frontiers in
COMPUTATIONAL NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 19 April 2010
doi: 10.3389/fncom.2010.00009

Pooling and correlated neural activity

Robert J. Rosenbaum*, James Trousdale and Krešimir Josić

Department of Mathematics, College of Natural Sciences and Mathematics, University of Houston, Houston, TX, USA

Edited by:
Philipp Berens,
Baylor College of Medicine, USA
Max Planck Institute for Biological Cybernetics, Germany

Reviewed by:
Nestor Parga,
Columbia University, USA
John A. Hertz,
Niels Bohr Institute, Denmark
Arvind Kumar,
University of Freiburg, Germany

****Correspondence:***
Robert J. Rosenbaum,
University of Houston,
Department of Mathematics, Houston, TX 77204-3008, USA.
e-mail: robertr@math.uh.edu

Correlations between spike trains can strongly modulate neuronal activity and affect the ability of neurons to encode information. Neurons integrate inputs from thousands of afferents. Similarly, a number of experimental techniques are designed to record pooled cell activity. We review and generalize a number of previous results that show how correlations between cells in a population can be amplified and distorted in signals that reflect their collective activity. The structure of the underlying neuronal response can significantly impact correlations between such pooled signals. Therefore care needs to be taken when interpreting pooled recordings, or modeling networks of cells that receive inputs from large presynaptic populations. We also show that the frequently observed runaway synchrony in feedforward chains is primarily due to the pooling of correlated inputs.

Keywords: correlation, pooling, synchrony, feedforward networks, synfire chains

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 Josic´, 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 postsynpatic 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; Ostojić 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 $C_{xy}(t) = \mathrm{cov}(x(s), y(s+t))$. 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,

$$\gamma_{xy} = \int_{-\infty}^{\infty} C_{xy}(t)dt, \quad \sigma_x^2 = \gamma_{xx}, \quad \rho_{xy} = \frac{\gamma_{xy}}{\sigma_x \sigma_y}. \tag{1}$$

While the asymptotic correlation, ρ_{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 $\{x_i\}_{i=1}^{n_x}$ and $\{y_j\}_{j=1}^{n_y}$, define the pooled variables, $X = \Sigma_i x_i$ and $Y = \Sigma_i y_i$. Since covariance is bilinear ($\mathrm{cov}(\Sigma_i x_i, \Sigma_j y_j) = \Sigma_{ij}\mathrm{cov}(x_i, y_j)$) the variance and covariance of the pooled variables are

$$\sigma_X^2 = \sum_{i=1}^{n_x}\sum_{\substack{j=1\\ j\neq i}}^{n_x} \sigma_{x_i}\sigma_{x_j}\rho_{x_i x_j} + \sum_{i=1}^{n_x}\sigma_{x_i}^2, \quad \text{and} \quad \gamma_{XY} = \sum_{i=1}^{nx}\sum_{j=1}^{ny}\sigma_{x_i}\sigma_{y_j}\rho_{x_i y_j},$$

and similarly for σ_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

$$\rho_{XY} = \frac{\bar{\rho}xy}{\sqrt{\left[w_x\bar{\rho}_{xx} + \frac{1}{n_x}\left(\frac{\bar{v}_x}{\overline{\sigma_x\sigma_y}} - w_x\bar{\rho}_{xx}\right)\right]\left[w_y\bar{\rho}_{yy} + \frac{1}{n_y}\left(\frac{\bar{v}_y}{\overline{\sigma_x\sigma_y}} - w_y\bar{\rho}_{yy}\right)\right]}} \tag{2}$$

$$= \frac{\bar{\rho}_{xy}}{\sqrt{\bar{\rho}_{xx}\bar{\rho}_{yy}}} + \mathcal{O}\left(\frac{1}{\sqrt{n_x n_y}}\right), \tag{3}$$

where

$$w_x = \frac{\overline{\sigma_x\sigma_x}}{\overline{\sigma_x\sigma_y}}, \qquad \bar{v}_x = \frac{1}{n_x}\sum_{i=1}^{n_x}\sigma_{x_i}^2,$$

$$\overline{\sigma_x\sigma_y} = \frac{1}{n_x n_y}\sum_{i=1}^{n_x}\sum_{j=1}^{n_y}\sigma_{x_i}\sigma_{y_i}, \quad \overline{\sigma_x\sigma_x} = \frac{1}{n_x(n_x-1)}\sum_{i=1}^{n_x}\sum_{\substack{j=1\\ j\neq 1}}^{n_x}\sigma_{x_i}\sigma_{x_j},$$

$$\bar{\rho}_{xy} = \frac{1}{n_x n_y\overline{\sigma_x\sigma_y}}\sum_{i=1}^{n_x}\sum_{j=1}^{n_y}\sigma_{x_i}\sigma_{y_j}\rho_{x_i y_j},$$

$$\bar{\rho}_{xx} = \frac{1}{n_x(n_x-1)\overline{\sigma_x\sigma_x}}\sum_{i=1}^{n_x}\sum_{\substack{j=1\\ j\neq 1}}^{n_x}\sigma_{x_i}\sigma_{y_j}\rho_{x_i x_j}$$

and similarly for w_y, v_y, $\overline{\sigma_y\sigma_y}$, and $\bar{\rho}_{yy}$. 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, $\bar{\rho}_{xy}$ represents the average correlation between x_i and y_j pairs, weighted by the product of their standard deviations, and similarly for $\bar{\rho}_{xx}$ and $\bar{\rho}_{yy}$. Correlation between weighted sums can be obtained by substituting $x_i \to w_{x_i}x_i$ and $y_j \to w_{y_j}y_j$ for weights w_{x_i} and w_{y_j} and making the appropriate changes to the terms in the equation above (e.g., $\sigma_{x_i} \to |w_{x_i}|\sigma_{x_i}, \rho_{x_i y_j} \to \mathrm{sign}(w_i w_j)\rho_{x_i y_j}$). 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 $i = 1,\ldots,n_x$ and $j = 1, \ldots, n_y$, simplifies these expressions. In particular, $\overline{v_x} = \overline{\sigma_x\sigma_x} = \sigma_x^2$, $\overline{\sigma_x\sigma_y} = \sigma_x\sigma_y$, and

$$\rho_{XY} = \frac{\bar{\rho}_{xy}}{\sqrt{\left[\bar{\rho}_{xx} + \frac{1}{n_x}\left(1-\bar{\rho}_{xx}\right)\right]\left[\bar{\rho}_{yy} + \frac{1}{n_y}\left(1-\bar{\rho}_{yy}\right)\right]}}. \tag{4}$$

Assuming further that the populations are symmetric, $\sigma_x = \sigma_y = \sigma$, $n_x = n_y = n$, and $\bar{\rho}_{xx} = \bar{\rho}_{yy}$, the expression above simplifies to

$$\rho_{XY} = \frac{\rho^b}{\rho^w + \frac{1}{n}\left(1-\rho^w\right)} = \frac{\rho^b}{\rho^w} + \mathcal{O}\left(\frac{1}{n}\right), \tag{5}$$

where $\rho^b = \bar{\rho}_{xy}$ is the average pairwise correlation between the two populations and $\rho^w = \bar{\rho}_{xx} = \bar{\rho}_{yy}$ is the average pairwise correlation within each population. Eq. (5) was derived in Bedenbaugh and Gerstein (1997) in an examination of correlations between multiunit recordings. In Chen et al. (2006), a version of Eq. (5) with $\rho^w = \rho^b$ is derived in the context of correlations between two VSD signals. The asymptotic, $\rho_{xy} \to 0$, limit when $\rho^w = \rho^b$ 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 jth excitatory input to cell k is labeled $\mathrm{e}_{j,k}(t)$ ($k = 1$ or 2). Similarly, $\mathrm{i}_{j,k}(t)$ denotes the jth inhibitory input to cell k. Each cell receives n_e excitatory and n_i inhibitory inputs with individual rates ν_e and ν_i respectively.

Each of the excitatory and inhibitory inputs to cell k, are stationary spike trains modeled by point processes, $\mathrm{e}_{j,k}(t) = \Sigma_i\,\delta(t - t_{j,k}^i)$ and $\mathrm{i}_{j,k}(t) = \Sigma_i\,\delta(t - s_{j,k}^i)$ where $\{t_{j,k}^i\}$ and $\{s_{j,k}^i\}$ are input spike times. We assume that the spike trains are stationary in a multivariate sense (Stratonovich, 1963). The pooled excitatory and inhibitory inputs to neuron k are $E_k(t) = \Sigma_{j=1}^{n_{\mathrm{e}_k}}\mathrm{e}_{j,k}(t)$, and $I_k(t) = \Sigma_{j=1}^{n_{\mathrm{i}_k}}\mathrm{i}_{j,k}(t)$.

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 5ms. The resulting processes are

Poisson with cross-covariance functions proportional to a double exponential, $C_{xy}(t) \sim e^{-|t|/5}$. Note that since each input is Poisson, $\sigma_e^2 = \nu_e$ and $\sigma_i^2 = \nu_i$.

