INTRODUCTION

Cortical neurons integrate inputs from thousands of afferents. Similarly, a variety of experimental techniques record the pooled activity of large populations of cells. It is therefore important to understand how the structured response of a neuronal network is reflected in the pooled activity of cell groups.

It is known that weak dependencies between the response of cell pairs in a population can have a significant impact on the variability and signal-to-noise ratio of the pooled signal (Shadlen and Newsome, 1998; Salinas and Sejnowski, 2000; Moreno-Bote et al., 2008). It has also been observed that weak correlations between cells in two populations can cause much stronger correlations between the pooled activity of the populations (Bedenbaugh and Gerstein, 1997; Chen et al., 2006; Gutnisky and Josić, 2010; Renart et al., 2010). We give a simple example of this effect in Figure 1C: Weak correlations were introduced between the spiking activity of cells in two non-overlapping presynaptic pools each providing input to a postsynaptic cell (see diagram in Figure 1B). The activity between pairs of excitatory, and pairs of inhibitory cells was correlated, but excitatory–inhibitory pairs were uncorrelated. Even without shared inputs and with background noise, pooling resulted in strong correlations in postsynaptic membrane voltages. The connectivity in the presynaptic network was irrelevant – it only mattered that the inputs to the downstream neurons reflected the pooled activity of the afferent populations. A similar effect can cause large correlations between recordings of multiunit activity (MUA) or recordings of voltage sensitive dyes (VSD), even when correlations between cells in the recorded populations are small (Bedenbaugh and Gerstein, 1997; Chen et al., 2006; Stark et al., 2008). The effect is the same, but in this case pooling occurs at the level of a recording device rather than a downstream neuron (compare Figures 1A,B).

We present a systematic overview, as well as extensions and applications of a number of previous observations related to this phenomenon. Using a linear model, we start by examining the potential effects of pooling on recordings from large populations obtained using VSD or MUA recording techniques. These techniques are believed to reflect the pooled postsynaptic activity of groups of cells. We extend earlier models introduced to examine the impact of pooling on correlations (Bedenbaugh and Gerstein, 1997; Chen et al., 2006; Nunez and Srinivasan, 2006), and show that heterogeneities in the presynaptic pools can have subtle effects on correlations between pooled signals.

Since neurons respond to input from large presynaptic populations, pooling also impacts the activity of single cells and cell pairs. As observed in Figure 1C, pooling can inflate weak correlations between afferents. However, excitatory–inhibitory correlations (Okun and Lampl, 2008) can counteract this amplification, as shown in Figure 1D (Hertz, 2010; Renart et al., 2010). We examine these effects analytically by modeling the subthreshold activity of postsynaptic cells as a filtered version of the inputs received (Tetzlaff et al., 2008). The impact of correlated subthreshold activity on the output spiking statistics is a nontrivial question which we address only briefly (Moreno-Bote and Parga, 2006; de la Rocha et al., 2007; Ostojic et al., 2009).

The effects of pooling provide a simple explanation for certain aspects of the dynamics of feedforward chains. Simulations and in vitro experiments show that layered feedforward architectures give rise to a robust increase in synchronous spiking from layer to layer (Diesmann et al., 1999; Litvak et al., 2003; Reyes, 2003; Doiron et al., 2006; Kumar et al., 2008). We describe how output correlations in one layer impact correlations between the pooled inputs to the next layer. This approach is used to derive a mapping that describes how correlations develop across layers (Tetzlaff et al., 2003; Renart et al., 2010), and to illustrate that the pooling of correlated inputs is the primary mechanism responsible for the development of synchrony in feedforward chains. Examining how correlations are mapped between layers also helps explain why asynchronous states are rarely observed in feedforward networks in the absence of strong background noise (van Rossum et al., 2002;
Vogels and Abbott, 2005). This is in contrast to recurrent networks which can display stable asynchronous states (Hertz, 2010; Renart et al., 2010) similar to those observed in vivo (Ecker et al., 2010).

**MATERIALS AND METHODS**

**CORRELATIONS BETWEEN STOCHASTIC PROCESSES**

The cross-covariance of a pair of stationary stochastic processes, \( x(t) \) and \( y(t) \), is \( C_{xy}(t) = \text{cov}(x(s), y(s + t)) \). The auto-covariance function, \( C_{xx}(t) \), is the cross-covariance between a process and itself. The cross- and auto-covariance functions measure second order dependencies at time lag \( t \) between two processes, or a process and itself. We quantify the total magnitude of interactions over all time using the asymptotic statistics,

\[
\gamma_{xy} = \int C_{xy}(t)dt, \quad \sigma_{x}^2 = \sigma_{xx}, \quad \rho_{xy} = \gamma_{xy} / \sigma_{x} \sigma_{y}.
\]

While the asymptotic correlation, \( \rho_{xy} \), measures correlations between \( x(t) \) and \( y(t) \) over large timescales, the auto- and cross-covariance functions determine the timescale of these dependencies.

**CORRELATIONS BETWEEN SUMS OF RANDOM VARIABLES**

Given two collections of correlated random variables \( \{x_i\}_{i=1}^{n} \) and \( \{y_j\}_{j=1}^{m} \), define the pooled variables, \( X = \sum x_i \) and \( Y = \sum y_j \). Since covariance is bilinear (\( \text{cov}(\sum x_i, \sum y_j) = \sum \text{cov}(x_i, y_j) \)), the variance and covariance of the pooled variables are

\[
\sigma_X^2 = \sum_{i=1}^{n} \sum_{j=1}^{m} \sigma_{x_i} \sigma_{y_j} \rho_{x_i y_j}, \quad \text{and} \quad \gamma_{XY} = \sum_{i=1}^{n} \sum_{j=1}^{m} \sigma_{x_i} \sigma_{y_j} \rho_{x_i y_j},
\]

and similarly for \( \sigma_Y^2 \).

Using these expressions along with some algebraic manipulation, the correlation coefficient, \( \rho_{XY} = \gamma_{XY} / \sqrt{\sigma_X \sigma_Y} \), between the pooled variables can be written as

\[
\rho_{XY} = \frac{\overline{\rho_{xy}}}{\sqrt{\overline{\sigma_{xx}} \overline{\sigma_{yy}}}},
\]

where

\[
w_{x} = \frac{\sigma_{x} \sigma_{y}}{\sigma_{x} \sigma_{y}}, \quad \overline{\rho_{xy}} = \frac{1}{n_{x}} \sum_{i=1}^{n} \sigma_{x_i} \sigma_{y_j} \rho_{x_i y_j}, \quad \overline{\sigma_{xx}} = \frac{1}{n_{x}(n_{x} - 1)} \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{x_i} \sigma_{x_j}, \quad \overline{\sigma_{yy}} = \frac{1}{n_{y}(n_{y} - 1)} \sum_{i=1}^{m} \sum_{j=1}^{m} \sigma_{y_i} \sigma_{y_j},
\]

and similarly for \( w_{y}, \overline{\rho_{xy}}, \overline{\sigma_{xy}}, \overline{\sigma_{yy}}, \overline{\sigma_{xx}} \). In deriving Eq. (3) we assumed that all pairwise statistics are uniformly bounded away from zero in the asymptotic limit.

Each overlined term above is a population average. Notably, \( \overline{\rho_{xy}} \) represents the average correlation between \( x \) and \( y \), pairs, weighted by the product of their standard deviations, and similarly for \( \overline{\rho_{xx}} \) and \( \overline{\rho_{yy}} \). Correlation between weighted sums can be obtained by substituting \( \sum x_i \rightarrow w_x \), and \( \sum y_j \rightarrow w_y \), for weights \( w_x \) and \( w_y \), and making the appropriate changes to the terms in the equation above (e.g., \( \sigma_{x_i} \rightarrow \sqrt{w_x} \sigma_{x_i}, \rho_{x_i y_j} \rightarrow \text{sign}(w_x w_y) \rho_{x_i y_j} \)). Overlap between the two populations can be modeled by taking \( \rho_{x_i y_j} = 1 \) for some pairs.

