response theory (Lindner et al. 2005; de la Rocha et al. 2007; Ostojic et al. 2009; Ostojic and Brunel 2011), which is related to classical Linear-Poisson models of neural spiking (Perkel et al. 1967). That is, we make the assumption that the change in a neuron’s instantaneous firing rate $v_i(t)$ due to the shared input signal can be represented by linearly filtering the common (perturbing) input $I_c$:

$$
v_i(t) = \langle y_i(t) \rangle I_c = v_{0,i} + \int_0^t K(s) I_c(t - s) ds
$$

where the filter $K(t)$ is 0 for $t < 0$ (causality) and $v_{0,i}$ is the “background” average firing rate of the independently firing neuron (Ostojic et al. 2009).

Equation 4 is extremely useful, because it isolates the common component of the response of neurons $i = 1 \text{ and } i = 2$, which is an enhanced (or depressed) tendency to emit spikes, at a rate determined by the filtered, common input. As a result, the cross-covariance function $C_{12}(\tau) = \langle (v_1(t) - v_{0,1})(v_2(t + \tau) - v_{0,2}) \rangle$ becomes

$$
C_{12}(\tau) = \text{Var}(I_c) \int_0^\infty K(s) K(s + \tau) ds
$$

where $K(t)$ is precisely given by the spike-triggered average STA($\tau$) (Gabbiani and Koch 1998; Dayan and Abbott 2001), different from the classical setting, but as in Ostojic et al. (2009), it is only the common component of the input current that is averaged in this procedure. Specifically, if the neuron produces $N$ spikes at times $t_k$, then STA($\tau$) = \frac{1}{N} \sum_{k=1}^N I_c(t_k - \tau)$ and in the limit of large $N$,

Below, we use this expression to derive $K(\tau)$ from numerically computed STAs.

Numerical simulations and estimates of spike count statistics. To compute spike count correlations and other statistical quantities, we performed Monte Carlo simulations of the circuit in Fig. 1. The governing Connor-Stevens equations (1, 2) were integrated using the stochastic Euler method with time step $\Delta t = 0.01$ ms for a total time $T_{\text{max}}$ of $8 \times 10^7$ ms. Random input currents were chosen at each time step using a standard random number generator (Marsaglia and Zaman 1994). To facilitate exploration of parameter space $\mu, \sigma, c$, and $\tau_A$, we distributed computations on parallel machines through the NSF Tera-grid program (http://www.teragrid.org). The simulation code was implemented in FORTRAN90 and distribution scripts in Python for running on clusters with and without PBS submission protocols. All code and scripts will be available at the modelDB site upon publication (http://senselab.med.yale.edu/modeldb).

We register spikes in our simulations at times when the membrane voltage exceeds $-30$ mV and maintains a positive slope in voltage for the next three time steps (0.03 ms). To avoid counting each spike more than once, we omit a 2-ms refractory period after each spike.

Spike count statistics were computed directly from the recorded spike times, based on a single long simulation, after discarding an initial transient (200 ms). When sampling spike counts over a time window $T$, we advance the window by $1/4 T$, resulting in $\sim 4 T_{\text{max}}/T$ (correlated) samples; consequently, our estimates of spike counts become noisier as $T$ increases. To estimate SEs on spike count statistics, we further divided the simulation into 10 equal time intervals ($8 \times 10^5$ ms each) and computed statistics on each subsimulation; SD, divided by $\sqrt{10}$, gives us an estimated SE. When appropriate, these are presented along with the mean estimates, as error bars.

Below, we also report STAs described above; these were computed using long simulations of length $8 \times 10^7$ ms for several sets of parameter values $\mu, \sigma$, and $\tau_A$. To compute these, the common input current $I_c$ was treated as the “signal” that was averaged and the private input as a “background” that was not. We used $c = 0.10$ in this computation. In our code, the history $I_c$ was continuously recorded for a duration into the past; when a spike was recorded, the STA was augmented by this current.

Finally, we generate auto- and cross-correlograms (shown in Fig. 4) by collecting interspike intervals from our simulations in 1-ms-long bins. These are used, after the standard normalization, as auto- and cross-covariance functions.

RESULTS

Rich structure of spike count correlations over short and long time scales. Our central findings contrast how different conductance-based neuron models produce correlated spiking when they receive overlapping fluctuating inputs, via the shared-input circuitry in Fig. 1. Specifically, we show how this correlation depends on the type I vs. type II excitability class of a neuron described by the well-studied Connor-Stevens model. As discussed above, neurons are often classified as type I vs. type II based on whether their firing rate-currentcurves are continuous (type I) vs. discontinuous (type II) at $\mu = I_{\text{th}}$, the threshold current above which periodic spiking can be elicited. Figure 2 demonstrates, as shown in (Rush and Rinzel 1995), that the Connor-Stevens model is type II when the maximal A-current conductance $g_A \approx 0$ mS/cm$, type I for $g_A \approx 60$ mS/cm$^2$, and displays a gradual transition in between. Thus we fix the neurons in the shared-input circuit to a point along the...
spectrum from type I to type II excitability by choosing different values of \( g_A \).

To compute levels of correlated spiking, we then fix the correlation in the input currents, that is, the fraction of the current variance that is shared vs. private to the two cells, to a preset value \( c \). For each value of \( c \) and \( g_A \), we compute spike count correlations for wide range of operating points for the neurons, as determined by a \(~200 \times 50\) grid of values for the mean current \( \mu \) and variance \( \sigma^2 \) (both \( \mu \) and \( \sigma \) are sampled at a resolution of 0.1 \( \mu \text{A/cm}^2 \) and \( \mu \text{A-ms}^{1/2}/\text{cm}^2 \), respectively). Specifically, we vary \( \mu \) over values centered at the threshold current \( I_{\text{th}}(g_A) \), from a minimum \( \mu = I_{\text{th}}(g_A) - 10(\mu \text{A/cm}^2) \) to a maximum \( \mu = I_{\text{th}}(g_A) + 10(\mu \text{A/cm}^2) \) for each value of \( g_A \). This enables us to cover, respectively, both subthreshold [i.e., fluctuation-driven, \( \mu < I_{\text{th}}(g_A) \)] and superthreshold [i.e., mean-driven, \( \mu > I_{\text{th}}(g_A) \)] firing regimes for each value of \( g_A \). We additionally vary \( \sigma \) over values of \( 0 < \sigma \leq 5(\mu \text{A-ms}^{1/2}/\text{cm}^2) \), so that we cover the range from nearly Poisson, irregular spiking to nearly periodic, oscillatory spiking. This is demonstrated by Fig. 3, which shows the Fano factor of spike counts over a long time window (\( T = 256\) ms), a proxy for the squared interspike interval coefficient of variation (Gabbiani and Koch 1998), over the entire \( \mu, \sigma \) parameter space for three representative values of \( g_A \) (0, 30, and 60 \( \text{mS/cm}^2 \)). For each value of \( g_A \), the Fano factor spans a range from near zero (periodic) to one (i.e., Poisson-like) or higher.