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

$$C_m \frac{dV_k}{dt} = -g_L(V_k - V_L) - g_{E_k}(t)(V_k - V_E) - g_{I_k}(t)(V_k - V_I), \qquad (6)$$

with excitatory and inhibitory conductances determined by $g_{E_k}(t) = (E_k * \alpha_e)(t)$ and $g_{I_k}(t) = (I_k * \alpha_i)(t)$ where $*$ denotes convolution. We used synaptic responses of the form $\alpha_e(t) = \mathcal{E}\tau_e^{-2} t e^{-t/\tau_e}\Theta(t)$ and $\alpha_i(t) = \mathcal{I}\tau_i^{-2} t e^{-t/\tau_i}\Theta(t)$ 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, $\mathcal{E}$ or $\mathcal{I}$, with units nS·ms. This analysis can easily be extended to situations where each input, $e_{j,k}$ or $i_{j,k}$, has a distinct synaptic weight.

When examining spiking activity, we assume that when V_k crosses a threshold voltage, V^{th}, 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)

$$\beta = \frac{|V_L - V_E|\mathcal{E}\nu_e n_e}{|V_L - V_I|\mathcal{I}\nu_i n_i}.$$

When $\beta = 1$, the net excitation and inhibition are balanced and the mean free membrane potential equals V_L. In simulations, we set $V_L = -60$ mV, $V_E = 0$ mV, $V_I = -90$ mV, $\tau_e = 10$ ms, $\tau_i = 20$ ms, $C_m = 114$ pF, and $g_L = 4.086$ nS, giving a membrane time constant, $\tau_m = C_m/g_L = 27.9$ 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 $U = V_k - V_L$ so that Eq. (6) becomes

$$C_m \frac{dU}{dt} = \left(-g_L - g_E(t) - g_I(t)\right)U - g_E(t)\left(V_L - V_E\right) - g_I(t)\left(V_L - V_I\right).$$

Define the effective membrane time constant, $\tau_{eff} = C_m/(E[g_L + g_E(t) + g_I(t)]) = C_m/(g_L + n_e\nu_e\mathcal{E} + n_i\nu_i\mathcal{I}$ Substituting this average value in the previous equation yields the linear approximation to the conductance based model,

$$\frac{dU}{dt} = -\frac{1}{\tau_{eff}}U + J_k(t), \qquad (7)$$

where $J_k(t) = (-g_{E_k}(t)(V_L - V_E) - g_{I_k}(t)(V_L - V_I))/C_m$ is the total input current to cell k. Solving and reverting to the original variables gives the linear approximation $V_k(t) = (J_k * K)(t) + V_L$, where $K(t) = \Theta(t)e^{-t/\tau_{eff}}$ 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 Meterials 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, $X_1(t)$ and $X_2(t)$, 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

$$\rho_{X_1X_2} = \frac{\bar{\rho}_{12}}{\sqrt{\left[\bar{\rho}_{11} + \frac{1}{n}\left(1 - \bar{\rho}_{11}\right)\right]\left[\bar{\rho}_{22} + \frac{1}{n}\left(1 - \bar{\rho}_{22}\right)\right]}} = \frac{\bar{\rho}_{12}}{\sqrt{\bar{\rho}_{11}\bar{\rho}_{22}}} + \mathcal{O}\left(\frac{1}{n}\right). \qquad (8)$$

Here n represents the number of neurons recorded, $\bar{\rho}_{kk}$, $k = 1,2$ represents the average correlation between cells contributing to signal $X_k(t)$, and $\bar{\rho}_{12}$ 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 $\bar{\rho}_{11} = \bar{\rho}_{22}$ was derived by Bedenbaugh and Gerstein (1997).

When the two recording sites are nearby, so that $\bar{\rho}_{12} \approx \bar{\rho}_{11} \approx \bar{\rho}_{22}$, 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 $\bar{\rho}_{12}$, $\bar{\rho}_{11}$, and $\bar{\rho}_{22}$ 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 $\bar{\rho}_{11}$ and $\bar{\rho}_{22}$ increase by a larger factor than $\bar{\rho}_{12}$.

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,

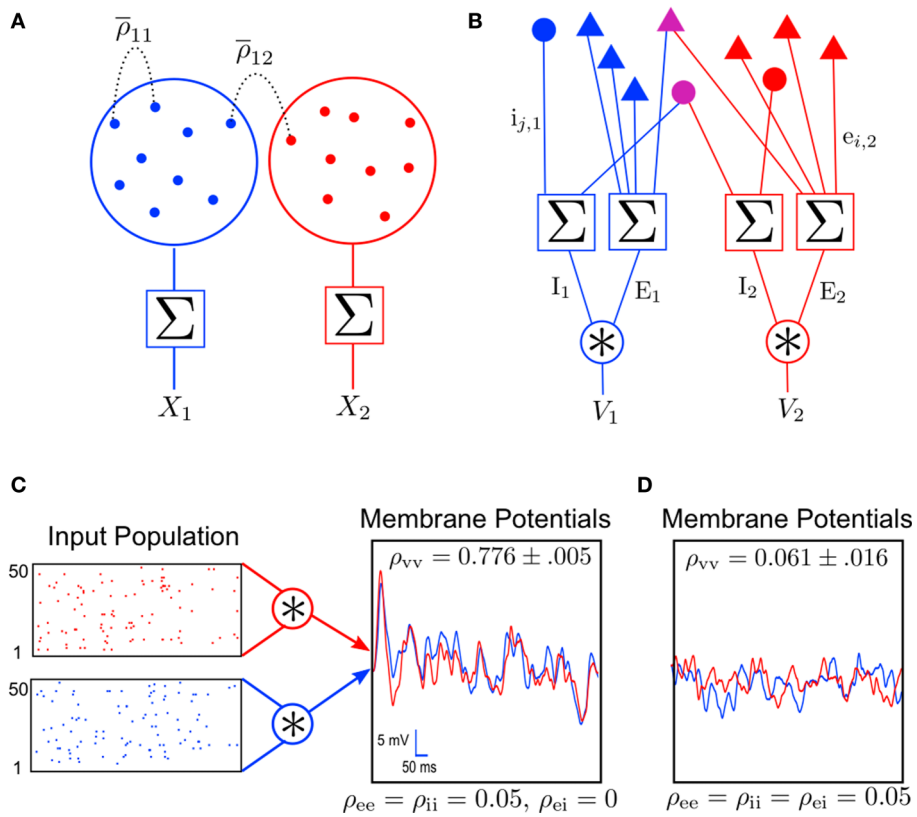

FIGURE 1 | Models of pooled recordings and the effects of pooling on correlations. (A) Pooling in experimental recordings. Cells from different populations are correlated with average correlation coefficient $\bar{\rho}_{12}$ and cells from the same population have average correlation coefficient $\bar{\rho}_{11}$ or $\bar{\rho}_{22}$. We examine correlations between the pooled signals, X_1 and X_2. **(B)** Pooled inputs to cell pairs. Individual excitatory (e) and inhibitory (i) inputs are correlated with coefficient ρ_{ee}, ρ_{ii}, and ρ_{ei}, respectively. The total input to a cell is the summed activity of its excitatory (E_k) and inhibitory (I_k) presynaptic population. The membrane potentials, V_1 and V_2, are obtained by filtering these inputs. **(C)** A simulation of the setup in **(B)** with background noise. Correlations between excitatory and between inhibitory cells were uniform ($\rho_{ee} = \rho_{ii} = 0.05$), but excitatory–inhibitory correlations were absent ($\rho_{ei} = 0$). The raster plot shows the activity in a subset of the input population. The correlation coefficient between the sub-threshold activity of the postsynpatic cells was $\rho_{vv} = 0.768 \pm 0.001$ s.e. Each cell receives 250 correlated and 250 uncorrelated excitatory Poisson inputs as well as 84 correlated and 84 uncorrelated inhibitory Poisson inputs ($n_e = 250$ and $n_i = 84$, $q_e = q_i = 1$, $\nu_e = 5$ Hz, $\nu_i = 7.5$ Hz, $\mathcal{I} = 4\mathcal{E}$, and $\mathcal{E} \approx 2.3$ nS·ms – see Materials and Methods and **Figure 3A** for notation and precise model description). **(D)** Same as **(B)**, but with $\rho_{ei} = 0.05$, $\nu_e = \nu_i = 5$ Hz, $\mathcal{I} = 6\mathcal{E}$ to maintain balance. In this case $\rho_{VV} = 0.0085 \pm 0.0024$ s.e.. The simulations were run 8000 times at 10 s each.

the imaged area is divided into 64 pixels, each 0.25 mm × .25 mm in size. The signal recorded from each pixel represents the pooled activity of $n \approx 1.25 \times 10^4$ neurons.