Assuming that variances are homogeneous within each population, that is \( \sigma_{x_i} = \sigma_{x} \) and \( \sigma_{y_j} = \sigma_{y} \) for \( i = 1, ..., n \) and \( j = 1, ..., m \), simplifies these expressions. In particular, \( \overline{\rho_{xx}} = \sigma_{x} \sigma_{x}, \overline{\rho_{yy}} = \sigma_{y} \sigma_{y}, \overline{\sigma_{xy}} = \sigma_{x} \sigma_{y} \), and

\[
\rho_{XY} = \frac{\overline{\rho_{xy}}}{\sqrt{\overline{\sigma_{xx}} \overline{\sigma_{yy}}}}.
\]

Assuming further that the populations are symmetric, \( \sigma_{x} = \sigma_{y} \), \( n_x = n_y = n \), and \( \overline{\rho_{xx}} = \overline{\rho_{yy}} \), the expression above simplifies to

\[
\rho_{XY} = \frac{\rho^x}{\rho^x + 1 - \rho^y} = \frac{\rho^x}{\rho^x + O\left( \frac{1}{n} \right)},
\]

where \( \rho^x = \overline{\rho_{xx}} \) is the average pairwise correlation between the two populations and \( \rho^y = \overline{\rho_{yy}} \) is the average pairwise correlation within each population. Eq. (5) was derived in Bedenbaugh and Gerstein (1997) in an examination of correlations between multielectrode recordings. In Chen et al. (2006), a version of Eq. (5) with \( \rho^y = \rho^x \) is derived in the context of correlations between two VSD signals. The asymptotic, \( \rho_{xy} \rightarrow 0 \), limit when \( \rho^y = \rho^x \) is discussed in Renart et al. (2010).

Note that the results above hold for correlations computed over arbitrary time windows. We concentrate on infinite windows, and discuss extensions in the Appendix.

**NEURON MODEL**

In the second part of the presentation, we consider two excitatory and two inhibitory input populations projecting to two postsynaptic cells. The \( j^{th} \) excitatory input to cell \( k \) is labeled \( e_{jk}(t) \) (\( k = 1 \) or 2). Similarly, \( i_{jk}(t) \) denotes the \( j^{th} \) inhibitory input to cell \( k \). Each cell receives \( n_e \) excitatory and \( n_i \) inhibitory inputs with individual rates \( \nu_e \) and \( \nu_i \), respectively.

Each of the excitatory and inhibitory inputs to cell \( k \) are stationary spike trains modeled by point processes, \( e_{jk}(t) = \sum \delta(t - t'_{ej}) \) and \( i_{jk}(t) = \sum \delta(t - s'_{ij}) \) where \( t'_{ej} \) and \( s'_{ij} \) are input spike times. We assume that the spike trains are stationary in a multivariate sense (Strattonovich, 1963). The pooled excitatory and inhibitory inputs to neuron \( k \) are \( E_k(t) = \sum_{j=1}^{n_e} e_{jk}(t) \), and \( I_k(t) = \sum_{j=1}^{n_i} i_{jk}(t) \).

To generate correlated inputs to cells, we used the multiple interaction process (MIP) method (Kuhn et al., 2003), then jittered each spike time independently by a random value drawn from an exponential distribution with mean 5 ms. The resulting processes are...
Poisson with cross-covariance functions proportional to a double exponential, \(C_{12}(t) \sim e^{-|t|/\tau_{12}}\). Note that since each input is Poisson, \(\sigma_i^2 = \nu_i\) and \(\sigma_j^2 = \nu_j\).

While the dynamics of the afferent population were not modeled explicitly, the response of the two downstream neurons was obtained using a conductance-based IF model. The membrane potentials of the neurons were described by

\[
C_m \frac{dV_k}{dt} = -g_{k}(V_k - V_l) - g_{k}(t)(V_k - V_e) - g_{k}(t)(V_k - V_i),
\]

with excitatory and inhibitory conductances determined by \(g_k(t) = (E_k + \alpha_k)(t)\) and \(g_k(t) = (I_k + \alpha_k)(t)\) where \(\ast\) denotes convolution. We used synaptic responses of the form \(\alpha_k(t) = \tau_{\alpha_k} e^{-t/\tau_{\alpha_k}} \Theta(t)\) and \(\sigma_k(t) = \tau_{\sigma_k} e^{-t/\tau_{\sigma_k}} \Theta(t)\) where \(\Theta(t)\) is the Heaviside function. The area of a single excitatory or inhibitory postsynaptic conductance (EPSC or IPSC) is therefore equal to the synaptic weight, \(E\) or \(I\), with units nS-ns. This analysis can easily be extended to situations where each input, \(e_{i,a}\) or \(i_{i,a}\), has a distinct synaptic weight.

When examining spiking activity, we assume that when \(V_k\) crosses a threshold voltage, \(V^\psi\), an output spike is produced and \(V_k\) is reset to \(V_l\). When examining sub-threshold dynamics, we considered the free membrane potential without threshold.

As a measure of balance between excitation and inhibition we used (Troyer and Miller, 1997; Salinas and Sejnowski, 2000)

\[
\beta = \frac{V_{l} - V_{e}}{V_{l} - V_{i}}.
\]

When \(\beta = 1\), the net excitation and inhibition are balanced and the mean free membrane potential equals \(V_l\). In simulations, we set \(V_e = -60\) mV, \(V_i = 0\) mV, \(V_l = -90\) mV, \(\tau_e = 10\) ms, \(\tau_i = 20\) ms, \(C_m = 114\) pF, and \(g_e = 4.086\) nS, giving a membrane time constant, \(\tau_m = C_m/g_e = 27.9\) ms. In all simulations except those in Figure 7, the cells are balanced (\(\beta = 1\)).

The conductance-based IF neuron behaves as a nonlinear filter in the sense that membrane potentials cannot be written as a linear transformation of the inputs. However, following Kuhn et al. (2004) and Coombes et al. (2007), we derive a linear approximation to the conductance based model. Let \(U = V_i - V_j\) so that Eq. (6) becomes

\[
C_m \frac{dU}{dt} = (-g_e(t) - g_i(t))(U - g_e(t))(V_i - V_e) - g_i(t)(V_i - V_i).
\]

Define the effective membrane time constant, \(\tau_{\alpha \alpha} = C_m / (E[g_e + g_i(t)] + g_i(t))]\), and \(\tau_{\sigma \sigma} = C_m / (E[g_e + I_j] + \nu_j E + I_j)\). Substituting this average value in the previous equation yields the linear approximation to the conductance based model,

\[
\frac{dU}{dt} = -\frac{1}{\tau_{\alpha \alpha}} U + J_k(t),
\]

where \(J_k(t) = (-g_e(t)(V_i - V_e) - g_i(t)(V_e - V_i))/C_m\) is the total input current to cell \(k\). Solving and reverting to the original variables gives the linear approximation \(V_k(t) = (J_k(t) + K(t))/\Theta(t)\), where \(K(t) = \Theta(t)e^{-t/\tau_{\alpha \alpha}}\) is the kernel of the linear filter induced by Eq. (7).

**RESULTS**

The pooling of signals from groups of neurons can impact both recordings of population activity and the structure of inputs to postsynaptic cells. We start by discussing correlations in pooled recordings using a simple linear model. A similar model is then used to examine the impact of pooling on the statistics of inputs to cells. For simplicity we assume that all spike trains are stationary. However, non-stationary results can be obtained using similar methods as outlined in the Section “Discussion.” Though all parameters are defined in the Materials and Methods, Tables 1 and 2 in the Appendix contain brief descriptions of parameters for quick reference. Also, Tables 3 and 4 summarize the values of parameters used for simulations throughout the article.

**CORRELATIONS BETWEEN POOLED RECORDINGS**

Pooling can impact correlations between recordings of population activity obtained from voltage sensitive dyes (VSDs), multi-unit recordings and other techniques. Such signals might each represent the summed activity of hundreds or thousands of neurons. Let two recorded signals, \(X_i(t)\) and \(X_j(t)\), represent the weighted activity of cells in two populations (see diagram in Figure 1A). If we assume homogeneity in the input variances and equal size of the recorded populations, Eq. (4) gives the correlation between the recorded signals

\[
\rho_{X_i X_j} = \frac{\bar{p}_{11}}{\sqrt{\bar{p}_{11} + \frac{1}{n}(1 - \bar{p}_{11})}} \frac{\bar{p}_{22} + \frac{1}{n}(1 - \bar{p}_{22})}{\sqrt{\bar{p}_{22} + \frac{1}{n}(1 - \bar{p}_{22})}} = \frac{\bar{p}_{22}}{\sqrt{\bar{p}_{11} \bar{p}_{22}}} + O\left(\frac{1}{n}\right).
\]

Here \(n\) represents the number of neurons recorded, \(\bar{p}_{11}, k = 1, 2\) represents the average correlation between cells contributing to signal \(X_i(t)\), and \(\bar{p}_{12}\) represents the average correlation between cells contributing to different signals. The averages are weighted so that cells that contribute more strongly to the recording, such as those closer to the recording site, contribute more to the average correlations (see Materials and Methods). Cells common to both recorded populations can be modeled by setting the corresponding correlation coefficients to unity. A form of Eq. (8) with \(\bar{p}_{11} = \bar{p}_{22}\) was derived by Bedenbaugh and Gerstein (1997).