For each set of parameters \( g_A, c, \mu, \) and \( \sigma \), we compute the Pearson’s correlation coefficient \( \rho_T \) between the spike counts that the neuron pair in Fig. 1 produces in time windows of length \( T \). Figure 4 summarizes the results, for inputs with 10% shared variance (\( c = 0.1 \)). Here, we view \( \rho_T \) over the entire \( \mu, \sigma \) parameter space for three representative values of \( g_A \) (0, 30, and 60 \( \text{mS/cm}^2 \)) and two different time windows (\( T = 4\) ms and \( T = 128\) ms). Values of \( \rho_T \) depend in a strong but systematic way on all of the parameters we have introduced. As we move down a column, we see major qualitative differences in the levels of correlation that emerge at different points through the type II (\( g_A = 0 \)) to type I (\( g_A = 60 \)) spectrum. Within each panel, the operating point set by input current and variance (\( \mu, \sigma \)) has a strong impact on \( \rho_T \). Finally, the levels and trends in \( \rho_T \) depend strongly on the time scale \( T \). We now describe these trends in more detail; the sections that follow will give an explanation for how they arise.

We begin with Fig. 4, top, which shows correlation \( \rho_T \) for \( g_A = 0 \) and hence type-II excitability. First, note that correlations are overall quite weak. The largest values of \( \rho_T \) obtained are \( \approx 0.04 \), indicating that \( \approx 40\% \) or less of correlations in input currents are ever transferred into correlations in output spikes. Moreover, the level of correlations \( \rho_T \) and their dependence on input parameters \( \mu \) and \( \sigma \) appear roughly similar for both short and long time scales \( T \). In both cases, for a fixed value of \( \mu \), a general trend is that \( \rho_T \) gradually increases with fluctuation strength \( \sigma \). For a fixed value of \( \sigma \), in general \( \rho_T \) first increases and then decreases with \( \mu \); the dependence is slightly more complex at longer \( T \). Significantly non-zero values of \( \rho_T \) are present for \( \mu < I_{\text{th}} \) as \( \sigma \) becomes appreciably high; this reflects the bistable firing dynamics of the underlying deterministic system, which supports both a stable resting state and a stable spiking trajectory for \( \mu < I_{\text{th}} \).

For type I excitability at \( g_A = 60 \) (Fig. 4, bottom), the picture is dramatically different. First, there is a marked difference between correlation elicited on short vs. long time scales \( T \), with much stronger correlations observed for larger \( T \). Moreover, correlations produced by type I neurons over longer time scales \( T \) are much higher than those observed for type II neurons at any time scale: the largest values of \( \rho_T \) obtained for type I are \( \approx 0.08 \), indicating that \( \approx 80\% \) of correlations in input currents can be transferred into spike correlations. Conversely, correlations for type I neurons are strongly suppressed at short time scales, where \( \approx 10\% \) of input correlations are transferred. Overall, trends in \( \rho_T \) as \( \mu \) and \( \sigma \) vary are similar to those found previously: correlations increase with \( \sigma \), and first increase, then decrease, with \( \mu \).

Correlation transfer in the intermediate model, \( g_A = 30 \), displays trends between those of the type I (\( g_A = 60 \)) and type II (\( g_A = 0 \)) cases. As when \( g_A = 60 \), spike count correlations \( \rho_T \) are very low for short time windows \( T \) and attain intermediate to high values \( \approx 60\% \) of input correlations transferred for longer \( T \). As for both \( g_A = 60 \) and \( g_A = 0 \), \( \rho_T \) increases with noise magnitude \( \sigma \) and displays a nonmonotonic trend with mean current \( \mu \).

We obtain additional insight into how spike count correlations depend on type I vs. type II spike generation and the time scale \( T \) by choosing matched values of input parameters \( \mu \) and \( \sigma \) and comparing spike count correlations produced for different values of the A-current conductance \( g_A \). We first concentrate on the \( \mu \) and \( \sigma \) values indicated by squares and circles in

Fig. 3. Fano factor of spike counts over a long time window (\( T = 256\) ms) for a \( \approx 200 \times 50 \) grid of values for the mean current \( \mu \) and variance \( \sigma^2 \). From top to bottom, type II to type I, \( g_A = 0 \) (A), \( g_A = 30 \) (B), and \( g_A = 60 \) mS/cm\(^2\) (C). Markers indicate relative location of (\( \mu, \sigma \))-pair: subthreshold by 1 \( \mu \text{A/cm}^2 \) (diamond), superthreshold by 2 \( \mu \text{A/cm}^2 \) with low noise (circle) and high noise (square), and superthreshold with matched Fano factors (asterisk, see text).
Fig. 4. Both of these points indicate superthreshold inputs $\mu = I_{\text{inh}}(g_A) + 2$ (\mu A/cm$^2$) for all $g_A$ values. The square corresponds to higher noise $\sigma = 5$ \mu A-ms$^{1/2}$/cm$^2$, and the circle to lower noise $\sigma = 1$ \mu A-ms$^{1/2}$/cm$^2$. In Fig. 5, we plot $\rho_T$ for a full range of $T$ values from 1 to 200 ms, for nine values of $g_A$ between $g_A = 0$ and $g_A = 60$ (thus filling in intermediate values of $g_A$ and $T$ between those in Fig. 4). For both superthreshold cases, we see that type II neurons transfer more input correlation into output (spike) count correlation at small $T$, while type I neurons transfer more at large $T$; this transition occurs, roughly, at a value $T_{\text{switch}}$ indicated by the dotted line. We note that for the low noise case $\sigma = 1$ (Fig. 5B), the trends appear less ordered as $g_A$ varies; as we will see in the next section, this is because the cross-covariance function is more oscillatory here, so that $\rho_T$ has not yet converged to its asymptotic large $T$ value.

Subthreshold points, denoted by diamonds in Fig. 4, were also compared: the mean input current is chosen to be $\mu = I_{\text{inh}}(g_A) - 1$ (\mu A/cm$^2$), and the noise magnitude to be $\sigma = 5$ \mu A-ms$^{1/2}$/cm$^2$. Here, the differing dynamical structure between type II and type I neurons is evident in the firing statistics (see Table 1): while the bistable type II neuron ($g_A = 0$) sustains a substantial firing rate, the monostable type I neuron ($g_A = 60$) barely fires at this level of input current. The correlation coefficient $\rho_T$ is also very low for $g_A = 60$ at all time windows (Fig. 5C); this is consistent with the relationship between correlation and firing rate identified in earlier studies (de la Rocha et al. 2007; Shea-Brown et al. 2008). Overall, note that the correlation coefficient $\rho_T$ increases steadily with $T$ for the type I neurons (high $g_A$) but stays roughly constant over a broad range of $T$ for type II neurons (low $g_A$). While, we do not observe a clear value of $T$ switch for all values of $g_A$ for the subthreshold point in Fig. 5C, we see the same relative trends as for superthreshold points. Below, we will see how this effect follows from filtering properties of type I vs. type II cells.