We model correlations between the signals, $X_1(t)$ and $X_2(t)$, 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

$$r(d) = \begin{cases} B + \dfrac{(1-B)(1+\cos(d\pi))^{\lambda}}{2} & \text{stimulus present} \\ B & \text{stimulus absent.} \end{cases} \tag{9}$$

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 $\rho_{jk} = S\sqrt{r(d_j)r(d_k)}\,e^{-\alpha D_{j,k}}$ where d_j and d_k are the distances from each cell to the center of the retinotopic image of the stimulus, $D_{j,k}$ is the distance between cells j and k, α 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 $\bar{\rho}_{11} = Sr(d_1)$ and $\bar{\rho}_{22} = Sr(d_2)$. The average correlation between cells in different pixels is $\bar{\rho}_{12} = S\sqrt{r(d_1)r(d_2)}\,e^{-\alpha D_{1,2}}$.

In this case, whether a stimulus is present or not, the correlation between the pooled signals is of the form $\rho_{X_1X_2} = e^{-\alpha D_{1,2}} + \mathcal{O}(1/n)$. 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.

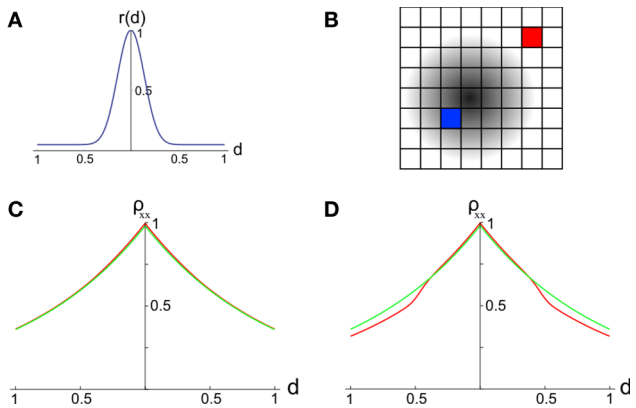

FIGURE 2 | The effect of pooling on recordings of stimulus dependent correlations. (A) The response amplitude of a model neuron as a function of its distance from the retinotopic image of a stimulus [Eq. (9)] with $B = 0.05$ and $\lambda = 10$. **(B)** A diagram of our model. Signals $X_1(t)$ and $X_2(t)$ are recorded from two pixels (red and blue squares). The activity in response to a stimulus is shown as a gradient centered at some pixel (the center of the retinotopic image of the stimulus). **(C)** The prediction of the correlation between two pixels obtained using the stimulus-dependent model considered in the text with stimulus present (red) and absent (green). We assumed that one pixel is located at the stimulus center ($d_1 = 0$). Parameters are as in **(A)** with $\alpha = 1$, $S = 0.1$, and $n = 1.25\times10^4$. A stimulus dependent change in correlations is undetectable. **(D)** Same as in **(C)**, except that baseline activity, B, was scaled by 0.5 in the presence of a stimulus. Compare to Figure 2f in Chen et al., 2006.

However, in Supplementary Figure 3 of Chen et al. (2006) the presence of a stimulus apparently results in a slight decrease in correlations between more distant pixels. In **Figure 2D** this effect was reproduced using the alternative model described above, with the additional assumption that baseline activity, B, decreases in the presence of a stimulus (Mitchell et al., 2009). The effect can also be reproduced by assuming that spatial correlation decay, α, increases when a stimulus is present.

As this example shows, care needs to be taken when inferring underlying correlation structures from pooled activity. The statistical structure of the recordings can depend on pairwise correlations between individual cells in a subtle way, and different underlying correlation structures may be difficult to distinguish from the pooled signals. However, downstream neurons may also be insensitive to the precise structure of pairwise correlations, as they are driven by the pooled input from many afferents.

CORRELATIONS BETWEEN THE POOLED INPUTS TO CELLS

We next examine the effects of pooling by relating the correlations between the activity of downstream cells to the pairwise correlations between cells in the input populations (see **Figure 1B**). The idea that pooling amplifies correlations carries over from the previous section. However, the presence of inhibition and non-instantaneous synaptic responses introduces new issues.

A homogeneous population with overlapping and independent inputs

For simplicity, we first consider a homogeneous population model (see **Figure 3A**). Each cell receives n_e inputs from a homogeneous pool of inputs with pairwise correlation coefficients ρ_{ee} and an additional $q_e n_e$ inputs from an outside pool of independent inputs. The two cells share $p_e n_e$ of the inputs drawn from the correlated pool. Processes in the independent pool are uncorrelated with all other processes. All excitatory inputs have variance σ_e^2.

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

$$\rho_{E_1E_2} = \frac{\rho_{ee} + \frac{p_e}{n_e}(1-\rho_{ee})}{\rho_{ee} + \frac{1}{n_e}(1-\rho_{ee}+q_e)}. \tag{10}$$

A form of this equation, with $p_e = 0$ 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, $\rho_{E_1E_2} = p_e$. When $\rho_{ee} > 0$ and n_e is large, pooled excitatory inputs are highly correlated, even when pairwise correlations in the presynaptic pool, ρ_{ee}, are small, and the neurons do not share inputs ($p_e = 0$). 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, $\rho_{I_1I_2}$, 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

$$\rho_{E_1I_2} = \rho_{I_1E_2} = \frac{\rho_{ei}}{\sqrt{\left(\rho_{ee} + \frac{1}{n_e}(1-\rho_{ee}+q_e)\right)\left(\rho_{ii} + \frac{1}{n_i}(1-\rho_{ii}+q_i)\right)}}. \tag{11}$$

Interestingly, since $|\rho_{E_1I_2}| \le 1$, pairwise excitatory–inhibitory correlations obey the bound $|\rho_{ei}| \le \sqrt{\rho_{ee}\rho_{ii}} + \mathcal{O}(1/\sqrt{n_e n_i})$. Combining this inequality with Eq. (10) and the analogous equation for $\rho_{I_1I_2}$, it follows that $|\rho_{E_1I_2}| \le \sqrt{\rho_{E_1E_2}\rho_{I_1I_2}} + \mathcal{O}(1/\sqrt{n_e n_i})$ 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, $\rho_{E_1E_2} = \bar{\rho}_{e_1e_2} / \sqrt{\bar{\rho}_{e_1e_1}\bar{\rho}_{e_2e_2}} + \mathcal{O}(1/\sqrt{n_{e_1}n_{e_2}})$. The term $\bar{\rho}_{e_1e_2}$ is a weighted average of the correlation coefficients *between* the two excitatory populations, and $\bar{\rho}_{e_1e_1}$ and $\bar{\rho}_{e_2e_2}$ are weighted averages of the correlations *within* each excitatory input population.

To illuminate this result, we assume symmetry between the populations: Let $n_{e_k} = n_e$ and $\sigma_{e_k^j} = \sigma_e$ for $k = 1,2$ and $j = 1,2,\ldots,n_e$, and assume $\bar{\rho}_{e_1e_1} = \bar{\rho}_{e_2e_2}$. The average "within" and "between" correlations, are $\rho_{ee}^{w} = \bar{\rho}_{e_1e_1} = \bar{\rho}_{e_2e_2}$ and $\rho_{ee}^{b} = \bar{\rho}_{e_1e_2}$ respectively (see **Figure 3B**). Under these assumptions, Eq. (5) can be applied to obtain (See also Bedenbaugh and Gerstein, 1997)

$$\rho_{E_1E_2} = \frac{\rho_{ee}^{b}}{\rho_{ee}^{w} + \frac{1}{n_e}(1-\rho_{ee}^{w})} = \frac{\rho_{ee}^{b}}{\rho_{ee}^{w}} + \mathcal{O}\left(\frac{1}{n_e}\right) \tag{12}$$

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, $\rho^{w}_{ee} > \rho^{b}_{ee}$. This reduces the correlation between the inputs, regardless of the input population size.

An increase in correlations in the presynaptic pool can also decorrelate 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 increased 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 $|\rho_{E_1E_2}| \le 1$, it follows that $|\rho^{b}_{ee}| \le |\rho^{w}_{ee}| + \mathcal{O}(1/n_e)$.