When the two recording sites are nearby, so that \(\bar{p}_{12} = \bar{p}_{11} = \bar{p}_{22}\), even small correlations between individual cells are amplified by pooling so that the correlations between the recorded signals can be close to 1. This effect was observed in experiments and explained in similar terms in Stark et al. (2008).

A significant stimulus-dependent change in correlations between individual cells might be reflected only weakly in the correlation between the pooled signals. This can occur, for instance, in recordings of large populations when \(\bar{p}_{12}, \bar{p}_{11}\), and \(\bar{p}_{22}\) are increased by the same factor when a stimulus is presented. Similarly, an increase in correlations between cells can actually lead to a decrease in correlations between recorded signals when \(\bar{p}_{11}\) and \(\bar{p}_{22}\) increase by a larger factor than \(\bar{p}_{12}\).

To illustrate these effects, we construct a simple model of stimulus dependent correlations motivated by the experiments in Chen et al. (2006), in which VSDs were used to record the population response in visual area V1 during an attention task. In their experiments,
the imaged area is divided into 64 pixels, each 0.25 mm × 0.25 mm in size. The signal recorded from each pixel represents the pooled activity of \( n = 1.25 \times 10^4 \) neurons.

We model correlations between the signals, \( X_i(t) \) and \( X_j(t) \), recorded from two pixels in the presence or absence of a stimulus (see Figure 2B), using a simplified model of stimulus dependent rates and correlations. The firing rate of a cell located at distance \( d \) from the center of the retinotopic image of a stimulus is

\[
   r(d) = \begin{cases} 
   B & \text{stimulus present} \\
   B + \frac{(1-B)(1+\cos(d\pi))^\lambda}{2} & \text{stimulus absent.} 
   \end{cases}
\]

(9)

Here, \( B \in [0,1] \) represents baseline activity and \( \lambda \geq 1 \) controls the rate at which activity decays with \( d \). Both \( d \) and \( r \) were scaled so that their maximum value is 1 (see Figure 2A).

We assume that the correlation between the responses of two neurons is proportional to the geometric mean of their firing rates (de la Rocha et al., 2007; Shea-Brown et al., 2008), and that correlations decay exponentially with cell distance (Smith and Kohn, 2008; see however Poort and Roelfsema, 2009; Ecker et al., 2010). We therefore model the correlation between two cells as

\[
   \rho_{k,k'} = S \sqrt{r(d_{k})r(d_{k'})} e^{-\alpha d_{k,k'}}
\]

where \( d_k \) and \( d_{k'} \) are the distances from each cell to the center of the retinotopic image of the stimulus, \( D_s \) is the distance between cells \( j \) and \( k \), \( \alpha \) is the rate at which correlations decay with distance, and \( S \leq 1 \) is a constant of proportionality.

If pixels are small compared to the scales at which correlations are assumed to decay, then the average correlation between cells within the same pixel are \( \bar{\rho}_{ij} = Sr(d_i) \) and \( \bar{\rho}_{ij'} = Sr(d_{i'}) \). The average correlation between cells in different pixels is \( \bar{\rho}_{ij} = S \sqrt{r(d_{i})r(d_{j})} e^{-\alpha d_{ij}} \).

In this case, whether a stimulus is present or not, the correlation between the pooled signals is of the form \( \rho_{XX'} = e^{-\alpha d_{ij}} + \mathcal{O}(1/n) \). Thus, even significant stimulus dependent changes in correlations would be invisible in the recorded signals. This overall trend is consistent with the results in Chen et al. (2006) (compare Figure 2C to their Figure 2f). In such settings, it is difficult to conclude whether pairwise correlations are stimulus dependent or not from the pooled data.
The two cells share $p_{n_e}$ of the inputs drawn from the correlated pool. Processes in the independent pool are uncorrelated with all other processes. All excitatory inputs have variance $\sigma_{ee}^2$.

The correlation between the pooled excitatory inputs is given by (see Appendix)

$$\rho_{ee} = \frac{\rho_{ee} + \rho_{ee}^b(1 - \rho_{ee})}{\rho_{ee} + \frac{1}{n_e}(1 - \rho_{ee} + q_e)}.$$  \hspace{1cm} (10)

A form of this equation, with $\rho_{ee} = 0$ and $q_e = 0$, is derived in Chen et al. (2006). In the absence of correlations between processes in the input pools, $\rho_{ee} = 0$, the correlation between the pooled signals is just the proportion of shared inputs, $\rho_{pEI} = \rho_{ee}$. When $\rho_{ee} > 0$ and $n_e$ is large, pooled excitatory inputs are highly correlated, even when pairwise correlations in the presynaptic pool $\rho_{ee}$ are small, and the neurons do not share inputs ($p_{ee} = 0$). Even when most inputs to the downstream cells are independent ($q_e > 1$), correlations between the pooled signals will be nearly 1 for sufficiently large input pools (see Figure 4A).

Under analogous homogeneity assumptions for the inhibitory pools, the correlation, $\rho_{EI}$, between the pooled inhibitory inputs is given by an equation identical to Eq. (10), and the correlation between the pooled excitatory and inhibitory inputs is given by

$$\rho_{EI} = \frac{\rho_{EI} + \rho_{EI}^b(1 - \rho_{EI})}{\rho_{EI} + \frac{1}{n_e}(1 - \rho_{EI} + q_i)}.$$  \hspace{1cm} (11)

Interestingly, since $|\rho_{EI}|| \leq 1$, pairwise excitatory–inhibitory correlations obey the bound $|\rho_{EI}| \leq \sqrt{|\rho_{ee}\rho_{ii}|} + O(1/\sqrt{n_e n_i})$. Combining this inequality with Eq. (10) and the analogous equation for $\rho_{EI}$, it follows that $|\rho_{EI}| \leq \sqrt{|\rho_{ee}\rho_{ii}|} + O(1/\sqrt{n_e n_i})$ for homogeneous populations. These are a result of the non-negative definiteness of covariance matrices.

**Heterogeneity and the effects of spatially dependent correlations**

We next discuss how heterogeneity can dampen the amplification of correlations due to pooling. In the absence of any homogeneity assumptions on the excitatory input population (see the population model in the Materials and Methods), Eq. (3) gives the pooled excitatory signals, $\mathbf{p}_{EI} = \mathbf{p}_{ee} / \sqrt{\mathbf{p}_{ee} \mathbf{p}_{ii}} + O(1/\sqrt{n_e n_i})$. The term $\mathbf{p}_{ee}$ is a weighted average of the correlation coefficients between the two excitatory populations, and $\mathbf{p}_{ee}$ and $\mathbf{p}_{ii}$ are weighted averages of the correlations within each excitatory input population.

To illuminate this result, we assume symmetry between the populations: Let $n_e = n_i$ and $\sigma_{ee} = \sigma_{ii}$ for $k = 1,2$ and $j = 1,2,\ldots,n_k$, and assume $\mathbf{p}_{ee} = \mathbf{p}_{ii}$. The average “within” and “between” correlations, are $\rho_{ee}^w = \mathbf{p}_{ee} = \mathbf{p}_{ee} = \mathbf{p}_{ee}$ respectively (see Figure 3B). Under these assumptions, Eq. (5) can be applied to obtain (See also Bedenbaugh and Gerstein, 1997)

$$\rho_{EE} = \frac{\rho_{ee}^b + \rho_{ee}^b(1 - \rho_{ee})}{\rho_{ee} + \frac{1}{n_e}(1 - \rho_{ee}^w)} = \frac{\rho_{ee}^b + \frac{1}{n_e}}{\rho_{ee} + \frac{1}{n_e}}.$$  \hspace{1cm} (12)
which is plotted in Figure 4A (green line) and Figure 4B. For large \( n_e \), the correlation between the pooled signals is the ratio of “between” and “within” correlations.