Because spike generation mechanisms vary widely as $g_A$ changes, the neuron models with matched input statistics at different values of $g_A$ in Fig. 5, A–C, do not all have the same firing variability. In Fig. 6A, we address this by showing that the same trends in $\rho_T$ persist if we select values of $\mu$ to maintain constant firing variability for each value of $g_A$ [see Table 1; variability measured via large-time ($T = 256$ ms) Fano factor]. Here, we fix $\sigma = 5$ \mu A-ms$^{1/2}$/cm$^2$; the required current value for each $g_A$ is indicated with a red asterisk in Fig. 4.

By considering a wider range of input current parameters $\mu$ and $\sigma$, we can find operating points at which both firing variability and rate are matched. If we choose $\mu = 9.35$ \mu A/cm$^2$ and $\sigma = 7.3$ \mu A-ms$^{1/2}$/cm$^2$ for $g_A = 30$, or $\mu = 22.75$ \mu A/cm$^2$ and $\sigma = 8.1$ \mu A-ms$^{1/2}$/cm$^2$ for $g_A = 60$, we find that the cell fires at 113 Hz with a large-time ($T = 256$ ms) Fano factor of 0.059, matching the output statistics of the superthreshold, $g_A = 0$ operating point. Once again, we see that the trends in $\rho_T$ remain (Fig. 6B).

In sum, for matched values of the mean and variance of input currents, a pair of superthreshold type II (vs. type I) neurons will produce greater spike count correlations $\rho_T$ at short time scales $T$. For a wide range of choices for the mean and variance, there will be a value of $T_{\text{switch}}$ where this relationship reverses, so that type I (vs. type II) neurons produce greater $\rho_T$ for $T > T_{\text{switch}}$. For matched subthreshold currents, similar trends are present; overall, the presence of a time $T_{\text{switch}}$ depends on how the input statistics are chosen.

Finally, we note that the general trends observed here carry over, largely unchanged, to different values of $c$. Figure 7
covariance functions shown in the Fig. 4, approximate time window to type I cells transferring more as $A$, $B$ superthreshold; $B$ ductance characteristics of input currents and the level of the A-current conductance $g_A$ vary. We now demonstrate this via the cross-covariance functions for neuron pairs and how they differ as the characteristics of input currents and the level of the A-current conductance $g_A$ vary. We now demonstrate this via the cross-covariance functions shown in the Fig. 4, right-most column; these are for the superthreshold high noise cases ($\sigma = 5 \mu A\cdot ms^{-1/2}/cm^2$) discussed in the previous section.

To make the connection, recall that the spike count covariance, $\text{Cov}(n_1, n_2)$, measured over a window of duration $T$ is given by the integral of the cross-covariance function $C_{12}(\tau)$ against a triangular kernel of width $T$ (see Methods, and Eq. 3). Thus, for short windows $T$, only the central peak of $C_{12}(\tau)$ contributes to spike count covariance. In the limit of long windows $T \rightarrow \infty$, the spike count covariance is simply the integral of the cross-covariance function, multiplied by $T$, over the whole $\tau$-axis. Spike-count correlation $\rho_T$ is then given by the ratio of spike count covariance to the spike count variance.

As we will show below, $\rho_T$ and $\frac{\text{Cov}(n_1, n_2)}{T}$ often show similar trends with $T$. Both quantities are of interest: while $\rho_T$ gives a normalized metric of correlation, $\frac{\text{Cov}(n_1, n_2)}{T}$ is the relevant quantity to analyze impact on downstream excitable cells (see below, Readout of correlated spiking by downstream cells).

Armed with these relationships between $C_{12}(\tau)$, $\text{Cov}(n_1, n_2)$, and $\rho_T$, we revisit the trends observed in the previous section and explore their origin. Starting in Fig. 4, top right, note that $C_{12}(\tau)$ has a much larger central peak, and hence short-$T$ spike count covariance, for type II excitability ($g_A = 0$) than for type I (Fig. 4, bottom right, $g_A = 60$). This is also clear in Fig. 8A, where we plot spike count covariance vs. $T$. Over long windows $T$, the trend reverses. For $g_A = 0$, the cross-covariance function shows oscillations with significant negative and positive lobes. These lobes tend to cancel as $C_{12}(\tau)$ is integrated over long windows $T$. This cancellation results in little overall change in values of spike-count covariance computed at increasingly long values of $T$. For type I excitability, however, $C_{12}(\tau)$ is mostly positive, so that spike count covariance increases with $T$.

As discussed above, spike-count correlation $\rho_T$ is given by the ratio of spike-count covariance and variance. Comparing Figs. 8A and 5A it is clear that spike-count correlation and spike-count covariance display the same trends as $g_A$ is varied. For example, for very short times $T$ spike-count correlation $\rho_T$
is given by the ratio of the peak in $C_{22}(\tau)$ to that in $A_{1}(\tau)$; this ratio is also larger for type II vs. type I excitability.

For other operating points $(\mu, \sigma)$, while trends in spike-count correlation and spike-count covariance do not exactly agree, the same relative trends persist. Specifically, the general pattern that type II cells produce greater covariance over short time windows, and that this trend disappears or reverses for larger time windows, holds for each of the superthreshold operating points explored here. Overall, the major trends in spike count correlations stem from the presence vs. absence of large negative lobes in cross-covariance functions for type II vs. type I neurons. We next describe how this difference arises via the distinct filtering properties of the two neuron types.

**Common-input STAs reliably predict spike count covariance.**

The previous section showed how trends in spike count covariances for type I vs. type II neurons follow from the presence of both strongly negative and positive lobes in crosscovariance functions for type II neurons. Here, we explain the origin of this phenomenon. *Equations 5 and 6 (see METHODS*) provide the key link, in which the cross-covariance function is given in terms of a cell’s STA, which is an estimate of the filter through which cells turn incoming currents into time-dependent spiking rates. Here, as in Ostojic et al. (2009), we define the common-input STA as an average of the common current only that precedes spikes over a single long realization:

$$\text{STA} = \left( \frac{1}{N} \sum_{k=1}^{N} I(t_k - \tau) \right), \tag{7}$$

where the $t_k$ are the $N$ spike times from the realization.

We first show that the prediction of spike count covariances from STAs is accurate. Figure 8, *left*, shows close agreement between spike count covariances computed from “full” numerical simulation (thin lines) vs. predictions from STAs via *Eq. 5* (heavy solid lines). Next, we examine the shape of the STAs themselves (Fig. 8, *right*).