The correlation, $\rho_{I_1I_2}$, 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

$$\rho_{E_1I_2} = \rho_{I_1E_2} = \frac{\rho^{b}_{ei}}{\sqrt{\left(\rho^{w}_{ee} + \frac{1}{n_e}\left(1-\rho^{w}_{ee}\right)\right)\left(\rho^{w}_{ii} + \frac{1}{n_i}\left(1-\rho^{w}_{ii}\right)\right)}} = \frac{\rho^{b}_{ei}}{\sqrt{\rho^{w}_{ee}\rho^{w}_{ii}}} + \mathcal{O}\left(\frac{1}{\sqrt{n_e n_i}}\right).$$

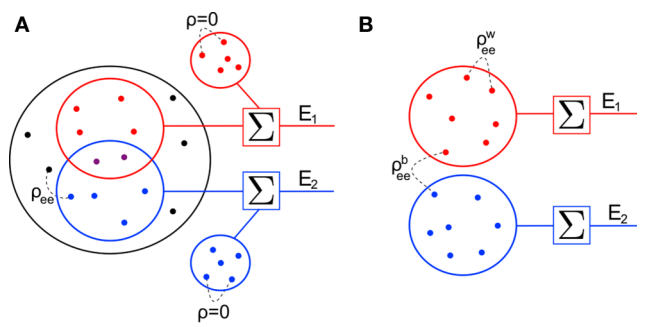

FIGURE 3 | Two population models considered in the text. (A) Homogeneous population with overlap and independent inputs: A homogeneous pool of correlated inputs (large black circle) with correlation coefficient between any pair of processes equal to ρ_{ee}. Each cell draws n_e inputs (larger red and blue circles) from this homogeneous input pool. Of these n_e correlated inputs, $p_e n_e$ are shared between the two neurons (purple dots). In addition, each cell receives $q_e n_e$ independent inputs (smaller red and blue circles), for a total of $n_e + q_e n_e$ inputs. All inputs have variance σ^2_e. **(B)** A population model with distinct "within" and "between" correlations: Each cell receives n_e inputs. The average correlation between two inputs to the same cell is ρ^{w}_{ee}, and between inputs to different cells is ρ^{b}_{ee}.

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: $\sigma_{E_k} = \sigma_E$, $\sigma_{I_k} = \sigma_I$, $\rho_{E_1I_2} = \rho_{E_2I_1}$, $\rho_{E_1I_1} = \rho_{E_2I_2}$, $\rho_{E_1E_1} = \rho_{E_2E_2}$ and $\rho_{I_1I_1} = \rho_{I_2I_2}$. The analysis is similar in the asymmetric case.

In the Section "Materials and Methods", we derive a linear approximation of the free membrane potentials,

$$V_k(t) = \left(J_k * K\right)(t) + V_L,$$

where $J_k(t) = -(g_{E_k}(t)(V_L - V_E) + g_{I_k}(t)(V_L - V_I))/C_m$ are the total input currents and $K(t) = \Theta(t)e^{-t/\tau_{eff}}$ for $k = 1,2$. Under this approximation, the correlation, $\rho_{V_1V_2}$, between the membrane potentials is

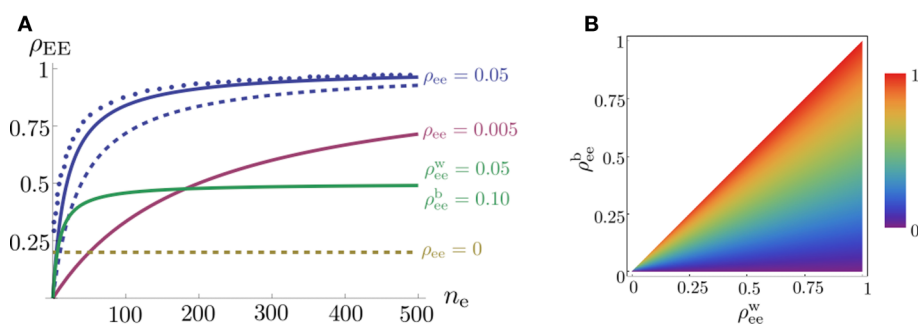

FIGURE 4 | The effect of pooling on correlations between summed input spike trains. (A) The correlation coefficient between the pooled excitatory spike trains (ρ_{EE}) is shown as a function of the size of the correlated excitatory input pool (n_e) for various parameter settings. The solid blue line was obtained by setting $\rho_{ee} = 0.05$ for the population model in **Figure 3A** in the absence of shared or independent inputs ($p_e = q_e = 0$). The dashed line illustrates the decorrelating effects of the addition of n_e independent inputs ($q_e = 1$, $q_e = 0$, $\rho_{ee} = 0.05$). The dotted blue line shows that shared inputs increase correlations, but have a diminishing effect on ρ_{EE} with increasing input population size ($p_e = 0.2$, $q_e = 0$, $\rho_{ee} = 0.05$). The solid pink line shows the effect of reducing the pairwise input correlations ($\rho_{ee} = 0.005$, $p_e = q_e = 0$). The dashed tan line was obtained with uncorrelated inputs so that correlations reflected shared inputs alone ($p_e = 0.2$, $\rho_{ee} = q_e = 0$). The green line was obtained with disparity in the "within" and "between" correlations ($\rho^{b}_{e} = 0.05$ and $\rho^{w}_{e} = 0.1$) using the model in **Figure 3B. (B)** The correlations coefficient, ρ_{EE}, between the pooled inputs as a function of the within and between correlations (ρ^{b}_{e} and ρ^{w}_{e}) for $n_e = 50$. Note that the pooled correlation is relatively constant along lines through the origin. Thus, changing ρ^{b}_{e} and ρ^{w}_{e} by the same proportion does not affect the pooled correlation.

equal to the correlation, $\rho_{in} = \rho_{J_1 J_2}$, between the total input currents and can be written as a weighted average of the pooled excitatory and inhibitory spike train correlations (see Appendix),

$$\rho_{V_1V_2} \approx \rho_{in} = \frac{W_E^2\rho_{E_1E_2} + W_I^2\rho_{I_1I_2} - 2W_EW_I\rho_{E_1I_2}}{W_E^2 + W_I^2 - 2W_EW_I\rho_{E_1I_1}} \tag{13}$$

where $\rho_{E_1E_2}, \rho_{E_1I_2}, \rho_{I_1I_2}$ and $\rho_{E_1I_1}$ are derived above, and $W_E = \mathcal{E}|V_E - V_L|\sigma_E$ and $W_I = \mathcal{I}|V_I - V_L|\sigma_I$ 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 ($\rho_{E_1E_2}$), and between the inhibitory inputs ($\rho_{I_1I_2}$). Contributions coming from excitatory–inhibitory correlations ($\rho_{E_1I_2}$ and $\rho_{E_2I_1}$) 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).

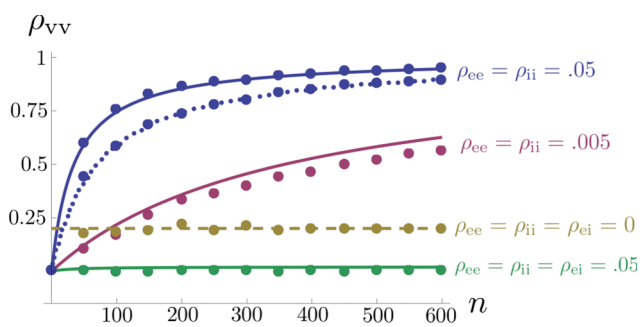

FIGURE 5 | The effects of pooling on correlations between postsynaptic membrane potentials. Results of the linear approximation (solid, dotted, and dashed lines) match simulations (points). For the solid blue line, $\rho_{ee} = \rho_{ii} = 0.05$, and $\rho_{ei} = p_e = p_i = q_e = q_i = 0$. The total number of excitatory and inhibitory inputs to each cell was $n = n_e + q_e n_e$, and $n_i + q_i n_i$ respectively. Here $n_i = n_e/3$, with other parameters given in the Section "Materials and Methods." The dotted blue line was obtained by including independent inputs, $q_e = q_i = 1$. The pink line was obtained by decreasing input correlations to $\rho_{ee} = \rho_{ii} = 0.005$. The solid green line was obtained by including excitatory–inhibitory correlations, $\rho_{ei} = 0.05$, so that total input correlations canceled. The dashed tan line was obtained by setting $\rho_{ee} = \rho_{ii} = \rho_{ei} = q_e = q_i = 0$ and $p_e = p_i = 0.2$ so that correlations are due to input overlap alone. In all cases, $\mathcal{E} = \frac{590}{n}$ nS·ms, and $\mathcal{I} = 4\mathcal{E}$. Standard errors are smaller than twice the radii of the points.

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 $k + 1$ 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 $\mathcal{E}|V_E - V_L| = \mathcal{I}|V_E - V_L|$. Spike trains driving the first layer are statistically homogeneous with pairwise correlations ρ_0.