This observation has implications for a situation ubiquitous in the cortex. A neuron is likely to receive afferents from cells that are physically close. The activity of nearby cells may be more strongly correlated than the activity of more distant cells (Chen et al., 2006; Smith and Kohn, 2008). We therefore expect that pairwise correlations within each input pool are on average larger than correlations between two input pools, that is, \( \rho_{ee}^w > \rho_{ee}^b \). This reduces the correlation between the inputs, regardless of the input population size.

An increase in correlations in the presynaptic pool can also correlate the pooled signals. If correlations within each input pool increase by a greater amount than correlations between the two pools, then the variance in the input to each cell will increase by a larger amount than the covariance between the inputs. As a consequence the correlations between the pooled inputs will be reduced. Modulations in correlation have been observed as a consequence of attention in V4 (Cohen and Maunsell, 2009; Mitchell et al., 2009; but apparently not in V1, Roelfsema et al., 2004). Such changes may be, in part, a consequence of small changes in “within” correlations between neurons in V1.

Equation 12 implies that correlations between large populations cannot be significantly larger than the correlations within each population. Since \( |\rho_{ei}b| \leq 1 \), it follows that \( |\rho_{ee}^w| \leq |\rho_{ee}^b| + \mathcal{O}(1/n_e) \).

The correlation, \( \rho_{ii,b} \), between the pooled inhibitory inputs is given by an identical equation to Eq. (12) and the correlation between the pooled excitatory and inhibitory inputs is given by

\[
\rho_{ii,e} = \rho_{ii,b} = \frac{\rho_{ee}^b}{\sqrt{\rho_{ee}^w + \frac{1}{n_e} (1-\rho_{ee}^w) \left( \rho_{ee}^w + \frac{1}{n_e} (1-\rho_{ee}^w) \right)}} = \frac{\rho_{ee}^b}{\sqrt{\rho_{ee}^w \rho_{ee}^w} + \mathcal{O}(1/n_e)}.
\]

**Correlations between the free membrane potentials**

We now look at the correlation between the free membrane potentials of two downstream neurons. The free membrane potentials are obtained by assuming an absence of threshold or spiking activity. For simplicity we assume symmetry in the statistics of the inputs to the postsynaptic cells: \( \sigma_{ee} = \sigma_{ii} = \sigma_{ei} = \sigma_{bi} = \rho_{ee}^b \), \( \rho_{ee}^b \), \( \rho_{ee}^w \), and \( \rho_{ii}^w \). The analysis is similar in the asymmetric case.

In the Section “Materials and Methods”, we derive a linear approximation of the free membrane potentials, \( V_i(t) = (J_i * K)(t) + V_i, \) where \( J_i(t) = -(g_e(t)(V_{ii} - V_i) + g_i(t)(V_{ei} - V_i))/C_m \) are the total input currents and \( K(t) = \Theta(t)e^{-t/\tau_m} \) for \( k = 1,2 \). Under this approximation, the correlation, \( \rho_{ii,V_i,V_j} \), between the membrane potentials is
equal to the correlation, $\rho_{m} = \rho_{i,j}$, between the total input currents and can be written as a weighted average of the pooled excitatory and inhibitory spike train correlations (see Appendix),

$$\rho_{V_{i}} = \rho_{m} = \frac{W_{E}^2\rho_{E,E} + W_{I}^2\rho_{I,I} - 2W_{E}W_{I}\rho_{E,I}}{W_{E}^2 + W_{I}^2 - 2W_{E}W_{I}}$$

where $\rho_{E,E}, \rho_{I,I}, \rho_{E,I}$, and $\rho_{E,I}$ are derived above, and $W_{E} = \mathcal{E}[V_{E} - V_{E}\sigma_{E}]$ and $W_{I} = \mathcal{I}[V_{I} - V_{I}\sigma_{I}]$ are weights for the excitatory and inhibitory contributions to the correlation. In Figure 5, we compare this approximation with simulations.

The correlation between the membrane potentials has positive contributions from the correlation between the excitatory inputs ($\rho_{E,E}$), and between the inhibitory inputs ($\rho_{I,I}$). Contributions coming from excitatory–inhibitory correlations ($\rho_{E,I}$ and $\rho_{E,I}$) are negative, and can thus decorrelate the activity of downstream cells. This “cancellation” of correlations is observed in Figures 1D and 5, and can lead to asynchrony in recurrent networks (Hertz, 2010; Renart et al., 2010).

**IMPLICATIONS FOR SYNCHRONIZATION IN FEEDFORWARD CHAINS**

Feedforward chains, like that depicted in Figure 6A, have been studied extensively (Diesmann et al., 1999; van Rossum et al., 2002; Litvak et al., 2003; Reyes, 2003; Tetzlaff et al., 2003; Câteau and Reyes, 2006; Doiron et al., 2006; Kumar et al., 2008). In such networks, cells in a layer necessarily share some of their inputs, leading to correlations in their spiking activity (Shadlen and Newsome, 1998). Frequently, spiking in deeper layers is highly synchronous (Reyes, 2003; Tetzlaff et al., 2003). However, in the presence of background noise, correlations can remain negligible (van Rossum et al., 2002; Vogels and Abbott, 2005).

Feedforward chains amplify correlations as follows: When inputs to the network are independent, small correlations are introduced in the second layer by overlapping inputs. The inputs to each subsequent layer are pooled from the previous layer. The amplification of correlations by pooling is the primary mechanism for the development of synchrony (Compare solid and dotted blue lines in Figure 4A). Overlapping inputs serve primarily to “seed” synchrony in early layers. The internal dynamics of the neurons and background noise can decorrelate the output of a layer, and compete with the correlation amplification due to pooling.

We develop this explanation by considering a feedforward network with each layer containing $N_{E}$ excitatory and $N_{I}$ inhibitory cells. Each cell in layer $k$ receives $n_{E}$ excitatory and $n_{I}$ inhibitory inputs selected randomly from layer $k$. For simplicity we assume that all excitatory and inhibitory cells are dynamically identical and $\mathcal{E}[V_{E} - V_{E}] = \mathcal{I}[V_{I} - V_{I}]$. Spike trains driving the first layer are statistically homogeneous with pairwise correlations $\rho_{ij}$.

![FIGURE 6](image-url) The development of synchrony in a feedforward chain can be understood using a model dynamical system (Tetzlaff et al., 2003).

(A) Schematic diagram of the network. Each layer consists of $N_{E}$ excitatory and $N_{I}$ inhibitory cells. Each cell in layer $k$ receives precisely $n_{E}$ excitatory and $n_{I}$ inhibitory, randomly selected inputs from layer $k-1$. (B) Stages of processing in the feedforward network. Inputs from layer $k-1$ are pooled with overlap, and drive the cells in layer $k$. (C) The correlation transfer map described by the pooling function, $\rho_{ij} = \mathcal{I}[V_{E} - V_{E}]$ (solid blue line). Cobweb shows the development of correlations in the discrete dynamical system defined by $\rho_{ij}^{n+1} = T(\rho_{ij}^{n})$ with $\rho_{ij} = 0$. Nearly perfect correlations develop by the fifth layer. The identity is shown as a dashed line. (D) Closer to balance (β=1), the correlating effects of pooling are weakened, and the model develops a stable fixed point close to $\beta = 0$. However, cells may no longer decorrelate their inputs in the balanced regime, and fluctuations in the input statistics due to random connectivity can destabilize the fixed point and lead to synchrony. The shaded region in the inset represents the region two standard deviations away from the mean (blue line) when randomness in the overlap is taken into account (see Appendix). The standard deviations were calculated using Monte Carlo simulations. In C and D, $N_{E} = 12000$ and $n_{E} = 600$. In C, $N_{I} = 8000$ and $n_{I} = 400$. In D, $N_{E} = 10500$ and $n_{E} = 512$ to obtain approximate balance (β = 6000/525). Filled black circles represent stable fixed points and open black circles represent unstable fixed points.
To explain the development of correlations, we consider a simplified model of correlation propagation (see also Renart et al., 2010 for a recurrent version). In the model, any two cells in a layer share the expected proportion $p_\beta = n_i/N_q$ of their excitatory inputs and $p_\beta = n_i/N_q$ of their inhibitory inputs (the expected proportions are taken with respect to random connectivity). We also assume that inputs are statistically identical across a layer.