For each operating point we consider, the type II STA ($g_A = 0$) has a pronounced negative lobe. Functionally, this corresponds to a “differentiating” mode through which inputs are processed: negative currents sufficiently far in the past tend to drive more vigorous spiking. Biophysically, this corresponds to the kinetics of ionic currents, for example, inward currents that can be de-inactivated through hyperpolarization. The differentiating filtering property with respect to the total input to a cell has been found before for neurons with type II excitability (e.g., Àguera y Arcas et al. 2003; Prescott et al. 2008); here, we show that this persists for common-input STAs in the presence of background (independent) input, as for correlation transfer of a weak common input to pairs of cells.

By contrast, type I common-input STAs show a less prominent negative lobe or none at all. The resulting filtering of inputs is characterized as “integrating”: a purely positive filter is applied to past inputs to determine firing rate (cf. Dayan and Abbott 2001; Àguera y Arcas et al. 2003; Mato and Samengo 2008; Prescott et al. 2008).

The consequences for spike cross-covariance functions are straightforward. Trends are most pronounced for the superthreshold, high $\sigma$ case (Fig. 8A). Here, the pronounced negative lobe in the type II STA ($g_A = 0$) leads to a similar negative lobe in the crosscovariance, and hence a sharp decrease, following an initial increase, of spike count covariance as a function of time window $T$. The type I STA ($g_A = 60$) is positive, leading to a spike count covariance that steadily increases until it overtakes the type II value at $T \approx 20$ ms. These trends are also reflected in the spike count correlation $r_T$, as described previously. Moreover, analogous plots for superthreshold, high $\sigma$ points where either $I$ spike-count variability (Fano factor; Fig. 9A), or $2$ both variability and firing rate (Fig. 9B) are maintained across $g_A$ values, show the same trends.

In the low $\sigma$ case (Fig. 8B), the STAs have similar characteristics; moreover, there is significant ringing in the STA. This
occurs at the characteristic firing frequency and reflects the cell’s oscillatory spiking: inputs that occur more than one period into the past affect the timing of multiple spikes in the future. By the time the predicted (and actual) covariances reach a limiting value, they are close to zero, and possibly too variable to order definitively. It appears that covariance is still larger for type II than for type I at $T = 200$ ms. In the subthreshold case (Fig. 8C), the STA for the type I neurons is very small, consistent with the very low firing rate here. The type II neuron shows a more robust response, similar in magnitude but less oscillatory than for the superthreshold regime.

**Trends in spike-generating dynamics mirror trends in STAs and transferred correlations.** The transition from type II to type I spike generation in the Connor-Stevens model, as manifest in the progression from discontinuous to continuous spike-frequency vs. current curves in Fig. 2, can also be characterized via the type of bifurcation that governs the transition from quiescence to periodic spiking as increasingly strong currents are injected (see Izhikevich 2007; Rinzel and Ermentrout 1998; Guckenheimer and Holmes 1983 for general references; and Rush and Rinzel 1995 and p. 96 of Ermentrout and Terman 2010 for treatment of the Connor-Stevens model specifically).

For $0 < g_A_m < 46$ mS/cm$^2$, the transition occurs via a subcritical Hopf bifurcation, as voltage trajectories jump from a stable rest state to a preexisting stable periodic orbit (limit cycle). This transition is schematized in Fig. 2, top left. As this figure shows, for smaller values of $g_A$ in this range, the frequency of this cycle is high ($\gtrsim 60$ Hz). The voltage-conductance dynamics near both stable structures, the stable rest state and the limit cycle, is oscillatory. This creates a resonator property (see Izhikevich 2007 and references therein): if they are properly timed, both negative and positive inputs cooperatively produce spikes or cause them to occur earlier than they would in the absence of inputs. This is reflected in the negative and positive lobes in the STA $K(t)$ for the $g_A = 0$ cases (see Fig. 11): recall that the STA is the filter applied to incoming currents to determine firing rates.
By contrast, for large $g_A > 58 \text{ mS/cm}^2$, a saddle-node on invariant circle bifurcation occurs. As sketched in Fig. 2, top right, in this case there is a pair of fixed points that form a “barrier” to spike generation for subthreshold values of $\mu$, and the shadow, or “ghost” of these fixed points still affects dynamics for super-threshold $\mu$, producing slow dynamics in their vicinity. The inputs that will elicit or accelerate spikes are those that will push trajectories past the fixed points, or their ghost, in a distinguished direction. These inputs therefore tend to have a single (positive) sign. This is referred to as an integrator property and gives rise to the mostly positive STAs seen in Fig. 8 for the $g_A = 60$ cases. [This argument breaks down for very low-variance (low $\sigma$) inputs, as we will see in a later section.]

Between these two extremes in $g_A$, the minimum frequency in response to a ramp current decreases steadily, creating a gradual shift between type I and type II behavior. This gradual shift is mirrored in the neural dynamics, in which the slow regions in the state space become increasingly dominant. This transition is clear in the STAs, and therefore spike count covariances, shown in Fig. 8. For example, for $g_A = 30$, we find both distinctly “type II”-like and “type I”-like aspects in the high noise (Fig. 8A) and subthreshold (Fig. 8C) covariance trends, respectively. In the former, spike count covariance increases, then decreases, with $T$; reflecting an oscillating STA; the end result is that the (normalized) covariance at $T = 200 \text{ ms}$ is lower than the covariance at $T = 1 \text{ ms}$. In the latter, the normalized covariance steadily increases with $T$, reflecting a nonnegative STA.

Readout of correlated spiking by downstream cells. How could the difference in spike count correlation between type I and type II cells impact neural circuits? We explore this impact in a simple network, in which correlated type I or type II cells collectively converge to drive a neuron downstream (see Fig. 10A).

In more detail, the drive comes from a population of $N = 200$ identical type I ($g_A = 60 \text{ mS/cm}^2$) or type II ($g_A = 0 \text{ mS/cm}^2$) upstream neurons; we refer to these as population I and population II, respectively. The upstream populations receive correlated inputs with $\epsilon = 0.5$ and values of $\mu$ and $\sigma$ that yield matched levels of variability, as for the parameter set identified with asterisks in Fig. 4 (for population I, $\mu = 18 \mu A/cm^2$ and $\sigma = 5 \mu A\text{-ms}^{1/2}/cm^2$; population II, $\mu = -6 \mu A/cm^2$ and $\sigma = 5 \mu A\text{-ms}^{1/2}/cm^2$). This yields firing rates of $\nu_I = 63.5 \text{ Hz}$ for neurons in population I and $\nu_{II} = 113 \text{ Hz}$ in population II. Each upstream neuron has a single, instantaneous (delta function) synapse onto the downstream neuron of strength $g_I$ or $g_{II}$; the relative size of the excitatory postsynaptic potentials are chosen so that the mean driving current is equal for each population ($\nu_I g_I = \nu_{II} g_{II}$, so that $g_I = 0.825 \text{ mV}$, $g_{II} = 0.5 \text{ mV}$).