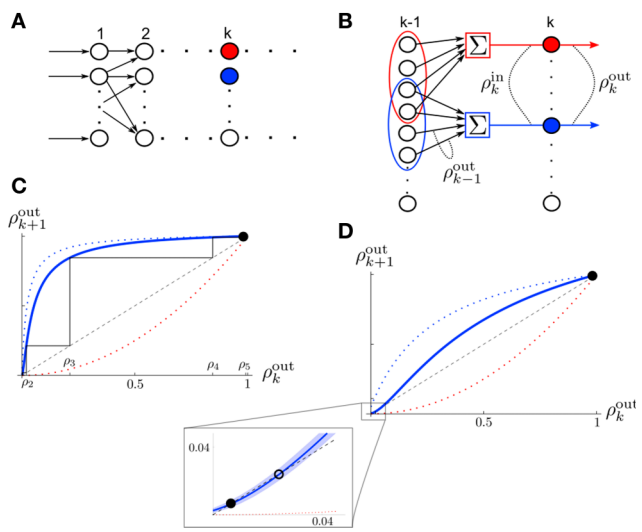

FIGURE 6 | The development of synchrony in a feedforward chain can be understood using a model dynamical system (Tetzlaff et al., 2003). **(A)** Schematic diagram of the network. Each layer consists of N_e excitatory and N_i inhibitory cells. Each cell in layer k receives precisely n_e excitatory and 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, $P(\rho)$ (blue dotted line), is composed with the decorrelating transfer function, $S(\rho) = \rho^2$ (red dotted line), to obtain the mapping, $T = S \circ P$ (solid blue line). Cobwebs show the development of correlations in the discrete dynamical system defined by $\rho_{k+1}^{out} = T(\rho_k^{out})$ with $\rho_0 = 0$. Nearly perfect correlations develop by the fifth layer. The identity is shown as a dashed line. **(D)** Closer to balance ($\beta \approx 1$), the correlating effects of pooling are weakened, and the model develops a stable fixed point close to $\rho = 0$. 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, $N_i = 10500$ and $n_i = 525$ to obtain approximate balance ($\beta = 600/525$). 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 $p_e = n_e/N_e$ of their excitatory inputs and $p_i = n_i/N_i$ 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 ρ_k^{in} and ρ_k^{out} 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, $\rho_{in}^1 = P(\rho_0)$ and $\rho_{k+1}^{in} = P(\rho_k^{out})$, for $k \geq 1$, where (see Appendix and Tetzlaff et al., 2003 for a similar derivation)

$$P(\rho) = \frac{\rho(\beta-1)^2 + \frac{1}{n_i}(1-\rho)\left(\beta p_e + p_i\right)}{\rho(\beta-1)^2 + \frac{1}{n_i}(1-\rho)(1+\beta)}. \quad (14)$$

Here β measures the balance between excitation and inhibition (see Materials and Methods). From our assumptions, $\beta = n_e/n_i$. 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, ρ_k^{in}, to correlations in their output spike trains, ρ_k^{out}. We assume that there is a transfer function, S, so that $\rho_k^{out} = S(\rho_k^{in})$ at each layer k. We additionally assume that $S(0) = 0$ and $S(1) = 1$, that is uncorrelated (perfectly correlated) inputs result in uncorrelated (perfectly correlated) outputs. We also assume that the cells are decorrelating, $|\rho| > |S(\rho)| > 0$ 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; Ostojić 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, $\rho_{k+1}^{out} = T(\rho_k^{out})$, with $\rho_1^{out} = S(\rho_0)$.

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 ($\rho_0 = 0$). In this case all of the correlations between the total inputs to the second layer come from shared inputs,

$$\rho_2^{in} = P(0) = \frac{n_e p_e + n_i p_i}{n_e + n_i}.$$

These correlations are then reduced by the second layer of cells, $\rho_2^{out} = S(\rho_2^{in}) = T(0) > 0$, and subsequently amplified by pooling and input sharing before being received by layer 3, $\rho_3^{in} = P(\rho_2^{out})$. 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 $\rho_0 = 0$, 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 n_e is large, $\rho > 0$ and $\beta \neq 1$]. Therefore, shared inputs seed correlated activity at the first layer, and pooling drives the development of larger correlations. When $\rho_0 = 0$, we cannot expect large correlations before layer 3, but when $\rho_0 > 0$ large correlations can develop by layer 2.

To verify this conclusion, we constructed a two-layer feedforward network with no overlap between inputs ($P_e = P_i = 0$). In **Figure 7A**, the inputs to layer 1 were independent ($\rho_0 = 0$), and the firing of cells in layer 2 was uncorrelated. In **Figure 7B**, we introduced small correlations ($\rho_0 = 0.05$) 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 $P_e = P_i = 0.05$). Inputs to layer 1 were independent ($\rho_0 = 0$), and hence outputs from layer 1 uncorrelated. Dependencies between inputs to layer 2 were weak and due to overlap alone, $\rho_2^{in} = P(0) = 0.05$. 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 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 Maršálek 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 $\beta \approx 1$, correlations between the pooled inputs to a layer are due to overlap alone, $\rho_k^{in} = P(\rho_{k-1}^{out}) \approx (p_e + p_i)/2$ 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 ρ when $\beta \approx 1$ (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 ρ_{xy}). 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

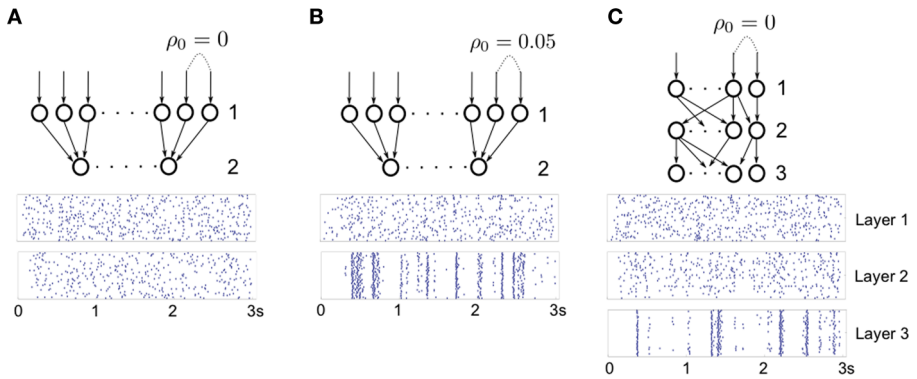

FIGURE 7 | Development of synchrony in feedforward networks. (A) A feedforward network with no overlap and independent, Poisson input. For excitatory cells, we set $\mathcal{E} \approx 1.55$nS·ms, and $\mathcal{I} \approx 4.67$nS·ms. For inhibitory cells, $\mathcal{E} \approx 3.61$nS·ms, and $\mathcal{I} \approx 10.82$nS·ms. **(B)** Same as A, except inputs to layer 1 are correlated with coefficient $\rho_0 = 0.05$. The network is highly synchronized in the second layer, even though inputs do not overlap. **(C)** Same as A, except for the presence of overlapping inputs ($p_e = p_i = 0.05$). Correlations due to overlap in the input to layer 2 result in average correlations of 0.05 between input currents. Layer 3 cells in C synchronize (Compare with layer 2 in B). In all three figures, each cell in the first layer was driven by excitatory Poisson inputs with rate $\nu_0 = 100$ Hz.

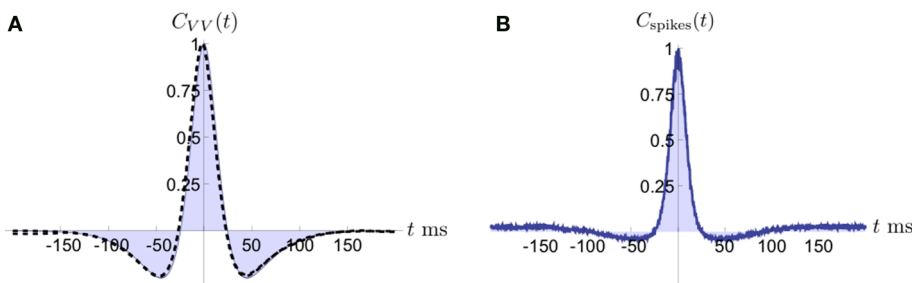

FIGURE 8 | Cross-covariance functions between membrane potentials and output spike trains. (A) The cross-covariance function between membrane potentials, scaled so that its maximum is 1. The linear approximation in Eq. (16) (blue, shaded) agrees with simulations of the full conductance-based model (black dashed line). Differences between simulations with and without threshold are too small to be observable (8000 simulations 10s each; simulations with and without threshold are shown). Parameters are as in **Figure 1D.** The cells are balanced with $\rho_{ee} = \rho_{ii} = \rho_{ei} = 0.05$ so that the correlation between the membrane potentials over long time windows is essentially zero ($\rho_{vv} = 0.0085 \pm 0.0024$ s.e. unthresholded, and $\rho_{vv} = 0.0174 \pm 0.0024$ s.e. thresholded). However, correlations over shorter time windows are positive as indicated by the central peak in the cross-covariance function. **(B)** The cross-covariance between the output spike trains is mostly positive. The correlation between the output spike trains was $\rho_{\text{spikes}} = 0.1570 \pm 0.0033$ s.e. (500 simulations of 100s each with same parameters as in A.

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; Ostojić 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 (Gasparini 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 $\rho_{E_1E_2E_3} = 1 - \mathcal{O}(1/n_e)$ where n_e is the size of each population and $\rho_{E_1E_2E_3}$ 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_e. In particular σ_E^2 and $\gamma_{E_1E_2}$ both behave asymptotically like $n_e^2\rho_{ee}\sigma_e^2 + \mathcal{O}(n_e)$, and their ratio, $\rho_{E_1E_2} = \gamma_{E_1E_2}/\sigma_E^2$, 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 ρ_{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 auto- and cross-covariance 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 $R_{xy}(s, t) = \mathrm{cov}\,(x(s), y(s+t))$, but there is no natural generalization of the asymptotic statistics defined in Eq. (1).