For a pair of cells in layer $k \geq 1$, let $\rho_{\beta}^i$ and $\rho_{\beta}^o$ represent the correlation coefficient between the total input currents and output spike trains respectively. The outputs from layer $k$ are pooled (with overlap) to obtain the inputs to layer $k+1$. Using the results developed above, $\rho_{\beta}^i = P(\rho_\beta)$ and $\rho_{\beta}^o = P(\rho_{\beta}^o)$, for $k \geq 1$, where (see Appendix and Tetzlaff et al., 2003 for a similar derivation)

$$P(\rho) = \frac{p_\beta(1-p)(\rho_\beta + p_\beta)}{p_\beta(1-p)(1+\beta)}.$$  
(14)

Here $\beta$ measures the balance between excitation and inhibition (see Materials and Methods). From our assumptions, $\beta = n_i/n_q$. With imbalance ($\beta \neq 1$) and a large number of cells in a layer, pooling amplifies small correlations, $P(\rho) > \rho$, as discussed earlier.

To complete the description of correlation transfer from layer to layer, we relate the correlations between inputs to a pair of cells, $\rho_{\beta}^i$, to correlations in their output spike trains, $\rho_{\beta}^o$. We assume that there is a transfer function, $T$, so that $\rho_{\beta}^o = T(\rho_{\beta}^i)$ at each layer $k$. We additionally assume that $T(0) = 0$ and $T(1) = 1$, that is uncorrelated (perfectly correlated) inputs result in uncorrelated (perfectly correlated) outputs. We also assume that the cells are decorrelating, $|p| > |S(p)| > 0$ for $\rho \neq 0,1$ (Shea-Brown et al., 2008). This is an idealized model of correlation transfer, as output correlations depend on cell dynamics and higher order statistics of the inputs (Moreno-Bote and Parga, 2006; de la Rocha et al., 2007; Barreiro et al., 2009; Ostojic et al., 2009).

Correlations between the spiking activity of cells in layers $k+1$ are related to correlations in layer $k$ by the layer-to-layer transfer function, $T = S \circ P$. The development of correlations across layers is modeled by the dynamical system, $\rho_{\beta}^o = T(\rho_{\beta}^i)$, with $\rho_{\beta}^i = S(\rho_\beta)$.

When the network is not balanced ($\beta \neq 1$), pooling amplifies correlations at each layer and the activity between cells in deeper layers can become highly correlated (see Figure 6C). The output of the first layer is uncorrelated if the individual inputs are independent ($\rho_\beta = 0$). In this case all of the correlations between the total inputs to the second layer come from shared inputs,

$$\rho_{\beta}^o = P(0) = \frac{n_i p_\beta + n_i p_\beta}{n_i + n_i}.$$

These correlations are then reduced by the second layer of cells, $\rho_{\beta}^o = S(\rho_{\beta}^o) = T(0) > 0$, and subsequently amplified by pooling and input sharing before being received by layer 3, $\rho_{\beta}^o = P(\rho_{\beta}^o)$. This process continues in subsequent layers. If the correlating effects of pooling and input sharing dominate the decorrelating effects of internal cell dynamics, correlations will increase from layer to layer (see Figure 6C).

When $\rho_\beta = 0$, overlapping inputs increase the input correlation to layer 2, but have a negligible effect on the mapping once correlations have developed since the effects of pooling dominate [see Eq. (14) and the dashed blue line in Figure 4A which shows that the effects of input overlaps are small when $\eta_i$ is large, $p > 0$ and $\beta \neq 1$]. Therefore, shared inputs seed correlated activity at the first layer, and pooling drives the development of larger correlations. When $\rho_\beta = 0$, we cannot expect large correlations before layer 3, but when $\rho_\beta > 0$ large correlations can develop by layer 2.

To verify this conclusion, we constructed a two-layer feedforward network with no overlap between inputs ($P_i = P_\beta = 0$). In Figure 7A, the inputs to layer 1 were independent ($\rho_\beta = 0$), and the firing of cells in layer 2 was uncorrelated. In Figure 7B, we introduced small correlations ($\rho_\beta = 0.05$) between inputs to layer 1. These correlations were amplified by pooling so that strong synchrony is observed between cells in layer 2. We compared these results with a standard feedforward network with overlap in cell inputs (Figure 7C, where $P_i = P_\beta = 0.05$). Inputs to layer 1 were independent ($\rho_\beta = 0$), and hence outputs from layer 1 uncorrelated. Dependencies between inputs to layer 2 were weak and due to overlap alone, $\rho_{\beta}^o = P(0) = 0.05$. Cells in layer 3 received pooled inputs from layer 2, and their output was highly correlated.

These results predict that correlations between spike trains develop in deeper layers, but they do not directly address the timescale of the correlated behavior. In simulations, spiking becomes tightly synchronized in deeper layers (see for instance Litvak et al., 2003; Reyes, 2003; and Figure 7). This can be understood using results in Marásék et al. (1997) and Diesmann et al. (1999) where it is shown that the response of cells to volleys of spikes is tighter than the volley itself. The firing of individual cells in the network becomes bursty in deeper layers and large correlations are manifested in tightly synchronized spiking events. Alternatively, one can predict the emergence of synchrony by observing that pooling increases correlations over finite time windows (see next section and Appendix) and therefore the analysis developed above can be adapted to correlations over small windows.

**Balanced feedforward networks**

In the simplified feedforward model above, when excitation balances inhibition, that is $\beta = 1$, correlations between the pooled inputs to a layer are due to overlap alone, $\rho_{\beta}^o = P(\rho_{\beta}^o) = (p_\beta + p_\beta)/2$ for all $k$. The correlating effects of this map are weak, and this would seem to imply that cells in balanced feedforward chains remain asynchronous. Indeed, our model of correlation propagation displays a stable fixed point at low values of $p$ when $\beta = 1$ (see Figure 6D). However, in practice, synchrony is difficult to avoid without careful fine-tuning (Tetzlaff et al., 2003), and almost always develops in feedforward chains (Litvak et al., 2003). We provide some reasons for this discrepancy.

Our focus so far has been on correlations over infinitely large time windows (see Materials and Methods where we define $\rho_\beta$). Even when the membrane potentials are nearly uncorrelated over large time windows, differences between the excitatory and inhibitory synaptic time constants can cause larger correlations over smaller time windows (Renart et al., 2010). This can, in turn, lead to...
significant correlations between the output spike trains. We discuss this effect further in the Appendix and give an example in Figure 8. In this example, the correlations between the membrane potentials over long windows are nearly zero due to cancellation (see Figure 8A where $\rho_{VV} = 0.0174 \pm 0.0024$ s.e. with threshold present), but positive over shorter timescales. The cross-covariance function between the output spike trains is primarily positive, yielding significant spike train correlations ($\rho_{\text{spikes}} = 0.1570 \pm 0.0033$ s.e.). Therefore, the assumption that pairs of cells decorrelate their inputs may not be valid in the balanced case.

Another source of discrepancies between the idealized model and simulations of feedforward networks are inhomogeneities, which become important when balance is exact. Note that Eq. (14) is an approximation obtained by ignoring fluctuations in connectivity from layer to layer. In a random network, inhomogeneities will be introduced by variability in input population overlaps. To fully describe the development of correlations in a feedforward network, it is necessary to include such fluctuations in a model of correlation propagation. The asynchronous fixed point that appears in the balanced case has a small basin of attraction and fluctuations induced by input inhomogeneities could destroy its stability (see Figure 6D). Other sources of heterogeneity can further destabilize the asynchronous state (see Appendix).

It has been shown that asynchronous states can be stabilized through the decorrelating effects of background noise (van Rossum et al., 2002; Vogels and Abbott, 2005). To emulate these effects, a third transfer function, $N$, can be added to our model. The correlation transfer map then becomes $T(\rho) = S \circ N \circ P(\rho)$. Sufficiently strong background noise can increase decorrelation from input to output of a layer, and stabilize the asynchronous fixed point.

**DISCUSSION**

We have illustrated how pooling and shared inputs can impact correlations between the inputs and free membrane voltages of postsynaptic cells in a feedforward setting. The increase in correlation due to pooling was discussed in a simpler setting in (Bedenbaugh...
and Gerstein, 1997; Super and Roelfsema, 2005; Chen et al., 2006; Stark et al., 2008), and similar ideas were also developed for the variance alone in (Salinas and Sejnowski, 2000; Moreno-Bote et al., 2008). The saturation of the signal-to-noise ratio with increasing population size observed in (Zohary et al., 1994) has a similar origin. Our aim was to present a unified discussion of these results, with several generalizations.

Other mechanisms, such as recurrent connectivity between cells receiving the inputs, can modulate correlated activity (Schneider et al., 2006; Ostojic et al., 2009). Importantly, the cancellation of correlations may be a dynamic phenomenon in recurrent networks, as observed in (Hertz, 2010; Renart et al., 2010). On the other hand, neurons may become entrained to network oscillations, resulting in more synchronous firing (Womelsdorf et al., 2007). A full understanding of the statistics of population activity in neuronal networks will require an understanding of how these mechanisms interact to shape the spatiotemporal properties of the neural response.

The results we presented relied on the assumption of linearity at the different levels of input integration. These assumptions can be expected to hold at least approximately. For instance, there is evidence that membrane conductances are tuned to produce a linear response in the subthreshold regime (Morel and Levy, 2009). The assumptions we make are likely to break down at the level of single dendrites where nonlinear effects may be much stronger (Johnston and Narayanan, 2008). The effects of correlated inputs to a single dendritic branch deserve further theoretical study (Gasparrini and Magee, 2006; Li and Ascoli, 2006).

We demonstrated that the structure of correlations in a population may be difficult to infer from pooled activity. For instance, a change in pairwise correlations between individual cells in two populations causes a much smaller change in the correlation between the pooled signals. With a large number of inputs, the change in correlations between the pooled signals might not be detectable even when the change in the pairwise correlations is significant.

While we discussed the growth of second order correlations only, higher order correlations also saturate with increasing population size. For example, in a 3-variable generalization of the homogeneous model from Figure 3A, it can be shown that $\rho_{E,E,E} = 1 - O(1/n)$, where $n$ is the size of each population and $\rho_{E,E,E}$ is the triple correlation coefficient (Stratonovich, 1963) between the pooled signals $E_1$, $E_2$, and $E_3$. The reason that higher order correlations also saturate follows from the generalization of the following observation at second order: Pooling amplifies correlations because the variance and covariance grow asymptotically with the same rate in $n$. In particular $\sigma_{E,E}$ and $\gamma_{E,E}$ both behave asymptotically like $n^{1/2} \sigma_{E,E}^2 + O(n)$, and their ratio, $\rho_{E,E} = \gamma_{E,E} / \sigma_{E,E}$, approaches unity (Bedenbaugh and Gerstein, 1997; Salinas and Sejnowski, 2000; Moreno-Bote et al., 2008).

We concentrated on correlations over infinitely long time windows (see Materials and Methods where we define $\rho_{xy}$). However, pooling amplifies correlations over finite time windows in exactly the same way as correlations over large time windows. Due to the filtering properties of the cells, the timescale of correlations between downstream membrane potentials may not reflect that of the inputs. We discuss this further in the Appendix where the autocorrelation and cross-correlation functions between the membrane potentials are derived.

To simplify the presentation, we have so far assumed stationary. However, since Eq. (2) applies to the Pearson correlation between any pooled data, all of the results on pooling can easily be extended to the non-stationary case. In the non-stationary setting, the cross-covariance function has the form $R_{xy}(t) = \text{cov}(x(t), y(t+t))$, but there is no natural generalization of the asymptotic statistics defined in Eq. (1).

Correlated neural activity has been observed in a variety of neuronal populations (Gawne and Richmond, 1993; Zohary et al., 1994; Vaadia et al., 1995), and has been implicated in the propagation and processing of information (Oram et al., 1998; Maynard et al., 1999; Romo et al., 2003; Tiesinga et al., 2004; Womelsdorf et al., 2007; Stark et al., 2008), and attention (Steinmetz et al., 2000; Mitchell et al., 2009). However, correlations can also introduce redundancy and decrease the efficiency with which networks of neurons represent information (Zohary et al., 1994; Gutnisky and Dragoi, 2008; Goard and Dan, 2009). Since the joint response of cells and recorded signals can reflect the activity of large neuronal populations, it will be important to understand the effects of pooling to understand the neural code (Chen et al., 2006).

**APPENDIX**

**DERIVATION OF EQ. (10)**

Equation (10) can be derived from Eq. (2). However, we find that it is more easily derived directly. We will calculate the variance, $\sigma_{E,E}^2 = \sigma_{E,E}^2$, and covariance $\gamma_{E,E}$ between the pooled signals.

The covariance is given by the sum of all pairwise covariances between the populations, $\gamma_{E,E} = \sum_{i \neq j} \sigma_{E_i,E_j} \rho_{E_i,E_j}$. Each cell receives $n_i + q_n n$ inputs so that there are $(n_i + q_n)^2$ terms that appear in this sum. However, the $q_n^2$ “independent” inputs from each pool are uncorrelated with all other inputs and therefore don’t contribute to the sum. Of the remaining $n_i^2$ pairs, $n^2 \rho_{xy}$ are shared and therefore have correlation $\rho_{xy} = 1$. These shared processes therefore collectively contribute $n^2 \rho_{xy} \sigma_{xy}^2$ to $\gamma_{E,E}$. The remaining $n_n^2 - n^2 \rho_{xy}$ processes are correlated with coefficient $\rho_{xy}$ and collectively contribute $(n_n^2 - n^2 \rho_{xy}) \rho_{xy} \sigma_{xy}^2$. The pooled covariance is thus

$$\gamma_{E,E} = (n_n^2 - n^2 \rho_{xy}) \rho_{xy} \sigma_{xy}^2 + n^2 \rho_{xy} \sigma_{xy}^2.$$ 

The variance is given by the sum of all pairwise covariances within a population, $\sigma_{E,E}^2 = \sum_{i \neq j} \sigma_{E_i,E_j} \rho_{E_i,E_j}$. As above, there are $n_i + q_n n$ neurons in the population, so that the sum has $(n_i + q_n)^2$ terms. Of these, $n_i + q_n$ are “diagonal” terms ($E = E$), each contributing $\sigma_{E,E}$, for a total contribution of $(n_i + q_n) \sigma_{E,E}$ to $\sigma_{E,E}$. The processes from the independent pool do not contribute any additional terms. This leaves $n_i (n_i - 1)$ correlated pairs which each contribute $\sigma_{E,E}^2$ to the overall contribution of $n_i (n_i - 1) \sigma_{E,E}$, giving

$$\sigma_{E,E}^2 = \gamma_{E,E} = n_i (n_i - 1) \sigma_{E,E}^2 + (n_i + q_n) \sigma_{E,E}.$$ 

$$\rho_{E,E} = \gamma_{E,E} / \sigma_{E,E}$$

can be simplified to give Eq. (10). Equations for $\rho_{xy}$, and $\rho_{E,E}$ can be derived identically.

**FINITE-TIME CORRELATIONS AND CROSS-COVARIANCES**

Throughout the text, we concentrated on correlations over large time windows. However, the effects of pooling described by Eq. (2) apply to the correlation, $\rho_{xy}(t)$, between spike counts over any time window of
size, defined by $p_{\text{sp}}(t) = \text{cov}(N_x(t), N_y(t))/\sqrt{\text{var}(N_x(t)) \text{var}(N_y(t))}$ where $N_x(t) = \int_0^t x(s) ds$ is the spike count over $[0, t]$ for the spike train $x(t)$. The equation also applies to the instantaneous correlation at time $t$, defined by $R_{xy}(t) = C_{xy}(t)/\sqrt{C_{xx}(t)C_{yy}(t)}$. Thus pooling increases correlations over all timescales equally.

However, the cell filters the pooled inputs to obtain the membrane potentials and, as a result, the correlations between membrane potentials is “spread out” in time (Tetzlaff et al., 2008). To quantify this effect, we derive an approximation to the auto- and cross-covariance functions between the membrane potentials.

The pooled input spike trains are obtained from a weighted sum of the individual excitatory and inhibitory spike trains (see Materials and Methods). As a result, a cross-covariance function between the pooled spike trains are just sums of the individual cross-covariance functions, $C_{xy}(t) = \sum_{\nu \in \nu, \psi \in \psi} C_{xy}(t)$ for $X = E_1, E_2, I_1, I_2$ and $x, y = e, i$ accordingly. Thus only the magnitude of the cross-covariance functions is affected by pooling. The change in magnitude is quadratic in $n_1$ or $n_2$. This is consistent with the observation that pooling amplifies correlations equally over all timescales.

The conductances are obtained by convolving the total inputs with the synaptic filter kernels, $g_{\text{eff}}(t) = (E_x * \alpha_x)(t)$, and $g_{\text{eff}}(t) = (I_x * \alpha_x)(t)$. Thus, the cross-covariance between the conductances can therefore be written as a convolution of the cross-covariance function between the input signals and the deterministic cross-covariance between the synaptic kernels (Tetzlaff et al., 2008). In particular, $C_{C_{\text{ee}}, \text{eff}}(t) = (C_{xy} * (\alpha_x * \alpha_x))(t)$

for $X = E_1, E_2, I_1, I_2$ and $x, y = e, i$ accordingly, where $(\alpha_x * \alpha_x)(t) = \int_{-\infty}^t \alpha_x(s) \alpha_x(t+s) ds$ is the deterministic cross-covariance between the synaptic filters, $\alpha_x$ and $\alpha_i$. Note that total correlations remain unchanged by convolution of the input spike trains with the synaptic filters, since the integral of the convolution will be equal to the product of the integrals (Tetzlaff et al., 2008).

The total input currents, $I_{\text{eff}}(t) = -g_{\text{eff}}(t)(V_e - V_I)/C_{\text{m}}$, obtained from the linearization of the conductance-based model described in the Section “Materials and Methods” are simply linear combinations of the individual conductances. The cross-covariance function between the input currents is therefore a linear combination of those between the conductances, $C_{I_{\text{eff}}, \text{eff}}(t) = [V_e - V_I \int C_{\text{eff}, s_{\text{h}, k}}(t)] + [V_e - V_I \int C_{\text{eff}, s_{\text{h}, k}}(t)] - 2[V_e - V_I][V_e - V_I] C_{\text{eff}, s_{\text{h}, k}}(t)$

Combining this result with Eq. (1), yields the correlation, $\rho_{\text{in}} = \rho_{I_{\text{eff}}, \text{eff}}$, between the total input currents given in Eq. (13).

Using the solution of the linearized equations described in the Section “Materials and Methods”, we obtain a linear approximation to the cross-covariance functions by a factor of $\tau_{\text{eff}}^2$, and therefore leaves the ratio of the areas, $\rho_{\text{in}} = \rho_{I_{\text{eff}}, \text{eff}}$, unchanged. Thus, the linear approximation predicts that $\rho_{\text{in}} = \rho_{\text{in}}$.

When the total inputs are strong, $\tau_{\text{eff}}$ is small and we can simplify Eq. (16) by approximating $(K \times K)(t)$ with a delta function with mass $\tau_{\text{eff}}^2$ so that $C_{\text{dy}}(t) = \tau_{\text{eff}}^2 C_{\text{dy}}(t)/2$ and similarly for $C_{\text{xy}}(t)$. This approximation is valid when the synaptic time constants are significantly larger than $\tau_{\text{eff}}$, which is likely to hold in high conductance states. We compare this approximation to cross-covariance functions obtained from simulations in Figure 8.

In all examples considered, the cross-covariance functions have exponentially decaying tails. We define the correlation time constant, $\tau_{\text{xy}} = \lim_{t \to \infty} -t/\ln(C_{\text{xy}}(t))$, as a measure of the decay rate of the exponential tail. If $t > \tau_{\text{xy}}$, then $x(s)$ and $y(t+s)$ can be regarded as approximately uncorrelated and $\gamma_{xy}(t) = \int_{t}^{t+s} (s/|s|)C_{\text{xy}}(s) ds$ (Stratonovich, 1963).

The time constant of a convolution between two exponentially decaying functions is just the maximum time constant of the two functions. Thus, from the results above, the correlation time constant between the membrane potentials is the maximum of the correlation time constants between the inputs, the synaptic time constants, and the effective membrane time constant $\tau_{\text{xy}, \text{eff}} = \max(\tau_{\text{yy}, \text{eff}}, \tau_{\text{yy}, \text{eff}}, \tau_{\text{yy}, \text{eff}}, \tau_{\text{yy}, \text{eff}}, \tau_{\text{yy}, \text{eff}}, \tau_{\text{yy}, \text{eff}}, \tau_{\text{yy}, \text{eff}}, \tau_{\text{yy}, \text{eff}})$ where $\tau_{\text{yy}, \text{eff}}, \tau_{\text{yy}, \text{eff}}, \tau_{\text{yy}, \text{eff}}, \tau_{\text{yy}, \text{eff}}, \tau_{\text{yy}, \text{eff}}, \tau_{\text{yy}, \text{eff}}, \tau_{\text{yy}, \text{eff}}, \tau_{\text{yy}, \text{eff}}$ are the time constants of the input spike trains and $\tau_{\text{yy}, \text{eff}}, \tau_{\text{yy}, \text{eff}}, \tau_{\text{yy}, \text{eff}}, \tau_{\text{yy}, \text{eff}}, \tau_{\text{yy}, \text{eff}}, \tau_{\text{yy}, \text{eff}}, \tau_{\text{yy}, \text{eff}}, \tau_{\text{yy}, \text{eff}}$ are synaptic time constants. Thus the cross-covariance functions between the membrane potentials are generally broader than the cross-covariance functions between the spike train inputs.

**DERIVATION OF EQ. (14)**

Consider a feedforward network where each layer consists of $N_s$ excitatory cells and $N_i$ inhibitory cells; each cell in layer $k$ receives $n_e$ excitatory and $n_i$ inhibitory inputs from layer $(k-1)$, and these connections are chosen randomly and independently across neurons in layer $k$. Then the degree of overlap in the excitatory and inhibitory inputs to a pair of cells in layer $k$ is a random variable. Following the derivation in Derivation of Eq. (10) in Appendix, $\rho_{\text{in}, \text{eff}} = \frac{\gamma_{\text{yy}, \text{eff}}}{\sigma_{\text{yy}}} = \frac{s_{\text{yy}}}{n_{\text{yy}}} \sigma_{\text{yy}}^2 + \frac{(n_{\text{yy}} - s_{\text{yy}})}{\sigma_{\text{xx}}} \rho_{\text{x}, \text{yy}} \sigma_{\text{yy}}^2$ where $s_{\text{yy}}$ denotes the number of common excitatory inputs between the two cells. To understand the origin of $s_{\text{yy}}$, suppose the $n_e$ excitatory inputs to cell 1 have been selected. Then the selection of the $n_i$ inhibitory inputs to cell 2 involves choosing, without replacement, from two pools: the first, of size $n_{\text{yy}}$ projects to cell 1, and the second, of size $(N_i - n_i)$, does not. Therefore, $s_{\text{yy}}$ is follows a hyper-geometric distribution with parameters $(N_i, n_i, n_{\text{yy}})$, and hence $n_{\text{yy}} = s_{\text{yy}}/n_{\text{yy}}$. In addition, this random variable is independently selected amongst each pair in layer $k$. Using the mean value of $s_{\text{yy}}$, we obtain Eq. (10).

For simplicity, we assume that $\mathcal{E}[|V_e - V_I|] = \mathcal{E}[|V_e - V_I|]$, so that $\beta = n_i/n_e$. If we assume that the statistics in the $(k-1)^{th}$ layer are uniform across all cells and cell types (i.e., $\rho_{\text{in}} = \rho_{\text{x}, \text{yy}} = \rho_{\text{x}, \text{yy}}^0$, and $\sigma_{\text{yy}} = \sigma_{\text{yy}}$), then by substituting Eq. (10) and the equivalent forms for $\rho_{\text{x}, \text{yy}}$ into Eq. (13), we may write the input correlations to the $k^{th}$ layer as $\rho_{\text{in}} = \frac{|V_e - V_I| \gamma_{\text{yy}, \text{eff}}}{|V_e - V_I| \sigma_{\text{yy}}^2 + |V_e - V_I| \sigma_{\text{yy}}^2 - 2|V_e - V_I| \gamma_{\text{yy}, \text{eff}}}$.
Substituting the values of the covariances and variances, and dividing the numerator and denominator by \((\mathbb{E}V_e - V_e)\sigma^2\), we get

\[
\rho_i = \frac{s_i + (n_i^2 - s_i)p + s_j + (n_j^2 - s_j)p - 2n_i n_j p}{n_i + (n_i^2 - n_i)p + n_j + (n_j^2 - n_j)p - 2n_i n_j p}.
\]

Rearranging terms and dividing numerator and denominator by \(n_i^2\), along with the substitution \(\beta = n_i/n_j\), we have

\[
\rho_i = \frac{p(1 - \beta)^2 + \frac{1}{n_i}(s_i + s_j)(1 - p)}{p(1 - \beta)^2 + \frac{1}{n_i}(n_i + n_j)(1 - p)} \quad (17)
\]

This equation takes into account the variations in overlap due to finite size effects since \(s_i\) and \(s_j\) are random variables. Eq. (14) in the text represents the expected value \(\mathbb{E}(\rho_i) = \langle\rho_i\rangle\) which can be obtained by replacing the variables \(s_i\) and \(s_j\) in Eq. (17) with their respective means, \((s_i) = n_i p_i\) and \((s_j) = n_j p_j\). The expectation above is taken over realizations of the random connectivity of the feedforward network.

To calculate the standard deviation for the deviation in Figure 6D, we ran Monte Carlo simulations, drawing \(s_i\) and \(s_j\) from a hyper-geometric distribution and calculating the resulting transfer, \(\mathcal{S}(\rho_i) = \rho_i^*\) using Eq. (17). Note, however, that Eq. (17) and the inset in Figure 6D, do not account for all of the effects of randomness which may destabilize the balanced network. In deriving Eq. (17), we assumed that the statistics in the second layer were uniform. However, variations in the degree of overlap in one layer will cause inhomogeneities in the variances and rates at the next layer. In a feedforward setting, these inhomogeneities are compounded at each layer to destabilize the asynchronous fixed point.

**DEFINITIONS AND VALUES OF VARIABLES USED IN THE TEXT**

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Definitions of variables pertaining to recordings.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(X_j(t), X_j(t))</td>
<td>Signals from two populations.</td>
</tr>
<tr>
<td>(\rho_{s,s_j})</td>
<td>Correlation between the signals.</td>
</tr>
<tr>
<td>(\rho_{s,j})</td>
<td>Average pairwise correlation between a cell in population (j) and a cell in population (k).</td>
</tr>
<tr>
<td>(r(d))</td>
<td>Firing rate of a cell at distance (d) from the center of a stimulus.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Definitions of variables pertaining to downstream cells.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subscripts e and I (i and l) denote excitation (inhibition).</td>
<td></td>
</tr>
<tr>
<td>(n_e, n_i)</td>
<td>Number of correlated inputs to a cell.</td>
</tr>
<tr>
<td>(v_e, v_i)</td>
<td>Input rates.</td>
</tr>
<tr>
<td>(\xi, \iota)</td>
<td>Synaptic weights.</td>
</tr>
<tr>
<td>(q_e, q_i)</td>
<td>Neurons received (n_e q_e (n_i q_i)) independent excitatory (inhibitory) inputs.</td>
</tr>
<tr>
<td>(\rho_{ee}, \rho_{ei}, \rho_{ie}, \rho_{ii})</td>
<td>Correlations between pairs of afferents.</td>
</tr>
<tr>
<td>(\rho_{ei}, \rho_{ie})</td>
<td>Correlations within or between two non-overlapping populations ((x,y = e,i)).</td>
</tr>
<tr>
<td>(p_e, p_i)</td>
<td>Proportion of shared input to the post-synaptic pair.</td>
</tr>
<tr>
<td>(N_e, N_i)</td>
<td>Number of cells per layer in feed-forward network model.</td>
</tr>
<tr>
<td>(E_{e,i})</td>
<td>Pooled input spike trains to cell (j).</td>
</tr>
<tr>
<td>(\sigma_{ee}, \sigma_{ii})</td>
<td>Standard deviation of pooled excitatory or inhibitory spike trains.</td>
</tr>
<tr>
<td>(\gamma_{ei}, \gamma_{ie}, \gamma_{ii})</td>
<td>Covariance between pooled spike trains.</td>
</tr>
<tr>
<td>(\rho_{ei}, \rho_{ie})</td>
<td>Correlations between pooled spike trains.</td>
</tr>
<tr>
<td>(\beta)</td>
<td>Measure of balance between excitation and inhibition in the inputs.</td>
</tr>
<tr>
<td>(\rho_{ee}^{e}, \rho_{ee}^{i})</td>
<td>Correlations between inputs to or outputs from cell pairs in a feedforward network.</td>
</tr>
<tr>
<td>(C_{e,i}(\tau))</td>
<td>Cross-covariance function between processes (X) and (Y).</td>
</tr>
<tr>
<td>(S(\rho))</td>
<td>Correlations between the pooled inputs to cells in the feedforward model.</td>
</tr>
<tr>
<td>(\sigma_{e,i})</td>
<td>Correlation between output spike trains in terms of input current correlations between cell pairs in the feedforward model.</td>
</tr>
</tbody>
</table>

**Table 3 | Parameter values for simulations of two downstream cells.** For fields with “var,” various values of the indicated parameters were used and are described in the captions. For all simulations, \(V_e = -60\) mV, \(V_i = 0\) mV, \(V_c = -90\) mV, \(C_m = 114\) pF, \(g_l = 4.086\) nS, \(\tau_s = 10\) ms, \(\tau_i = 20\) ms.

<table>
<thead>
<tr>
<th>(\rho_{ee}, \rho_{ei}, n_e, n_i, v_e, v_i, (\xi, \iota))</th>
<th>(nS·ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1C</td>
<td>0.05</td>
</tr>
<tr>
<td>Figure 1D</td>
<td>0.05</td>
</tr>
<tr>
<td>Figure 5</td>
<td>var</td>
</tr>
<tr>
<td>Figure 8</td>
<td>0.05</td>
</tr>
</tbody>
</table>

| Table 4 | Parameter values for simulations of feedforward networks. The parameter \(v_e\) is the input rate to the first layer, \((\xi, \iota)\) indicates synaptic weights for excitatory cells, and \((\xi, \iota)\) for inhibitory cells. For all simulations, \(V_e = -60\) mV, \(V_i = 0\) mV, \(V_c = -90\) mV, \(C_m = 114\) pF, \(g_l = 4.086\) nS, \(\tau_s = 10\) ms, \(\tau_i = 20\) ms. For Figure 6, theoretical values were obtained under the assumption that \(v_e = v_i\) and \(\xi|V_e - V_i| = \iota|V_e - V_i|\). |

<table>
<thead>
<tr>
<th>(n_e, n_i)</th>
<th>(N_e, N_i)</th>
<th>(\rho_{ee}, \rho_{ei})</th>
<th>(v_e) (Hz)</th>
<th>(\xi, \iota)</th>
<th>(nS·ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 6C</td>
<td>600, 400</td>
<td>12000, 8000</td>
<td>0.05</td>
<td>NA</td>
<td>0</td>
</tr>
<tr>
<td>Figure 6D</td>
<td>600, 525</td>
<td>12000, 10500</td>
<td>0.05</td>
<td>NA</td>
<td>0</td>
</tr>
<tr>
<td>Figure 7A</td>
<td>225, 75</td>
<td>NA</td>
<td>0</td>
<td>100</td>
<td>0.05</td>
</tr>
<tr>
<td>Figure 7B</td>
<td>225, 75</td>
<td>NA</td>
<td>0</td>
<td>100</td>
<td>0.05</td>
</tr>
<tr>
<td>Figure 7C</td>
<td>225, 75</td>
<td>4500, 1500</td>
<td>0.05</td>
<td>100</td>
<td>0</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENTS

We thank Jaime de la Rocha, Brent Doiron and Eric Shea-Brown for helpful discussions. We also thank the reviewers and the handling editor for numerous useful suggestions. This work was supported by NSF Grants DMS-0604429 and DMS-0817649 and a Texas ARP/ATP award.

REFERENCES


**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 26 November 2009; paper pending published: 21 December 2009; accepted: 24 March 2010; published online: 19 April 2010.