The total input received by the downstream neuron, $I_{ds}$, is thus the weighted sum of N upstream spike trains $y_I(t)$:

$$I_{ds}(t) = g_I \sum_j y_I(t) \quad \text{or} \quad I_{ds}(t) = g_{II} \sum_j y_{II}(t).$$  

(8)

When the population size $N$ is large, the summed signal has the same temporal characteristics as the cross-covariance between neuron pairs. Specifically, the autocovariance of the summed input is

$$A_{ds}(\tau) = E \left[ (I_{ds}(t) - \langle I_{ds} \rangle)(I_{ds}(t+\tau) - \langle I_{ds} \rangle) \right]$$

$$= N(N-1) \left[ \langle y_I(t) - \nu_I \rangle \langle y_I(t+\tau) - \nu_I \rangle \right]$$

$$+ N \left[ E \left[ (y_I(t) - \nu_I)(y_I(t+\tau) - \nu_I) \right] \right]$$

$$= N^2 \left[ \langle y_I(t) - \nu_I \rangle \langle y_I(t+\tau) - \nu_I \rangle \right]$$

$$= N^2 C_{y_I^2}(\tau).$$  

(9)

J Neurophysiol • doi:10.1152/jn.00928.2011 • www.jn.org
This relationship is evident in the peristimulus time histograms in Fig. 10B. For population II, fast fluctuations above and below the mean population output reflect the negative lobe in adjacent to its large peak. Meanwhile, fluctuations in the output of population I are less extreme and more gradual in time.

The downstream cell integrates \( I_d(t) \) via leaky integrate-and-fire (LIF) voltage dynamics (Dayan and Abbott 2001):

\[
\frac{dV}{dt} = -(V - V_r) + I_d(t)
\]

where \( \tau_{\text{LIF}} \) is the membrane time constant and \( V_r = -60 \text{ mV} \) is the rest voltage. Spikes are produced when the voltage \( V \) crosses \( V_{\text{thresh}} = -45 \text{ mV} \), at which point \( V \) is reset to \( V_r \).

For the parameters we have chosen, the downstream neuron is driven subthreshold, so that \( \langle I_d(t) \rangle \) is not sufficient to excite a spike; any spikes must be driven by fluctuations in \( I_d(t) \). Thus the variance of fluctuations in \( V(t) \) should give a rough estimate of how often membrane voltage will exceed the threshold and consequently the downstream firing rate. This variance is easy to compute for a passive membrane (i.e., neglecting spike-reset dynamics). First, note that

\[
V(t) = V_r + \int_{-\infty}^{t} I_d(s)L(t - s)ds
\]

where \( L \) is a one-sided exponential filter

\[
L(t) = \frac{1}{\tau} \exp(-t/\tau_{\text{LIF}}), \quad t \geq 0,
\]

\[
= 0, \quad t < 0.
\]

We compute the variance as follows, using the causality of \( L \) to take each upper limit of integration to infinity:

\[
E[V(t)^2] = E\left[ \int_{-\infty}^{t} I_d(s) L(t - s)ds \times \int_{-\infty}^{t} I_d(r) L(t - r)dr \right]
\]

\[
= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} ds dr L(t - s)L(t - r)E[I(s)I(r)]
\]

\[
= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} ds dr L(t - s)L(t - r)A_d(s - r)
\]

\[
= \int_{-\infty}^{\infty} dz A_d(z) \tilde{L}(z) = \tilde{L}(-z)
\]

where \( \tilde{L}(t) = L(-t) \); the last step involved the substitution \( z = s - r \) and switching the order of integration. This final interior integral can be evaluated in the Fourier domain: using
the properties that \( F[\tilde{L}](\omega) = F[L](\omega) \) and the fact that for real functions \( F[f](\omega) = F[Ff](\omega) \), we find
\[
F[(L \ast \tilde{L})](\omega) = |F[L](\omega)|^2 \\
= \frac{1}{1 + \omega^2 \tau_{\text{LIF}}^2}.
\]
Therefore (for example by consulting a transform table)
\[
(L \ast \tilde{L})(t) = \frac{1}{2\tau_{\text{LIF}}} \exp(-|t|/\tau_{\text{LIF}}).
\]
Substituting into Eq. 11, we see that the variance of the downstream cell’s voltage is given by a formula similar to that for the spike count covariances (Eq. 3); both involve integrating the cross-covariance function against a (roughly) triangular-shaped kernel, with time scale \( \tau_{\text{LIF}} \) in the former case and \( T \) in the latter.

Figure 10C shows the by-now familiar trends that this predicts. For short membrane time scales \( \tau_{\text{LIF}} \), type II populations drive greater voltage variance; this is precisely analogous to the finding that spike-count correlations are greater for type II cells over short time scales \( T \). For long \( \tau_{\text{LIF}} \), type I populations drive greater voltage variance, just as type I spike trains are more correlated over long time scales \( T \). In Fig. 10D, we compare this trend with actual firing rates elicited in the downstream cell (from numerical simulation). The general trends match, validating our simple prediction.

In sum, downstream neurons with short membrane time scales (\( \tau_{\text{LIF}} \leq 5 \) ms) are preferentially driven by type II cells upstream; for longer time scales, the preference shifts to type I cells. The effects are substantial. For example, for a downstream neuron with time constant \( \tau_{\text{LIF}} = 4 \) ms, the type II population elicits firing rates that are \( \approx 20\% \) larger than for the type I population; for \( \tau_{\text{LIF}} = 10 \) ms the trend reverses, with the type II populations producing firing rates that are doubled compared with the type I. Some implications are noted in the DISCUSSION.

We note that this result is not limited to the particular choices of operating point (i.e., \( \mu \) and \( \sigma \)) for the upstream populations in Fig. 10, A–D. For example, Fig. 10E demonstrates an analogous finding for upstream type I and type II cell populations with lower values of the DC input \( \mu \) and hence greater variability in single-cell spiking (specifically, matched Fano factors of 0.5). Here, we used the same values of \( g_A \) and \( \sigma \) as above, but for population I, took \( \mu = 13.7 \) \( \mu \text{A/cm}^2 \) (barely superthreshold); for population II, \( \mu = -12.45 \) \( \mu \text{A/cm}^2 \) (subthreshold).

Phase-response curves (PRCs) predict common-input STAs. Finally, we focus on superthreshold operating points, and show how the key properties of type I vs. type II spike generation that determine the filtering of common inputs in this regime can be understood via a commonly used and analytically tractable phase model for tonically spiking neurons. This provides a connection to previous results on correlation transfer (see below, and DISCUSSION).

The response of phase neurons to an additional small-amplitude current \( I(t) \) can be described by a phase model, a one-dimensional description which keeps track only of the progress of neuron along its periodic spiking orbit (or limit cycle). Identifying progress along the cycle with a phase \( \theta \in [0, 2\pi) \), this model is completely determined by a single function of phase \( Z(\theta) \), called a phase response curve or PRC (Ermentrout and Kopell 1984; Winfree 2001; Ermentrout and Terman 2010; Reyes and Fetz 1993):
\[
\frac{d\theta}{dt} = \omega + Z(\theta)I(t).
\]
We can interpret the meaning of this function by considering its effects on the timing of the next spike delivered at a particular phase of the limit cycle \( \varphi \). If \( Z(\varphi) > 0 \), then a positive input delivered at that particular phase will push the neuron further along, advancing the time of the next spike; if \( Z(\varphi) < 0 \), the same input would delay the time of the next spike.

Neurons that display type I spiking have a purely positive (or type I) PRC, while type II neurons show a PRC that has both positive and negative lobes (Ermentrout and Kopell 1984; Ermentrout 1996; Hansel et al. 1995; Brown et al. 2004). A purely positive PRC is characteristic of dynamics near a saddle node bifurcation, in which the system lingers near the ghost of its fixed points (as described previously in Trends in spike-generating dynamics); input in a specific direction is needed to force the system away and elicit a spike. A biphasic PRC reflects oscillatory structure in the phase space, in which correctly timed negative and positive inputs can cooperate to elicit a spike (as with a Hopf bifurcation).

Strong relationships between the PRC and the STA have been found for neurons close to the threshold for periodic spiking (i.e., \( \mu = I_{\text{inf}} \), see METHODS). Spike-triggered covariance analysis of both a type I phase model and the Wang-Busaki model show that the dominant linear “feature” (corresponding to the STA) qualitatively resembles the PRC (Mato and Samengo 2008) in the presence of sufficient current noise. In the (type II) Hodgkin-Huxley model, the two dominant “spike-associated” features identified through covariance analysis closely resemble the STA and its derivative; the STA, in turn, closely resembles the PRC (Agüera y Arcas et al. 2003).

In contrast, phase models in the oscillatory regime (far from the excitability threshold) are known to have an STA proportional to the derivative of the PRC (Ermentrout et al. 2007). In the APPENDIX, we generalize this result to the case of Fig. 1, where the relevant signal \( \xi \) is delivered on top of a noisy background (see Eq. 17); STA(t) \( \approx -Z'(t) \).

In Fig. 11, we test the accuracy of these relationships for the superthreshold points considered above. We show results for type I (\( g_A = 60 \), left) and type II neurons (\( g_A = 0 \), left) and compare the STA computed at two different noise levels to the shape of the PRC \( Z(\theta) \) and its (negative) derivative, labeled dPRC \( [-Z'(\theta)] \).

The time output of the STA has been scaled so that one period (\( T \)) maps onto the unit interval; likewise, the PRC is mapped onto the unit interval. At the lower level of noise, we have good correspondence between the STA and the dPRC in both cases. Notably, both type I and type II neurons have biphasic STAs. At high noise levels, while there is not a strong quantitative relationship between the STA and the PRC itself (unlike in the excitable regime explored by Agüera y Arcas et al. 2003), the PRC carries important clues about the qualitative behavior of the STA. The type II neuron retains the biphasic shape reflective of its PRC, while the type I neuron has shifted to a purely positive STA. In sum, by predicting the STA shape, the PRC gives important clues to the linear response (and hence common input transfer) that we observe in Fig. 8.
Type II (data not shown), we eventually recover results consistent with a very large (but probably biologically irrelevant) time window.

In Fig. 11, we show that this gives a close approximation to the asymptotic behavior has not been recovered even at low-noise points, the asymptotic approximation by comparing the limitations of this asymptotic approximation by computing for both low and high noise, for both type I (left) and type II (right) neurons. For simplicity of visualization, each curve has been normalized by its maximum; that is $Z(t) = \max[Z(t)]$, $\text{STA}(t) = \text{STA}(t)/\max[\text{STA}(t)]$, and $-Z(t) = -\max[Z(t)]$. In addition, the time axis has been scaled by the mean period in each case. Top: PRC, showing monophasic and biphasic shape for type I and type II neurons, respectively. Middle: high noise STA; the type I neuron has lost the negative lobe in its STA, while the type II neuron retains a negative component. Bottom: comparison of (dashed line) derivative PRCs with STA for the low noise (solid line) case. Both STAs have negative components.

Finally, we test an alternate result (cf. Barreiro et al. 2010; Abouzeid and Ermentrout 2011) that, in limited cases, relates PRCs to spike count correlations directly. For long $T$ and reasonably small $\sigma$ and $c$

$$\rho \approx c \left[ \frac{1}{2\pi} \int_{0}^{2\pi} Z(x) dx \right]^{2} + O(\sigma^{2}). \tag{13}$$

In Fig. 12, we show that this gives a close approximation to simulation results for $T = 200$ ms in the superthreshold, high-noise case (see Table 1). Moreover, we can gain insight into the limitations of this asymptotic approximation by comparing with the superthreshold, low-noise case. The results of Barreiro et al. (2010) and Abouzeid and Ermentrout (2011) are derived by taking the asymptotic limit $T \to \infty$ before considering $\sigma$ finite but small; in practice, the smaller the noise variance $\sigma$, the longer $T$ must be to see this effect. For our low-noise points, the asymptotic behavior has not been recovered even at $T = 1000$ ms (as may be seen in Fig. 5B). By using a very large (but probably biologically irrelevant) time window (data not shown), we eventually recover results consistent with the asymptotic prediction (Eq. 13).

**DISCUSSION**

Diverging connections, leading to overlapped input shared across multiple neurons, are a ubiquitous feature of neural anatomy. We study the interplay between this connectivity pattern and basic properties of spike generation in creating collective spiking across multiple neurons. We range spike generation over the fundamental categories of type I to type II excitability (Rinzel and Ermentrout 1998; Hodgkin 1948). The transition in excitability is produced by varying the A-current conductance $g_{A}$ within the well-studied Connor-Stevens neuron model.

Our principal finding is that excitability type plays a major role in how shared, i.e., correlated, input currents are transformed into correlated output spikes. Moreover, these differences depend strongly on the time scale $T$ over which correlations are assessed. At short time scales $T$, type II neurons tend to produce relatively stronger spike correlations for comparable input currents (Marella and Ermentrout 2008; Galán et al. 2007). At longer time scales, the opposite is generally true: for a broad range of input currents, type I neurons transfer most of the shared variance in their inputs (~80%) into shared variance in output spikes, while type II neurons transfer less than half (~40%).

We show that these results have direct implications for how downstream neurons with different membrane time constants will respond to type I vs. type II populations. Specifically, downstream neurons preferentially respond to populations that are strongly correlated on time scales similar to their membrane time constant. Interestingly, for the case we study, we find that the breakpoint between selectivity to type I vs. type II populations was for downstream membrane time constants of ~5 ms, easily within the ranges found experimentally.

This raises interesting possibilities for neuromodulation. The membrane time constant of the downstream cell could be changed by shunting effects of additional background inputs, leading to a switch in its sensitivity to different upstream populations. Alternatively, modulators applied to the upstream populations themselves could change their excitability from type I to type II (Stiefel et al. 2008, 2009), adjusting their impact on a downstream cell with a fixed membrane time constant.

Overall, we demonstrate and apply a general principle: the presence and balance among different membrane currents controls not only single-cell dynamics but also the strength and time scales of spike correlations in cell groups receiving common inputs. We show how this relationship can be understood. As a membrane current (here, $g_{A}$) is adjusted, the firing rate-current curves progressively transition (here, from type I to type II). At the same time, there is a transition in periodic orbit types that neural trajectories visit (here, ranging from orbits “near” a fixed point to relatively “isolated” orbits; Rush and Rinzel 1995). In turn, this produces a steady progression of STAs and hence the filters that neurons apply to shared input signals (here, from primarily integrating to primarily differentiating).
tering modes; Mato and Samengo 2008; cf. Agüera y Arcas et al. 2003). Basic formulas can then be used to translate these filtering properties into predictions for correlated spiking in neural pairs and populations (Ostojic et al. 2009) as well as the downstream impact of this cooperative activity. We anticipate that this approach will bear fruit in studies of the collective activity of a wide variety of neuron types.

**Relationship with prior work.** A number of prior studies have considered the problem of how spike-generating dynamics affect the transfer of incoming current correlations into outgoing spike correlations (Binder and Powers 2001; de la Rocha et al. 2007; Shea-Brown et al. 2008; Rosenbaum and Josić 2011; Tchumatchenko et al. 2010; Marella and Ermentrout 2008; Vilela and Lindner 2009; Barreiro et al. 2010; Ostojic et al. 2009; Tchumatchenko et al. 2010; Hong and De Schutter 2008). In particular, the studies (de la Rocha et al. 2007; Shea-Brown et al. 2008; Rosenbaum and Josić 2011) show that LIF neurons can transfer up to 100% of current correlations into spike count correlations. The level transferred increases with the firing rate at which single neurons are operating and the time scale $T$. These findings are simpler to state compared with the present results for conductance-based neuron models, for which 100% correlation transfer is never obtained, and trends with $T$ differ depending on $g_A$.

Other works (Hong and De Schutter 2008; Shea-Brown et al. 2008; Vilela and Lindner 2009) investigate correlation transfer in more complex spiking models. In particular, Shea-Brown et al. (2008) and Vilela and Lindner (2009) explore the full parameter space of input currents for the quadratic integrate-and-fire model, arguably that with the next level of complexity beyond the LIF model. These authors find similar trends in correlation transfer as the neurons’ operating points change but a limitation to 66% rather than 100% in correlation transfer. Meanwhile, Hong and De Schutter (2008) show complex dependencies on neural operating point for the Hodgkin-Huxley model. Taken together, these studies suggested that correlation transfer depends on spike-generating dynamics in rich and diverse ways.

This opened the door to a broader study, but exploring correlation transfer for the full space of possible spike-generating dynamics in neural models is a daunting task. The axis that spans from type I to type II excitability provides a natural focus. This has been explored using sinusoidal “normal form,” phase-reduced models (Marella and Ermentrout 2008; Galán et al. 2007; Barreiro et al. 2010; Abouzeid and Ermentrout 2011). These studies used simulations in the superthreshold regime, together with analysis in the limits of very short or very long time scales $T$, to show the same trend in correlation transfer over short vs. long $T$ that we find here for conductance-based models. A greater frequency of instantaneous (small $T$) spikes for type II vs. type I neurons was predicted using these simplified models (Marella and Ermentrout 2008; Galán et al. 2007); later, Barreiro et al. (2010) predicted the switch in relative correlation transfer efficiency from type II to type I models as $T$ increases.

The present study confirms the resulting predictions for the superthreshold, oscillatory firing regime using biophysical, conductance-based models. Here, we also explore correlation transfer for subthreshold, fluctuation-driven firing. For both cases, we explain the origins of variable correlation transfer via filtering properties of type I vs. type II cells, and demonstrate the impact on downstream neurons. The very recent study by Hong et al. (2012) uses the related (but not identical) characterization of cells as “integrators” vs. “coincidence detectors” and shows how the measures of synchrony and firing rate correlation differ for each model and depend on the mean input current they receive.

**Scope, limitations, and open questions.** The circuit model that we have studied, as illustrated in Fig. 1, is limited to a single, idealized feature of feedforward connectivity: overlapping inputs to multiple recipient cells. More realistic architecture could include delays in incoming inhibitory vs. excitatory inputs (Gabet et al. 2005). Interactions of shared-input circuitry with recurrent connectivity also pose important questions (Ly and Ermentrout 2009). This is especially so given the distinct properties of type I vs. type II cells in synchronization due to reciprocal coupling (Rinzel and Ermentrout 1998; Ermentrout and Terman 2010).

Other aspects of our biophysical and circuit dynamics are also idealized. For one, individual input currents fluctuated on arbitrarily fast time scales (i.e., as white noise processes). Relaxing this would be an interesting extension. While prior studies (de la Rocha et al. 2007) suggest that trends will persist for inputs with fast (but finite) time scales, new effects could arise for slower-time scale inputs representative of slower synapses or even network-level oscillations. Another addition would be for inputs to arrive via excitatory and inhibitory conductances, rather than currents; while previous studies with integrate-and-fire cells (de la Rocha et al. 2007) have found that this yields qualitatively similar results, there could be interesting interactions with underlying filtering properties in biophysical models. The same holds true for inputs that arrive at dendrites in multicompartment models. Finally, our focus on type I vs. type II dynamics captures some, but not all, relevant dynamical features: for the related coincidence detector models of (Hong et al. 2012), accurate predictions require second-order terms from the spike-triggered covariance.

Likewise, the circuitry of Fig. 10A that we used to investigate the impact of correlated spiking on downstream neurons was highly idealized. An especially appealing extension would be to note that inhibitory and excitatory neurons often have different excitability types. Thus downstream cells could receive input from both excitatory type I and inhibitory type II populations. Our results suggest that sensitivity to excitatory vs. inhibitory afferents would vary with membrane time constants downstream, possibly amplifying the modulatory effects identified here.

**APPENDIX: RELATING STAS AND SPIKE-GENERATING DYNAMICS IN THE PRESENCE OF COMMON AND INDEPENDENT NOISE**

To relate the common input STA to spike-generating dynamics, we extend a result in the literature to derive an explicit formula for the common input STA of a phase model, which captures the response of a tonically spiking neuron to a small-amplitude current $I(t)$. We emphasize that the resulting formula, and the calculation that yields it, are very similar to a relationship previously derived (Ermentrout et al. 2007) for the STA of a phase oscillator without background noise.

We consider a model that tracks only the phase of a neuron as it progresses along its periodic spiking orbit:
The function $Z(\theta)$, called a PRC (Ermentrout and Kopell 1984; Winfree 2001; Ermentrout and Terman 2010; Reyes and Fetz 1993), determines how a brief current injection applied at a specific phase of the cycle affects the timing of the next spike. By convention, the neuron is said to “spike” when $\theta$ crosses $2\pi$.

To begin, we assume that the phase model is forced by scaled zero-mean, stationary stochastic processes, which we also label $\xi(\cdot)$ and $\xi(\cdot)$. For now, $\xi(\cdot)$ and $\xi(\cdot)$ have unit variance and are differentiable with some finite correlation time $\tau$, although we will consider the limit $\tau \to 0$ (i.e., the white noise limit). We are interested in the average value of $\xi(\cdot)$ that precedes a spike; the term $\xi(\cdot)$ will play the role of a background noise. Assuming that the background noise process is scaled by a small constant $\sigma$, and that $\xi(\cdot)$ is scaled by an order of magnitude $c$ smaller still, we write the evolution equation of the phase model as

$$\frac{d\theta}{dt} = 1 + Z(\theta)[\sigma \xi(t) + \sigma \xi(t)], \quad \theta \in [0, T)$$

where $\sigma, c << 1$. Note that we have chosen our phase variable to have unit speed; i.e $\theta \in [0, T]$, where $T$ is the period of the unperturbed ($\sigma = 0$) oscillator. We proceed as in Ermentrout et al. (2007): writing $\theta$ as a series in the small parameters $\sigma$ and $c$

$$\theta(t) = \theta_0(t) + \sigma \theta_{10}(t) + \epsilon \theta_{01}(t) + \sigma^2 \theta_{20}(t) + \sigma \epsilon \theta_{11}(t) + \ldots$$

and matching terms of same order in the evolution equation, we find $\theta_0(t) = t$. We additionally find that $\theta_{01} = 0$, $\theta_{02} = 0$, and

$$\theta_{10}' = Z(t)\xi(t)$$
$$\theta_{20}' = Z(t)\xi(t)\theta_{10}(t)$$
$$\theta_{11}' = Z(t)\xi(t)$$

so that

$$\theta_{10}(t) = \int_0^T Z(s)\xi(s)ds$$
$$\theta_{20}(t) = \int_0^T Z(s)\xi(s)\int_0^T Z(r)\xi(r)dr ds$$
$$\theta_{11}(t) = \int_0^T Z(s)\xi(s)ds.$$

To compute the STA, we need to find the time of the next spike, assuming the neuron has just spiked [$\theta(0) = 0$]; in other words, the time $\tau$ when $\theta(\tau) = T$. As above, we expand

$$\tau = T + \sigma \tau_{10} + \sigma^2 \tau_{20} + \sigma \epsilon \tau_{11} + \ldots$$

Using our previous expressions for $\theta(\tau)$, and using the fact that $\tau = T + \sigma \tau_{10} + O(\sigma^2, \sigma \epsilon)$ to decompose the stochastic integrals, we find

$$\tau_{10} = -\int_T^T Z(s)\xi(s)ds$$
$$\tau_{20} = -\int_T^T Z(s)\xi(s)ds - \int_T^T Z'(s)\xi(s)\int_0^T Z(r)\xi(r)dr ds$$
$$\tau_{11} = -\int_T^T Z(s)\xi(s)ds.$$

Next, we use Taylor’s theorem for smooth functions to expand $\xi(\cdot)$ about $T - t$ to compute

$$\text{STA}(t) = \left< I_c(T - t) \right> = \sigma \epsilon \left< \xi(\cdot)(T + \sigma \tau + \sigma^2 \tau_{20} + \sigma \epsilon \tau_{11} - t) \right>$$
$$= \sigma \epsilon \left< \xi(\cdot)(T - t) + (\sigma \tau + \sigma^2 \tau_{20} + \sigma \epsilon \tau_{11}) \xi(\cdot)(T - t) \right>$$
$$+ \frac{1}{2} (\sigma \tau_{10} + \sigma^2 \tau_{20} + \sigma \epsilon \tau_{11})^2 \xi(\cdot)(T - t)$$
$$= \sigma \epsilon \left< \xi(\cdot)(T - t) + (\sigma \tau_{10} + \sigma^2 \tau_{20} + \sigma \epsilon \tau_{11}) \xi(\cdot)(T - t) \right>$$
$$+ \frac{1}{2} \sigma^2 \tau_{10} \xi(\cdot)(T - t) + O(\sigma^3, \sigma^2 \epsilon)$$

where we have kept terms up to second order both in our expression for $\tau$, and in our Taylor expansion of $\xi(\cdot)$. We can use the independence of $\xi(\cdot)$ and $\xi(\cdot)$ to eliminate a large number of terms, as

$$\left< \xi(\cdot)(\xi(\cdot) + s) \right> = \left< \xi(\cdot) \right> \left< \xi(\cdot) + s \right> = 0.$$

Similarly,

$$\left< \xi(\cdot)(\xi(\cdot) + s) \right> = \left< \xi(\cdot) \right> \left< \xi(\cdot) + s \right> = 0,$$

and so forth for expressions with higher derivatives. The only term that survives is

$$\text{STA}(t) = \sigma \epsilon \left< \xi(\cdot)(T - t) \right> - \sigma \epsilon \int_0^T Z(s)\xi(s)ds$$
$$= -\left( \sigma \epsilon \right)^2 \int_0^T Z(s)\xi(\cdot)(T - t)\xi(s)ds$$
$$= -\left( \sigma \epsilon \right)^2 \int_0^T Z(s)A_c(T - t - s)ds$$
$$= -\left( \sigma \epsilon \right)^2 \int_0^T Z(s)A_c(T - s)ds,$$

where $A_c$ is the autocovariance function of $\xi(\cdot)$ and we used integration by parts in the final step. Taking the white noise limit [$A_c(T - t - s) \to 8\delta(t - s)$] and using the periodicity of the PRC [$Z(T - t) = Z(-t)$], we recover a very similar expression as Ermentrout et al. (2007):

$$\text{STA}(t) = -\left( \sigma \epsilon \right)^2 Z(-t).$$

ACKNOWLEDGMENTS

We thank Kresimir Josic and Brent Doiron for helpful comments as this work progressed.
Present address of A. K. Barreiro: Dept. of Mathematics, Southern Methodist Univ., Dallas, TX.

GRANTS

This research was supported by National Science Foundation (NSF) Grants DMS-0817649 and CAREER DMS-1056125 and by a Career Award at the Scientific Interface from the Burroughs-Wellcome Fund (to E. Shea-Brown), by the Mary Gates Foundation at the University of Washington (to E. T. Thilo), and by NSF Teragrid allocation TG-IBN090004.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).
CORRELATION TRANSFER AND NEURAL EXCITABILITY

AUTHOR CONTRIBUTIONS

Author contributions: A.B., E.T, and E. S-B. each contributed to computation, coding, and writing; E.T. spearheaded large-scale distributed simulation; A.B. and E. S-B. conducted mathematical analysis and study design.

REFERENCES