Correlated neural activity has been observed in a variety of neural 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, $\sigma_{E_1}^2 = \sigma_{E_2}^2$, and covariance $\gamma_{E_1E_2}$ between the pooled signals.

The covariance is given by the sum of all pairwise covariances between the populations, $\gamma_{E_1E_2} = \Sigma_{e_1\in E_1, e_2\in E_2}\sigma_{e_1}\sigma_{e_2}\rho_{e_1e_2}$. Each cell receives $n_e + q_e n_e$ inputs so that there are $(n_e + q_e n_e)^2$ terms that appear in this sum. However, the $q_e n_e$ "independent" inputs from each pool are uncorrelated with all other inputs and therefore don't contribute to the sum. Of the remaining n_e^2 pairs, $n_e p_e$ are shared and therefore have correlation $\rho_{e_1e_2} = 1$. These shared processes therefore collectively contribute $n_e p_e \sigma_e^2$ to $\gamma_{E_1E_2}$. The remaining $n_e^2 - n_e p_e$ processes are correlated with coefficient ρ_{ee} and collectively contribute $(n_e^2 - n_e p_e)\rho_{ee}\sigma_e^2$. The pooled covariance is thus

$$\gamma_{E_1E_2} = \underbrace{(n_e^2 - p_e n_e)\rho_{ee}\sigma_e^2}_{\text{Correlated}} + \underbrace{n_e p_e \sigma_e^2}_{\text{Shared}}.$$

The variance is given by the sum of all pairwise covariances within a population, $\sigma_{E_1}^2 = \Sigma_{e_1\in E_1, e_2\in E_1}\sigma_{e_1}\sigma_{e_2}\rho_{e_1e_2}$. As above, there are $n_e + q_e n_e$ neurons in the population, so that the sum has $(n_e + q_e n_e)^2$ terms. Of these, $n_e + q_e n_e$ are "diagonal" terms ($e_1 = e_2$), each contributing σ_e^2, for a total contribution of $(n_e + q_e n_e)\sigma_e^2$ to $\sigma_{E_1}^2$. The processes from the independent pool do not contribute any additional terms. This leaves $n_e(n_e - 1)$ correlated pairs which each contribute $\sigma_e^2\rho_{ee}$ for a collective contribution of $n_e(n_e - 1)\sigma_e^2\rho_{ee}$, giving

$$\sigma_{E_1} = \sigma_{E_2} = \underbrace{n_e(n_e - 1)\sigma_e^2\rho_{ee}}_{\text{Correlated}} + \underbrace{(n_e + n_e q_e)\sigma_e^2}_{\text{Diagonal}}.$$

Now, $\rho_{E_1E_2} = \gamma_{E_1E_2}/\sigma_{E_1}$ can be simplified to give Eq. (10). Equations for $\rho_{I_1I_2}$ and $\rho_{E_1I_2} = \rho_{I_1E_2}$ 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 t, defined by $\rho_{xy}(t) = \mathrm{cov}(N_x(t), N_y(t)) / \sqrt{\mathrm{var}(N_x(t))\,\mathrm{var}(N_y(t))}$ 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 from a weighted sum of the individual excitatory and inhibitory spike trains (see Materials and Methods). As a result cross-covariance functions between the pooled spike trains are just sums of the individual cross-covariance functions, $C_{XY}(t) = \Sigma_{x\in X, y\in Y} C_{xy}(t)$ for $X, Y = \mathrm{E}_1, \mathrm{E}_2, \mathrm{I}_1, \mathrm{I}_2$ and $x, y = \mathrm{e}, \mathrm{i}$ accordingly. Thus only the magnitude of the cross-covariance functions is affected by pooling. The change in magnitude is quadratic in n_e or n_i. 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,

$$g_{\mathrm{E}_k}(t) = (\mathrm{E}_k * \alpha_\mathrm{e})(t), \quad \text{and} \quad g_{\mathrm{I}_k}(t) = (\mathrm{I}_k * \alpha_i)(t); \quad k = 1,2.$$

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,

$$C_{g_X g_Y}(t) = (C_{XY} * (\alpha_x \star \alpha_y))(t) \tag{15}$$

for $X, Y = \mathrm{E}_1, \mathrm{E}_2, \mathrm{I}_1, \mathrm{I}_2$ and $x, y = \mathrm{e}, \mathrm{i}$ accordingly, where $(\alpha_x \star \alpha_y)(t) = \int_{-\infty}^{\infty} \alpha_x(s)\alpha_y(t+s)\mathrm{d}s$ is the deterministic cross-covariance between the synaptic filters, α_x and α_y. Note that total correlations remain unchanged by convolution of the input spike trains with the synaptic filters, since the integral of a convolution will be equal to the product of the integrals (Tetzlaff et al., 2008).

The total input currents, $J_K(t) = -(g_{E_k}(t)(V_L - V_I))/C_m$, 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,

$$\begin{aligned} C_{J_h J_k}(t) = {} & |V_\mathrm{E} - V_L|^2\, C_{g_{\mathrm{E}_h} g_{\mathrm{E}_k}}(t) + |V_\mathrm{I} - V_L|^2\, C_{g_{\mathrm{I}_h} g_{\mathrm{I}_k}}(t) \\ & - 2|V_\mathrm{E} - V_L||V_\mathrm{I} - V_L|\, C_{g_{\mathrm{I}_h} g_{\mathrm{E}_k}}(t). \end{aligned}$$

Combining this result with Eq. (1), yields the correlation, $\rho_\mathrm{in} = \rho_{J_1 J_2}$, 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,

$$C_{V_h V_k}(t) \approx \frac{1}{2\tau_\mathrm{eff}} (C_{J_h J_k} * (K \star K))(t) \tag{16}$$

for $h,k = 1,2$ where $(K \star K)(t) = \frac{\tau_\mathrm{eff}}{2} e^{-|t|/\tau_\mathrm{eff}}$ is the cross-covariance between the linear kernel, K, and itself. The convolution with $(K \star K)(t)$ scales the area of both the auto- and cross-covariance functions by a factor of τ_eff^2, and therefore leaves the ratio of the areas, $\rho_{V_1 V_2}$ unchanged. Thus, the linear approximation predicts that $\rho_{V_1 V_2} \approx \rho_\mathrm{in}$.

When the total inputs are strong, τ_eff is small and we can simplify Eq. (16) by approximating $(K \star K)(t)$ with a delta function with mass τ_eff^2 so that $C_{V_1 V_2}(t) \approx \tau_\mathrm{eff} C_{J_1 J_2}(t)/2$ and similarly for $C_{V_1 V_1}(t)$. This approximation is valid when the synaptic time constants are significantly larger than τ_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, $\tau_{xy} = \lim_{t\to\infty} -t/\ln(C_{xy}(t))$, as a measure of the decay rate of the exponential tail. If $t \gg \tau_{xy}$, then $x(s)$ and $y(s+t)$ can be regarded as approximately uncorrelated and $\gamma_{xy} \approx \int_{-t}^{t} (t - |s|/t) C_{xy}(s)\mathrm{d}s$ (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 $\tau_{V_1 V_2} = \max\{\tau_{\mathrm{E}_1\mathrm{E}_2}, \tau_{\mathrm{I}_1\mathrm{I}_2}, \tau_{\mathrm{E}_1\mathrm{I}_2}, \tau_{\mathrm{E}_2\mathrm{I}_1}, \tau_\mathrm{e}, \tau_\mathrm{i}, \tau_\mathrm{eff}\}$ where $\tau_{\mathrm{E}_1\mathrm{E}_2}, \tau_{\mathrm{I}_1\mathrm{I}_2}, \tau_{\mathrm{E}_1\mathrm{I}_2}$, and $\tau_{\mathrm{E}_2\mathrm{I}_1}$ are the time constants of the input spike trains and τ_e and τ_i are synaptic time constants. Thus the cross-covariances 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_e 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,

$$\rho_{\mathrm{E}_1\mathrm{E}_2} = \frac{\gamma_{\mathrm{E}_1\mathrm{E}_2}}{\sigma_\mathrm{E}^2} = \frac{s_\mathrm{e}\sigma_\mathrm{e}^2 + (n_\mathrm{e}^2 - s_\mathrm{e})\rho_{\mathrm{ee}}\sigma_\mathrm{e}^2}{n_\mathrm{e}\sigma_\mathrm{e}^2 + (n_\mathrm{e}^2 - n_\mathrm{e})\rho_{\mathrm{ee}}\sigma_\mathrm{e}^2},$$

where s_e denotes the number of common excitatory inputs between the two cells. To understand the origin of s_e, suppose the n_e excitatory inputs to cell 1 have been selected. Then the selection of the n_e excitatory inputs to cell 2 involves choosing, without replacement, from two pools: the first, of size n_e, projects to cell 1, and the second, of size $(N_\mathrm{e} - n_\mathrm{e})$, does not. Therefore, s_e is follows a hyper-geometric distribution with parameters $(N_\mathrm{e}, n_\mathrm{e}, n_\mathrm{e})$, and has mean $n_\mathrm{e}^2 / N_\mathrm{e} = n_\mathrm{e} p_\mathrm{e}$. In addition, this random variable is independently selected amongst each pair in layer k. Using the mean value of s_e, we obtain Eq. (10).

For simplicity, we assume that $\mathcal{E}|V_\mathrm{E} - V_L| = \mathcal{I}|V_\mathrm{I} - V_L|$, so that $\beta = n_\mathrm{i}/n_\mathrm{e}$. If we assume that the statistics in the $(k-1)^\mathrm{st}$ layer are uniform across all cells and cell types (i.e., $\rho = \rho_\mathrm{ee} = \rho_\mathrm{ii} = \rho_\mathrm{ei} = \rho_{k-1}^\mathrm{out}$ and $\sigma_\mathrm{e} = \sigma_\mathrm{i}$) then by substituting Eq. (10) and the equivalent forms for $\rho_\mathrm{II}, \rho_\mathrm{EI}$ in to Eq. (13), we may write the input correlations to the kth layer as

$$\rho_\mathrm{in} = \frac{|V_\mathrm{E} - V_L|^2\, \gamma_{\mathrm{E}_1\mathrm{E}_2} + |V_\mathrm{I} - V_L|^2\, \gamma_{\mathrm{I}_1\mathrm{I}_2} - 2|V_\mathrm{E} - V_L||V_\mathrm{I} - V_L|\, \gamma_{\mathrm{E}_1\mathrm{I}_2}}{|V_\mathrm{E} - V_L|^2\, \sigma_\mathrm{E}^2 + |V_\mathrm{I} - V_L|^2\, \sigma_\mathrm{I}^2 - 2|V_\mathrm{E} - V_L||V_\mathrm{I} - V_L|\, \gamma_{\mathrm{E}_1\mathrm{I}_1}}.$$

Substituting the values of the covariances and variances, and dividing the numerator and denominator by $(\mathcal{E}|V_E - V_L|\sigma_e)^2$, we get

$$\rho_{in} = \frac{s_e + (n_e^2 - s_e)\rho + s_i + (n_i^2 - s_i)\rho - 2n_e n_i \rho}{n_e + (n_e^2 - n_e)\rho + n_i + (n_i^2 - n_i)\rho - 2n_e n_i \rho}.$$

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

$$\rho_{in} = \frac{\rho(1-\beta)^2 + \frac{1}{n_i^2}(s_e + s_i)(1-\rho)}{\rho(1-\beta)^2 + \frac{1}{n_i^2}(n_e + n_i)(1-\rho)} \tag{17}$$

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

To calculate the standard deviation for the inset in **Figure 6D**, we ran Monte Carlo simulations, drawing s_e and s_i from a hypergeometric distribution and calculating the resulting transfer, $S(\rho_{in}) = \rho_{in}^2$ 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 1 | Definitions of variables pertaining to recordings.

$X_1(t)$, $X_2(t)$	Signals from two populations.
$\rho_{X_1X_2}$	Correlation between the signals.
$\bar{\rho}_{jk}$	Average pairwise correlation between a cell in population j and a cell in population k.
$r(d)$	Firing rate of a cell at distance d from the center of a stimulus.

Table 2 | Definitions of variables pertaining to downstream cells. Subscripts e and E (i and I) denote excitation (inhibition).

n_e, n_i	Number of correlated inputs to a cell.
ν_e, ν_i	Input rates.
$\mathcal{E}$, $\mathcal{I}$	Synaptic weights.
q_e, q_i	Neurons received $n_e q_e$ ($n_i q_i$) independent excitatory (inhibitory) inputs.
ρ_{ee}, ρ_{ii}, ρ_{ei}	Correlations between pairs of afferents.
ρ_{xy}^{w}, ρ_{xy}^{b}	Correlations within or between two non-overlapping populations (x,y = e,i).
p_e, p_i	Proportion of shared input to the post-synaptic pair.
N_e, N_i	Number of cells per layer in feed-forward network model.
E_j, I_j	Pooled input spike trains to cell j.
σ_E, σ_I	Standard deviation of pooled excitatory or inhibitory spike trains.
$\gamma_{E_jE_k}$, $\gamma_{E_jI_k}$, $\gamma_{I_jI_k}$	Covariance between pooled spike trains.
$\rho_{E_jE_k}$, $\rho_{E_jI_k}$, $\rho_{I_jI_k}$	Correlations between pooled spike trains.
β	Measure of balance between excitation and inhibition in the inputs.
ρ_k^{in}, ρ_k^{out}	Correlations between inputs to or outputs from cell pairs in a feedforward network.
$C_{XY}(t)$	Cross-covariance function between processes X and Y.
$P(\rho)$	Correlations between the pooled inputs to cells in the feedforward model.
$S(\rho)$	Correlation between output spike trains in terms of input current correlations between cell pairs in the feedforward model.

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_L = –60 mV, V_E = 0 mV, V_I = –90 mV, C_m = 114 pF, g_L = 4.086 nS, τ_e = 10 ms, τ_i = 20 ms.

	ρ_{ee}, ρ_{ii}	ρ_{ei}	n_e, n_i	p_e, p_i	q_e, q_i	ν_e, ν_i (Hz)	$\mathcal{E}$,$\mathcal{I}$ (nS·ms)
Figure 1C	0.05	0	250, 84	0	1	5, 7.5	2.3, 9.2
Figure 1D	0.05	0.05	250, 84	0	1	5	2.3, 13.8
Figure 5	var	var	var, n_e/3	var	var	5, 7.5	590/n_e, 4$\mathcal{E}$
Figure 8	0.05	0.05	250, 84	0	1	5	2.3, 13.8

Table 4 | Parameter values for simulations of feedforward networks. The parameter ν_0 is the input rate to the first layer, $(\mathcal{E}, \mathcal{I})_e$ indicates synaptic weights for excitatory cells, and $(\mathcal{E}, \mathcal{I})_i$ for inhibitory cells. For all simulations, V_L = –60 mV, V_E = 0 mV, V_I = –90 mV, C_m = 114 pF, g_L = 4.086 nS, τ_e = 10 ms, τ_i = 20 ms. For **Figure 6,** theoretical values were obtained under the assumption that $\nu_e = \nu_i$ and $\mathcal{E}|V_E - V_L| = \mathcal{I}|V_I - V_L|$.

	n_e, n_i	N_e, N_i	p_e, p_i	ν_0 (Hz)	ρ_0	$(\mathcal{E},\mathcal{I})_e$, $(\mathcal{E},\mathcal{I})_i$ (nS·ms)
Figure 6C	600, 400	12000, 8000	0.05	NA	0	NA
Figure 6D	600, 525	12000, 10500	0.05	NA	0	NA
Figure 7A	225, 75	NA	0	100	0	(1.55,5.67),(3.61,10.82)
Figure 7B	225, 75	NA	0	100	0.05	(1.55,5.67),(3.61,10.82)
Figure 7C	225, 75	4500, 1500	0.05	100	0	(1.55,5.67),(3.61,10.82)

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

Barreiro, A., Shea-Brown, E., and Thilo, E. (2009). Timescales of spike-train correlation for neural oscillators with common drive. *Phys. Rev. E* 81, Arxiv preprint arXiv:0907.3924.

Bedenbaugh, P., and Gerstein, G. (1997). Multiunit normalized cross correlation differs from the average single-unit normalized correlation. *Neural. Comput.* 9, 1265–1275.

Câteau, H., and Reyes, A. (2006). Relation between single neuron and population spiking statistics and effects on network activity. *Phys. Rev. Lett.* 96, 58101.

Chen, Y., Geisler, W. S., and Seidemann, E. (2006). Optimal decoding of correlated neural population responses in the primate visual cortex. *Nat. Neurosci.* 9, 1412–1420.

Cohen, M., and Maunsell, J. (2009). Attention improves performance primarily by reducing interneuronal correlations. *Nat. Neurosci.* 12, 1594–1600.

Coombes, S., Timofeeva, Y., Svensson, C., Lord, G., Josić, K., Cox, S., and Colbert, C. (2007). Branching dendrites with resonant membrane: A Òsum-over-tripsÓ approach. *Biol. Cybern.* 97, 137–149.

de la Rocha, J., Doiron, B., Shea-Brown, E., Josić, K., and Reyes, A. (2007). Correlation between neural spike trains increases with firing rate. *Nature* 448, 802–806.

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

Doiron, B., Rinzel, J., and Reyes, A. (2006). Stochastic synchronization in finite size spiking networks. *Phys. Rev. E* 74, 30903.

Ecker, A., Berens, P., Keliris, G., Bethge, M., Logothetis, N., and Tolias, A. (2010). Decorrelated neuronal firing in cortical microcircuits. *Science* 327, 584–587.

Gasparini, S., and Magee, J. (2006). State-dependent dendritic computation in hippocampal CA1 pyramidal neurons. *J. Neurosci.* 26, 2088.

Gawne, T., and Richmond, B. (1993). How independent are the messages carried by adjacent inferior temporal cortical neurons? *J. Neurosci.* 13, 2758.

Goard, M., and Dan, Y. (2009). Basal forebrain activation enhances cortical coding of natural scenes. *Nat. Neurosci.* 12, 1444–1449.

Gutnisky, D., and Dragoi, V. (2008). Adaptive coding of visual information in neural populations. *Nature* 452, 220–224.

Gutnisky, D. A., and Josić, K. (2010). Generation of spatio-temporally correlated spike-trains and local-field potentials using a multivariate autoregressive process. *J. Neurophys.* doi:10.1152/jn.00518.2009.

Hertz, J. (2010). Cross-correlations in high-conductance states of a model cortical network. *Neural. Comput.* 22, 427–447.

Johnston, D., and Narayanan, R. (2008). Active dendrites: colorful wings of the mysterious butterflies. *Trends Neurosci.* 31, 309–316.

Kuhn, A., Aertsen, A., and Rotter, S. (2003). Higher-order statistics of input ensembles and the response of simple model neurons. *Neural Comput.* 15, 67–101.

Kuhn, A., Aertsen, A., and Rotter, S. (2004). Neuronal integration of synaptic input in the fluctuation-driven regime. *J. Neurosci.* 24, 2345.

Kumar, A., Rotter, S., and Aertsen, A. (2008). Conditions for propagating synchronous spiking and asynchronous firing rates in a cortical network model. *J. Neurosci.* 28, 5268.

Li, X., and Ascoli, G. A. (2006). Computational simulation of the input–output relationship in hippocampal pyramidal cells. *J. Comput. Neurosci.* 21, 191–209.

Litvak, V., Sompolinsky, H., Segev, I., and Abeles, M. (2003). On the transmission of rate code in long feedforward networks with excitatory-inhibitory balance. *J. Neurosci.* 23, 3006.

Maršálek, P., Koch, C., and Maunsell, J. (1997). On the relationship between synaptic input and spike output jitter in individual neurons. *Proc. Natl. Acad. Sci. U.S.A* 94, 735.

Maynard, E. M., Hatsopoulos, N. G., Ojakangas, C. L., Acuna, B. D., Sanes, J. N., Normann, R. A., and Donoghue, J. P. (1999). Neuronal interactions improve cortical population coding of movement direction. *J. Neurosci.* 19, 8083–8093.

Mitchell, J. F., Sundberg, K. A., and Reynolds, J. H. (2009). Spatial attention decorrelates intrinsic activity fluctuations in macaque area v4. *Neuron* 63, 879–888.

Morel, D., and Levy, W. (2009). The cost of linearization. *J. Comput. Neurosci.* 27, 259–275.

Moreno-Bote, R., and Parga, N. (2006). Auto-and crosscorrelograms for the spike response of leaky integrate-and-fire neurons with slow synapses. *Phys. Rev. Lett.* 96, 28101.

Moreno-Bote, R., Renart, A., and Parga, N. (2008). Theory of input spike auto- and cross-correlations and their effect on the response of spiking neurons. *Neural Comput.* 20, 1651–1705.

Nunez, P., and Srinivasan, R. (2006). Electric Fields of the Brain: The Neurophysics of EEG. New York, NY: Oxford University Press.

Okun, M., and Lampl, I. (2008). Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. *Nat. Neurosci.* 11, 535–537.

Oram, M. W., Földiák, P., Perrett, D. I., and Sengpiel, F. (1998). The 'ideal homunculus': decoding neural population signals. *Trends Neurosci.* 21, 259–265.

Ostojić, S., Brunel, N., and Hakim, V. (2009). How connectivity, background activity, and synaptic properties shape the cross-correlation between spike trains. *J. Neurosci.* 29, 10234–10253.

Poort, J., and Roelfsema, P. (2009). Noise correlations have little influence on the coding of selective attention in area v1. *Cereb. Cortex* 19, 543.

Renart, A., de la Rocha, J., Bartho, P., Hollender, L., Parga, N., Reyes, A., and Harris, K. (2010). The asynchronous state in cortical circuits. *Science* 327, 587–590.

Reyes, A. (2003). Synchrony-dependent propagation of firing rate in iteratively constructed networks in vitro. *Nat. Neurosci.* 6, 593–599.

Roelfsema, P., Lamme, V., and Spekreijse, H. (2004). Synchrony and covariation of firing rates in the primary visual cortex during contour grouping. *Nat. Neurosci.* 7, 982–991.

Romo, R., Hernández, A., Zainos, A., and Salinas, E. (2003). Correlated neuronal discharges that increase coding efficiency during perceptual discrimination. *Neuron* 38, 649–657.

Salinas, E., and Sejnowski, T. J. (2000). Impact of correlated synaptic input on output firing rate and variability in simple neuronal models. *J. Neurosci.* 20, 6193–209.

Schneider, A., Lewis, T., and Rinzel, J. (2006). Effects of correlated input and electrical coupling on synchrony in fast-spiking cell networks. *Neurocomputing* 69, 1125–1129.

Shadlen, M. N., and Newsome, W. T. (1998). The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J. Neurosci.* 18, 3870–96.

Shea-Brown, E., Josic´, K., de La ocha, J., and Doiron, B. (2008). Correlation and synchrony transfer in integrate-and-fire neurons: basic properties and consequences for coding. *Phys. Rev. Lett.* 100, 108102.

Smith, M. A., and Kohn, A. (2008). Spatial and temporal scales of neuronal correlation in primary visual cortex. *J. Neurosci.* 28, 12591–12603.

Stark, E., Globerson, A., Asher, I., and Abeles, M. (2008). Correlations between groups of premotor neurons carry information about prehension. *J. Neurosci.* 28, 10618–10630.

Steinmetz, P., Roy, A., Fitzgerald, P., Hsiao, S., Johnson, K., and Niebur, E. (2000). Attention modulates synchronized neuronal firing in primate somatosensory cortex. *Nature* 404, 187–190.

Stratonovich, R. (1963). Topics in the Theory of Random Noise: General Theory of Random Processes. Nonlinear Transformations of Signals and Noise. New York, NY: Gordon and Breach.

Super, H., and Roelfsema, P. (2005). Chronic Multiunit Recordings in Behaving Animals: Advantages and Limitations. Development, dynamics and pathology of neuronal networks: from molecules to functional circuits: proceedings of the 23rd International Summer School of Brain Research, held at the Royal Netherlands Academy of Arts and Sciences, Amsterdam, from 25–29 August 2003, 263.

Tetzlaff, T., Buschermöhle, M., Geisel, T., and Diesmann, M. (2003). The spread of rate and correlation in stationary cortical networks. *Neurocomputing* 52, 949–954.

Tetzlaff, T., Rotter, S., Stark, E., Abeles, M., Aertsen, A., and Diesmann, M., (2008). Dependence of neuronal correlations on filter characteristics and marginal spike train statistics. *Neural Comput.* 20, 2133–2184.

Tiesinga, P. H., Fellous, J.-M., Salinas, E., José, J. V., and Sejnowski, T. J. (2004). Inhibitory synchrony as a mechanism for attentional gain modulation. *J. Physiol.* (Paris) 98, 296–314.

Troyer, T., and Miller, K. (1997). Physiological gain leads to high ISI variability in a simple model of a cortical regular spiking cell. *Neural. Comput.* 9, 971–983.

Vaadia, E., Haalman, I., Abeles, M., Bergman, H., Prut, Y., Slovin, H., and Aertsen, A. (1995). Dynamics of neuronal interactions in monkey cortex in relation to behavioural events. *Nature* 373, 515–518.

van Rossum, M., Turrigiano, G., and Nelson, S. (2002). Fast propagation of firing rates through layered networks of noisy neurons. *J. Neurosci.* 22, 1956–1966.

Vogels, T. P., and Abbott, L. F., (2005). Signal propagation and logic gating in networks of integrate-and-fire neurons. *J. Neurosci.* 25, 10786–10795.

Womelsdorf, T., Schoffelen, J.-M., Oostenveld, R., Singer, W., Desimone, R., Engel, A. K., Fries, P., and Jun. (2007). Modulation of neuronal interactions through neuronal synchronization. *Science* 316, 1609–1612.

Zohary, E., Shadlen, M., and Newsome, W. (1994). Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature* 370, 140–143.

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.

Citation: Rosenbaum RJ, Trousdale J and Josić K (2010) Pooling and correlated neural activity. Front. Comput. Neurosci. ***4****:9. doi: 10.3389/fncom.2010.00009*

Copyright © 2010 Rosenbaum, Trousdale and Josić. This is an open-access article subject to an exclusive license agreement between the authors and the Frontiers Research Foundation, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited.