Estimating the contribution of assembly activity to cortical dynamics from spike and population measures

Michael Denker · Alexa Riehle ·
Markus Diesmann · Sonja Grün

Abstract The hypothesis that cortical networks employ the coordinated activity of groups of neurons, termed assemblies, to process information is debated. Results from multiple single-unit recordings are not conclusive because of the dramatic undersampling of the system. However, the local field potential (LFP) is a mesoscopic signal reflecting synchronized network activity. This raises the question whether the LFP can be employed to overcome the problem of undersampling. In a recent study in the motor cortex of the awake behaving monkey based on the locking of coincidences to the LFP we determined a lower bound for the fraction of spike coincidences originating from assembly activation. This quantity together with the locking of single spikes leads to a lower bound for the fraction of spikes originating from any assembly activity. Here we derive a statistical method to estimate the fraction of spike synchrony caused by assemblies—not its lower bound—from the spike data alone. A joint spike and LFP surrogate data model demonstrates consistency of results and the sensitivity of the method. Combining spike and LFP signals, we obtain an estimate of the fraction of spikes resulting from assemblies in the experimental data.

Keywords LFP · Synchrony · Oscillations ·
Network dynamics · Motor cortex

M. Denker (✉) · M. Diesmann · S. Grün
RIKEN Brain Science Institute, Wako City, Japan
e-mail: mdenker@brain.riken.jp

A. Riehle
Mediterranean Institute of Cognitive Neuroscience (INCM),
CNRS—Univ. Aix-Marseille 2, Marseille, France

M. Diesmann
Brain and Neural Systems Team, RIKEN Computational
Science Research Program, Wako City, Japan

1 Introduction

A common hypothesis concerning the processing of information by cortical networks involves the propagation of activity through synchronously firing groups of neurons, termed assemblies (Hebb 1949). A hallmark signature of an activated assembly is the repeated coincident firing of neurons in relation to a specific stimulus or behavioral demand resulting in coincidence counts that exceed the chance level estimated from the firing rates (Gerstein et al. 1989). Despite the inherent undersampling of state of the art multiple single-unit recordings, experimental studies indirectly substantiate the assembly idea with findings of such behavior-related significant synchronous spiking (e.g., Riehle et al. 1997; Kilavik et al. 2009). Independently thereof, a signal on the population level, like the mesoscopic local field potential (LFP), typically exhibits temporally structured oscillations commonly interpreted as correlated network activity. Synaptic transmembrane currents have been identified as the primary contributor to LFP generation (Mitzdorf 1985; Logothetis and Wandell 2004). Counterintuitively, although synchronized membrane potential oscillations of neurons in the vicinity (Katzner et al. 2009) of the recording electrode show strong correlations with the LFP (Poulet and Petersen 2008;

Okun et al. 2010), these synchronized potentials do not induce the same degree of coincident spiking between the same neurons. Correlated spiking appears largely independent of synchrony on the level of membrane potentials (Tetzlaff et al. 2008; Poulet and Petersen 2008).

In a recent study (Denker et al. 2010) we were able to demonstrate in data of the motor cortex of the awake behaving monkey the missing link between significant spike synchrony and the LFP. A conceptual model enabled us to derive lower bounds for the fraction of spike coincidences β originating from observed assembly activity identified by periods containing excess spike synchrony, and the fraction of spikes of neurons γ caused by assembly activity whether observed or not. The results were obtained by comparing the locking of spike coincidences to the LFP in time periods with significant synchrony to the locking outside of these periods.

In the remainder of this section we first introduce the Unitary Events analysis method (Section 1.1) used to quantify significant excess spike synchrony. We then review our earlier findings (Denker et al. 2010) on the phase locking between spike-spike coincidences and the LFP (Section 1.2) and the model-based interpretation of the data (Section 1.3). In the present work we remove the limitation to lower bounds of β and γ by evaluating a stochastic model of the composition of spike coincidence counts. Section 2 introduces the model and derives an estimator for β only based on the configuration of spikes in the respective time interval. In Section 3 we explain that the probability of spikes γ to be part of an assembly—in contrast to β—cannot be extracted from the spike trains alone. However, equipped with the estimate of β we are in the position to compute γ using the previously found relationships between the phase distributions of the spike-LFP coupling (Section 1.3). The possibility to extract a parameter that relates to macroscopic features of the network dynamics illustrates the relevance of the LFP as a population signal in addressing the undersampling problem.

We demonstrate the consistency of the concept with the help of a joint spike-LFP toy model. In the physiological range, the parameters can reliably be determined and the lower bounds obtained from the phase locking are compatible with the parameter estimates yielded if the spike statistics is considered in addition. Finally, we utilize our new tool to reanalyze the experimental data (Section 4) for β and γ. Although highly simplified the toy model enables us to discuss (Section 5) various aspects of the distribution of parameters obtained from the experimental data set.

1.1 Analysis of spike synchrony

In the following we briefly summarize the main results obtained in Denker et al. (2010) which serve as a starting point for our analysis. In the first step we analyzed simultaneously recorded single units from monkey motor cortex (cf., e.g., Roux et al. 2006 and Section 4.1 for details) for excess spike synchrony by employing the Unitary Events analysis method (Grün et al. 2002b; Grün 2009). The method compares the empirically measured spike coincidences to the number expected by chance given the neurons' firing rates. The expected number defines the mean of the Poisson distribution realizing the null-hypothesis of statistical independence. If the p-value of the empirically found number of coincidences evaluated by comparison to this distribution exceeds the significance level α, the coincidences are considered significant and are termed Unitary Events (UE). This analysis is performed in a sliding window fashion (windows of 100 ms) to account for the dynamics of correlations and the non-stationarities in the firing rates in time (cf. Riehle et al. 1997 for details). Per sliding window data are considered from the various trials, and typically only a small fraction of the spikes in a window is coincident (Grün et al. 2002b, 2010). Non-stationarity across trials is accounted for by calculating the expected number of coincidences as the sum of trial-by-trial expectancies (Grün et al. 2003). Coincident spike events with a temporal precision of up to 3 ms were collected by the multiple-shift method to avoid artifacts due to binning (Grün et al. 1999). As a result of this analysis we are able to detect the dynamics of spike synchrony and identify time intervals that exhibit UE, i.e. contain excess spike synchrony (compare Fig. 1(a), top). Non-UE time periods contain coincidence spike events at chance level only. Thus UEs indicate time windows where the neurons under consideration exhibit coordinated activity across trials. The (unknown) set of all neurons within the network that generate this correlated activity is termed the assembly.

1.2 Analysis of phase-locking

In the next step we investigated the relation of spikes to the LFP. In particular, we were interested if spikes in synchronous events, i.e., chance coincidences (CC) or UE, have a different relation to the LFP than spikes that are not involved in a coincidence, i.e. isolated spikes (ISO). Therefore we classified each spike of a data set into one of these three classes and analyzed the classes separately (see Fig. 1(a) for an illustration; color codes for ISO, CC, and UE are retained throughout
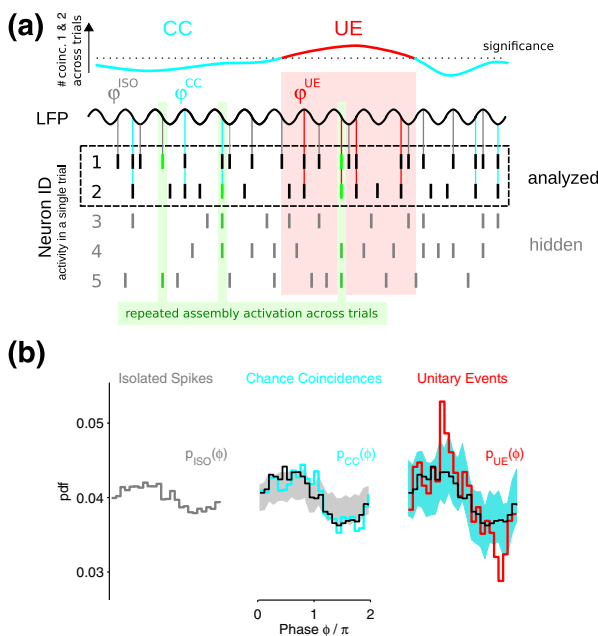
Fig. 1 Strongest phase locking of spikes to the LFP occurs during time periods of excess spike synchrony. (**a**) Sketch of the analysis: the number of pairwise synchronous spikes (precision: 3 ms) between two neurons observed across many trials (here, neurons 1 and 2, illustrated for one trial only) is evaluated in sliding windows of $T = 100$ ms (*top graph*). Synchronous spikes between the two neurons are classified as Unitary Events (UE; *red*) in windows where the observed number of coincidences exceeds a minimal number required to reach significance (*dotted line*). The latter is computed based on the null-hypothesis of independent firing considering the instantaneous firing rates. The *red shaded area* marks the center positions of UE windows. Windows not containing UE are classified as windows containing chance coincidences (CC; *cyan*). Spikes that are not part of an observed coincident event are classified as isolated spikes (ISO; *gray*). UE periods reflect the presence of excess spike synchrony caused by an assembly activated (*green*) repetitively across trials in relation to a behaviorally relevant event. Here, the assembly is sketched to consist of the observed neurons 1 and 2 and the unobserved neurons 4 and 5. In contrast, during CC periods each of the observed neurons may individually participate in other assemblies within the network (e.g., here, neuron 1 participates in an assembly that does not include neuron 2, and vice versa). The instantaneous phase ϕ of the LFP (beta) oscillation is extracted at spike times and pooled according to its classification as ISO (ϕ^{ISO}), CC (ϕ^{CC}), and UE (ϕ^{UE}). (**b**) Distributions of the LFP phases at spike occurrences for the three sets ISO (*left, gray*, $n = 240,455$), CC (*middle, cyan*, $n = 44,867$), and UE (*right, red*, $n = 12162$) from monkey motor cortex (see Section 4.1). The phase axes of the histograms span in total 1 period (2π) and are divided in 25 bins; phase π corresponds to the trough of the LFP oscillations. The *black curve in the middle and right graph* represents the expected phase distribution of coincident spikes assuming independent neurons. For comparison of the phase distributions, we extracted their variances (*middle*: gray, of ISO; *right*: cyan, of CC) by resampling from the set to the *left* with the same number of spikes as in the current set; each band encloses 95% of the 1,000 resampled distributions at each phase bin

this article). Spike triggered averages (STA) of the LFP revealed for all classes an oscillatory structure at about 17 Hz with spikes occurring preferentially at the decaying amplitude. However, the amplitude of the STAs derived for the different spike classes were strikingly different: consistent across the whole data set the STA triggered on UE spikes exhibited the largest amplitude, for CC spikes a smaller amplitude and for ISO the smallest.

However, the STA analysis cannot uncover the reason for the differences in STA amplitude, since it may be due to differences in the LFP amplitudes or due to different degrees of phase locking between spikes and the LFP. In order to disentangle these aspects we performed a phase-amplitude analysis of the LFP. Using a Hilbert transform of the LFP we gained the instantaneous phase and amplitude as time dependent functions, and extracted the respective measures at spike times (the detailed procedures will be explained in Section 4.1). The phase distribution exhibits the phase preferences by non-uniformity of the phase histograms (see Fig. 1(b)), and we observe that UE spikes express the strongest degree of phase locking, CC less (at the predicted chance level) and even less for ISO. A resampling procedure ensures that histograms are compared with comparative numbers of spikes: The phase histogram of CC (UE) is compared to resampled versions of the ISO (CC) phase histogram, each containing the same number of data points as the CC (UE) histogram (bands in Fig. 1(b), middle and right graph). The differences in phase locking were found to be consistent across the whole population of recorded neurons, and 68% of the neurons ($n = 123$ neuron pairs; see Section 4.1 for details) exhibit a stronger phase locking of UE than CC. In contrast we found much smaller or even negligible differences of the LFP amplitude measured as its envelope at spike times for the different spike classes.

1.3 Interpretation and conceptual model

Our analysis of the relation of spikes to the LFP revealed several unexpected results. Firstly, we found a difference in the locking degree for spikes involved in different categories of coincidences, i.e. UE vs CC. Following the hypothesis that active assemblies are expressed by coordinated spiking activity, UE coincidences are interpreted as a signature of such active assemblies. In the following, we will define the assembly as the set neurons exhibiting coordinated activity that is reflected in the excess synchronous events during UE periods. Our new finding on the phase locking of UE spikes to the LFP leads to the interpretation that

assembly activity occurs in a pulsed fashion, locked to the LFP oscillation. The fact that spikes of chance coincidences also occur phase locked seems contradictory. However, their degree of locking is fully explained by the (weak) locking of single spikes, which trivially leads to an enhanced locking of such spikes if occurring coincidently (cf., predictor in Fig. 1(b)). Thus the question is rather, why isolated spikes outside UE periods exhibit phase locking at all.

To get a better understanding of these puzzling observations, we developed a conceptual model which consistently explains all our findings (Fig. 2(a)). A direct conclusion of the experimental data, which show that UE periods exhibit a better locking to the LFP than CC periods, is that assembly spikes must be better locked to the LFP than spikes not participating in an assembly. Hence, one basic assumption of the model is that spikes involved in assemblies exhibit locking to the LFP while non-assembly spikes do not. According to our findings that individual spike trains typically contain a mixture of all spike categories (ISO, CC, and UE), it is reasonable to assume that an individual spike train is composed of assembly and non-assembly spikes. Due to the severe undersampling of the system caused by the limited number of recording electrodes, it is highly likely that also ISO spikes contain assembly spikes, but the corresponding partner neurons could not be identified. Consequently, the phase histogram of ISO must contain a mixture of unlocked and locked spikes. Assuming a uniform phase distribution $p_n(\phi)$ for non-assembly spikes and an unknown non-uniform phase distribution $p_a(\phi)$ for assembly spikes, the phase distribution of the ISO spikes $p_{\text{ISO}}(\phi)$ can be expressed by a weighted combination of the two components:

$$p_{\text{ISO}}(\phi) = (1 - \gamma) p_n(\phi) + \gamma p_a(\phi). \tag{1}$$

Thus also ISO spikes show locking, and its strength is determined by the factor γ.

CCs result from pairs of independent spike trains that express coincidences by chance. Under the above assumption that spike trains are composed of assembly and non-assembly spikes, chance coincidences are the result of combinations of different spike types: unlocked–unlocked spikes, locked–unlocked spikes (two possible combinations), and locked–locked spikes. Thus, $p_{\text{CC}}(\phi)$ is also expressed as a composition of $p_n(\phi)$ and $p_a(\phi)$:

$$\begin{aligned} p_{\text{CC}}(\phi) = & (1 - \gamma)^2 p_n(\phi) \\ & + 2\gamma(1 - \gamma) p_n(\phi) p_a(\phi) + \gamma^2 p_a^2(\phi). \end{aligned} \tag{2}$$

In case of time periods that contain UEs, coincidences resulting from an active assembly are present,
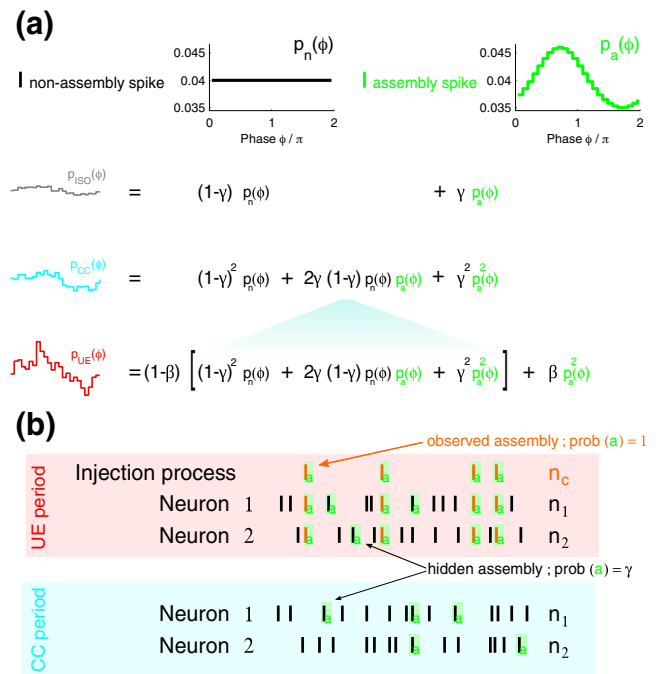
Fig. 2 Conceptual model to explain the differences in spike-LFP locking and definition of a suitable stochastic spike-LFP model. (**a**) An individual spike train is assumed to be composed of assembly spikes that exhibit locking to the LFP (non-uniform phase distribution $p_a(\phi)$, *top right*) and of non-assembly spikes that do not lock to the LFP (uniform phase distribution $p_n(\phi)$, *top left*). Their respective occurrences within a spike train is expressed by γ (and $1 - \gamma$) with γ indicating the overall probability that any spike of a neuron within the network is part of the assembly process (which might go unobserved). CC are composed of chance combinations of assembly and non-assembly spikes, whereas an UE period has contributions of assembly coincidences of fraction β and CC coincidences of fraction $1 - \beta$. The observed phase distributions for isolated spikes $p_{\text{ISO}}(\phi)$, for chance coincidence spikes $p_{\text{CC}}(\phi)$, and Unitary Event spikes $p_{\text{UE}}(\phi)$ are expressed as combinations of $p_n(\phi)$ and $p_a(\phi)$ by the three respective equations. (**b**) Stochastic spike-LFP model (details in Section 3.1) used to estimate parameters β and γ from data. UE periods (*shaded in red*) are modeled as independent spike trains of a given background rate, and excess coincidences due to an additional process whose spikes are inserted into both parallel spike trains (labeled by *green letter* 'a'). The latter is modeled as Poisson process with n_c spikes. The background processes of each neuron contain $n_1 - n_c$ and $n_2 - n_c$ spikes, respectively. A fraction γ of the background spikes is also part of an unobserved assembly, here selected randomly and also marked with the *green letter* 'a'. Labeled spikes exhibit locking to the LFP according to $p_a(\phi)$, non-labeled spikes do not lock (uniform phase distribution $p_n(\phi)$). A CC period (*shaded in cyan*) is realized as two independent Poisson processes with the same spike counts n_1 and n_2 for neuron 1 and 2, and a fraction γ of them marked as spikes of an unobserved assembly

but likely intermixed with chance coincidences that are not related to this assembly. Thus, the phase distribution of UE is assumed to be composed of a contribution of chance locking of a percentage of $(1 - \beta)$, and to a

percentage of β of assembly coincidences (see sketch in Fig. 2(a)):

$$
\begin{aligned}
p_{\mathrm{UE}}(\phi) = (1 - \beta) \cdot \big[&(1 - \gamma)^2 p_{\mathrm{n}}(\phi) \\
&+ \gamma(1 - \gamma) p_{\mathrm{n}}(\phi) p_{\mathrm{a}}(\phi) + \gamma^2 p_{\mathrm{a}}^2(\phi) \big] \\
&+ \beta p_{\mathrm{a}}^2(\phi).
\end{aligned} \tag{3}
$$

As a result, the enhanced locking of UE as compared to CC is a consequence of the presence of assembly coincidences i.e. $\beta > 0$.

2 Estimation of assembly activations from spike statistics

In terms of spike coincidences, β is the ratio of the number of coincidences resulting from assembly activation and the total number of coincidences in a given time window

$$
\beta = \frac{n_{\mathrm{c}}}{n_{\mathrm{emp}}}. \tag{4}
$$

The task therefore is to construct an estimate of n_{c} on the basis of the known properties of the observed spike trains. For two independently spiking neurons with n_1 spikes in one spike train and n_2 spikes in the other, the probability to observe k coincidences in a time window T is given by the hypergeometric distribution (Grün et al. 2003)

$$
\mathcal{H}_{T_h}^{n_1, n_2}(k) \tag{5}
$$

where we introduced the shorthand

$$
T_h = T/h \tag{6}
$$

and h is the resolution of the discretized time axis. We define $\mathcal{H} = 0$ outside of $[0, T_h]$. The expected number of coincidences is

$$
n_{\mathrm{exp}} = \frac{1}{T_h} \cdot n_1 n_2. \tag{7}
$$

The expectation value (Eq. (7)) is just the probability to observe a coincidence in a given bin $(n_1/T_h)(n_2/T_h)$ multiplied by the number of available time bins T_h. In the presence of n_{c} deterministic coincidences, however, only a reduced number of spikes of the two neurons is available to form chance coincidences

$$
n_{\mathrm{exp,c}} = n_{\mathrm{c}} + \frac{1}{T_h - n_{\mathrm{c}}} \cdot (n_1 - n_{\mathrm{c}})(n_2 - n_{\mathrm{c}}).
$$

The construction of the respective spike trains is illustrated in Fig. 2(b). Let us now assume that we measure the number of coincidences $\langle n_{\mathrm{emp}} \rangle$ averaged over many repetitions of the experiment with exactly the same values for n_{c}, n_1, n_2. In this case we can equate the

empirical average with the expectation value $\langle n_{\mathrm{emp}} \rangle = n_{\mathrm{exp,c}}$. In this expression n_{c} is the only unknown variable. Hence, we can express n_{c} in terms of n_1, n_2, and $\langle n_{\mathrm{emp}} \rangle$. Given only the triplet of a single realization $(n_1, n_2, n_{\mathrm{emp}})$ we can still hope that the measured n_{emp} is typical and write

$$
n_{\mathrm{emp}} = n_{\mathrm{c}} + \frac{1}{T_h - n_{\mathrm{c}}} \cdot (n_1 - n_{\mathrm{c}})(n_2 - n_{\mathrm{c}}). \tag{8}
$$

Multiplication with $(T_h - n_{\mathrm{c}})$

$$
\begin{aligned}
(T_h - n_{\mathrm{c}}) n_{\mathrm{emp}} = {} & (T_h - n_{\mathrm{c}}) n_{\mathrm{c}} \\
& + \left(n_1 n_2 - n_{\mathrm{c}} n_2 - n_{\mathrm{c}} n_1 + n_{\mathrm{c}}^2 \right)
\end{aligned}
$$

shows that the quadratic terms in n_{c} cancel

$$
-n_{\mathrm{c}} n_{\mathrm{emp}} - T_h n_{\mathrm{c}} = -T_h n_{\mathrm{emp}} - n_{\mathrm{c}}(n_1 + n_2) + n_1 n_2
$$

and collecting terms with n_{c} gives

$$
n_{\mathrm{c}} = \frac{T_h n_{\mathrm{emp}} - n_1 n_2}{T_h + n_{\mathrm{emp}} - (n_1 + n_2)}. \tag{9}
$$

Using Eq. (9) we can already compute β for all experimental UE periods with their individual triplets $(n_1, n_2, n_{\mathrm{emp}})$ and obtain the distribution of β as well as its mean β^{UE} (see Section 4). Clearly Eq. (9) is an approximation, the expression is negative for small n_{emp}, in particular when no coincidence has been observed ($n_{\mathrm{emp}} = 0$).

In order to capture coincidences with a temporal jitter larger than the resolution of the data h, the multiple-shift method (Grün et al. 1999) sums the coincidences over a range of $s = 2b + 1$ displacements $-b, \ldots, +b$ of one spike train with respect to the other. At each shift only some of the n_{c} coincidences are detected and a particular coincidence is by definition only detected at a single displacement. Under the assumption that in the s displacements the spike counts n_1 and n_2 are conserved, the relation (Eq. (8)) between n_{emp} and n_{c} holds if we replace the number of spikes available for chance coincidences by $s n_i - n_{\mathrm{c}}$ and the number of available bins by $s T_h - n_{\mathrm{c}}$. The estimate of the number of coincidences originating from assembly activation (Eq. (9)) then reads $n_{\mathrm{c}} = \left(s T_h n_{\mathrm{emp}} - s^2 n_1 n_2 \right) / \left(s T_h + n_{\mathrm{emp}} - s(n_1 + n_2) \right)$.

Next we consider a slightly more realistic model with a variable number of assembly activations n_{c}. The purpose of the model is to help us to understand the conditions under which we can reliably extract β. In the absence of any knowledge about the process generating the additional coincidences n_{c} we assume that each value consistent with the observed number of coincidences is equally likely. Thus, each possible value of n_{c} occurs with probability $1/(n_{\mathrm{emp}} + 1)$. For a

particular n_c, however, the probability to be consistent with the n_{emp} observed coincidences is again given by the hypergeometric distribution (Eq. (5))

$$\mathcal{H}_{T_h-n_c}^{n_1-n_c,\,n_2-n_c}\left(n_{emp}-n_c\right).$$

The probability that a particular n_c underlies the observation therefore is

$$\frac{1}{n_{emp}+1}\mathcal{H}_{T_h-n_c}^{n_1-n_c,\,n_2-n_c}\left(n_{emp}-n_c\right)$$

and the total probability to observe n_{emp} coincidences is

$$\sum_{i=0}^{n_{emp}}\frac{1}{n_{emp}+1}\mathcal{H}_{T_h-i}^{n_1-i,\,n_2-i}\left(n_{emp}-i\right).$$

Consequently, the expected n_c given the observed triplet (n_1, n_2, n_{emp}) is

$$n_c=\sum_{i=0}^{n_{emp}}\frac{i\cdot\frac{1}{n_{emp}+1}\mathcal{H}_{T_h-i}^{n_1-i,\,n_2-i}\left(n_{emp}-i\right)}{\sum_{j=0}^{n_{emp}}\frac{1}{n_{emp}+1}\mathcal{H}_{T_h-j}^{n_1-j,\,n_2-j}\left(n_{emp}-j\right)}$$

which reduces to

$$n_c=\frac{\sum_{i=0}^{n_{emp}}i\cdot\mathcal{H}_{T_h-i}^{n_1-i,\,n_2-i}\left(n_{emp}-i\right)}{\sum_{i=0}^{n_{emp}}\mathcal{H}_{T_h-i}^{n_1-i,\,n_2-i}\left(n_{emp}-i\right)}. \qquad (10)$$

Note that the denominator is not unity as the sum extends over probabilities from different hypergeometric distributions.

2.1 Correspondence of models of assembly activations

The intuitive approximation (Eq. (9)) is compatible with the more detailed model result (Eq. (10)). To see this, we first approximate the individual hypergeometric distributions in Eq. (10) by the one for the average n_c and extend the range of the sums to the maximum value $T_h - n_c$:

$$n_c=\frac{\sum_{i=0}^{T_h-n_c}i\cdot\mathcal{H}_{T_h-n_c}^{n_1-n_c,\,n_2-n_c}\left(n_{emp}-i\right)}{\sum_{i=0}^{T_h-n_c}\mathcal{H}_{T_h-n_c}^{n_1-n_c,\,n_2-n_c}\left(n_{emp}-i\right)}.$$

This reduces the denominator to unity. With the substitution $k = n_{emp} - i$ this reads

$$n_c=\sum_{k=n_{emp}-(T_h-n_c)}^{n_{emp}}\left(n_{emp}-k\right)\cdot\mathcal{H}_{T_h-n_c}^{n_1-n_c,\,n_2-n_c}\left(k\right).$$

The sum only has contributions for $k \geq 0$ and we again extend the range to the maximum value

$$n_c=\sum_{k=0}^{T_h-n_c}\left(n_{emp}-k\right)\cdot\mathcal{H}_{T_h-n_c}^{n_1-n_c,\,n_2-n_c}\left(k\right).$$

This suggests a decomposition of the sum to

$$n_c=n_{emp}\sum_{k=0}^{T_h-n_c}\mathcal{H}_{T_h-n_c}^{n_1-n_c,\,n_2-n_c}\left(k\right)$$

$$-\sum_{k=0}^{T_h-n_c}k\cdot\mathcal{H}_{T_h-n_c}^{n_1-n_c,\,n_2-n_c}\left(k\right)$$

where the first sum is unity and the latter term is the expectation value of k. Hence,

$$n_c=n_{emp}\cdot1-\frac{1}{T_h-n_c}\cdot\left(n_1-n_c\right)\left(n_2-n_c\right)$$

which is Eq. (8) implying the intuitive result (Eq. (9)).

2.2 Accuracy of the count of assembly activations

Before we turn to the experimental data in Section 4 we need to assess the accuracy of our estimator of n_c. To this end we construct a surrogate data set with parameters adapted to the experimental data. The data are organized into 32 blocks containing an increasing number of assembly activations from $n_c = 0$ to 31. A block is composed of $M = 2{,}700$ time windows, consistent with the number of UE windows found in the experimental data. A time window has a duration (Eq. (6)) of $T_h = 5{,}000$ at a resolution of $h = 1$ ms, corresponding to the 100 ms segments of UE analysis covering 50 trials. Each time window contains surrogate spike trains of two neurons with totals of exactly $n_1 = n_2 = 100$ spikes (corresponding to a rate of 20 Hz). In each spike train the $n_i - n_c$ non-assembly spikes are uniformly distributed over the time window.

Figure 3(a) shows the distribution of injected coincidences n_c in the data set organized by the total number of coincidences in each window n_{emp}. Here and in the following only windows containing a significant number of coincidences ($n_{emp} > n_\alpha$) are analyzed (cf. Section 1.1). The estimator (Eq. (9)) of n_c assigns a unique value to each value of n_{emp} because the spike counts are identical for all windows. Therefore at a given n_{emp} the estimate of n_c is identical for each window and thereby identical to the average over all windows. In Fig. 3(b) we observe that the estimate well describes the mean (red curve) of the actual distribution of n_c at a particular n_{emp}. The panel further demonstrates that the approximative (Eq. (9)) and the exact (Eq. (10)) estimator only start to deviate for n_{emp} below n_α. Figure 3(c) uses the same representation as Fig. 3(a) to illustrate the excellent correspondence between the distribution of the actual β and the estimated values. The total distribution of β (Fig. 3(d)) in the model is asymmetrical because of the lack of a typical n_c and the
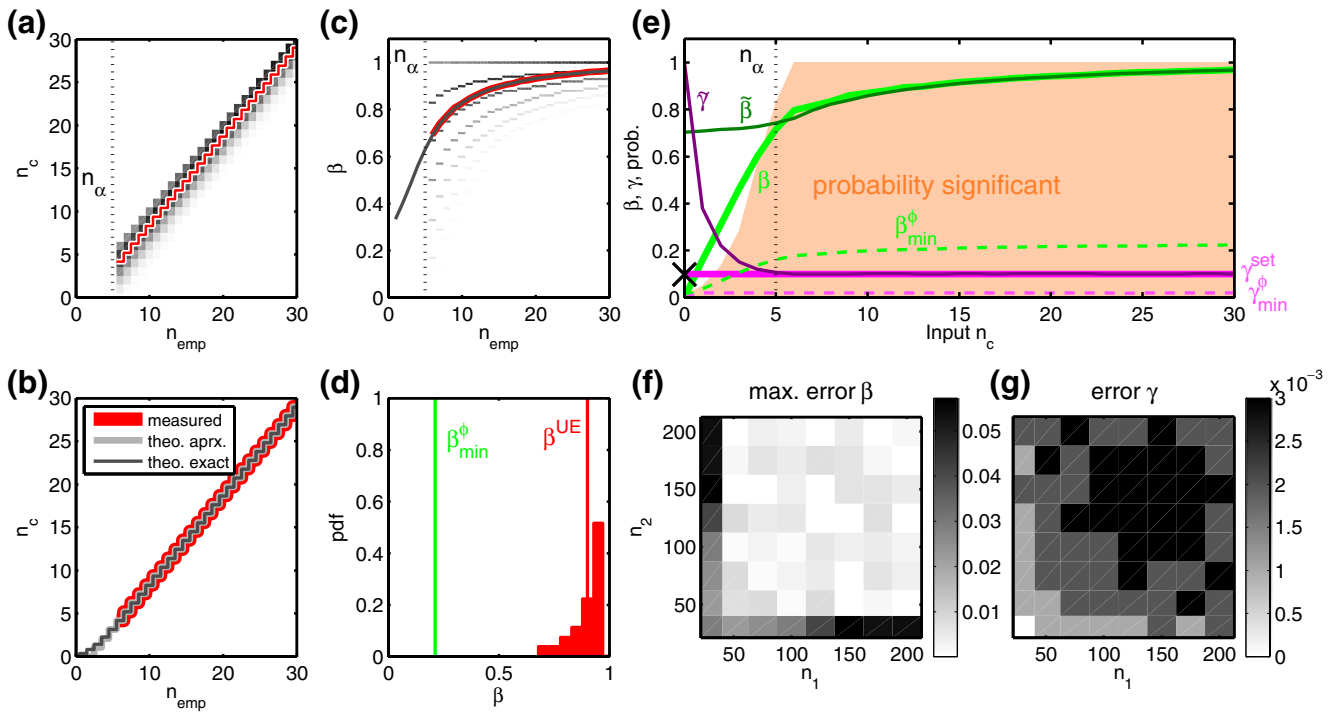
Fig. 3 Consistency and sensitivity of the estimation procedure in surrogate data. (**a**) The actual distribution (*gray shading*) of numbers of injected coincidences n_c and its measured mean (*red curve*) as a function of the empirical number n_{emp} of coincidences per time window. The simulated data is composed of blocks, each involving a different number of injected coincidences $n_c = 0, 1, \ldots 31$. A total of 2,700 windows are simulated per block, each with a total of exactly $n = 100$ spikes. Values extracted for all significant (UE analysis, $\alpha = 0.05$) windows per value of n_c are pooled and reordered according to n_{emp}. The *dotted line* represents the minimum number n_α of coincidences needed for a window to become significant. (**b**) Approximate (*light gray*) and exact (*dark gray*) theoretical prediction of n_c. For comparison, the *red curve* shows the average n_c (same curve as in panel (**a**)). Since the total number of spikes per window is fixed the expected number of coincidences and thus n_α are constant, as well as the estimated n_c for a given n_{emp}. (**c**) Distribution (*gray shading*) and measured mean (*red curve*) of $\beta = n_c n_{emp}^{-1}$ as a function of n_{emp}, depicted in analogy to panel (**a**). β assumes only discrete values. The *dark gray curve* shows the result from the exact theoretical prediction indicated by the *dark gray curve* in panel (**b**). (**d**) Distribution of β (*red bars*) across all observed

values of n_{emp} and the corresponding mean β^{UE} (*red line*). The *green line* shows that the lower bound β_{min}^ϕ calculated from the simulated phase distributions is well below the mean for our particular choice of $p_a(\phi)$. (**e**) Comparison of theoretical (*light green*, $\bar{\beta}$) and estimated (*dark green*, $\tilde{\beta}$) values of β from the UE analysis and corresponding estimates of γ (*magenta*) from the phase distributions as a function of n_c. *Dashed curves* show the minimum β_{min}^ϕ and corresponding γ_{min}^ϕ. The *shaded graph* indicates the probability for a time window created with given n_c to become significant in the UE analysis. The deviation of estimates from their theoretical value for $n_c < n_\alpha$ (*dotted line*) is an expected artifact due to the preselection of significant UE periods. The average β^{UE} from (**d**) yields a good estimate $\gamma = 0.098 \approx \gamma^{set}$ (*black cross on vertical axis*). (**f**) Absolute error in estimating β^{UE} for $n_{emp} = n_\alpha + 1$ (smallest possible n_{emp}) for different choices of the spike count n_1 and n_2 of the two neurons. Counts correspond to firing rates between 5 Hz and 40 Hz. The *shading* visualizes the absolute difference between β^{UE} and the theoretical prediction (*red and gray curves* in panel (**c**), taken at the data point to the *right* of n_α). (**g**) Absolute error $|\gamma^{set} - \gamma|$ in estimating the final result γ for different choices of n_1 and n_2

constant spike count. In conclusion, our measure is a faithful estimator of the average β in the data.

3 Estimating the assembly participation probability in a joint spike-LFP model

In the following section, we extend the simple spike model defined in Section 2 to include a representation of the LFP locking and verify numerically that the

results obtained in the previous section yield a reliable estimate of our model parameters β and γ.

3.1 Combined spike-LFP model

First, we choose a parameter γ^{set} as the assumed probability that any spike in the network is part of an assembly activation, in agreement with our conceptual model. In the following we simulate for each choice of n_c the set of $M = 2,700$ UE time windows as combinations

of an injection process and a background process for two neurons as described in Section 2.2. In addition, CC time windows are generated only by a background process with equal spike counts $n_1 = n_2 = 100$ (see Fig. 2(b) for an illustration). Again, only significant UE windows ($n_{\text{emp}} > n_\alpha$) and non-significant CC windows ($n_{\text{emp}} \leq n_\alpha$) are retained. To model the experimental results, we assign a label 'a' to all spikes that originate from an assembly activation (Fig. 2(b)). By definition of our assembly process, every spike that originates from the injection process receives a label. In addition, a random proportion γ^{set} of spikes from the background process is labeled (in both, UE and CC time windows). In our simulations, the overall probability for a spike to belong to an assembly is set to $\gamma^{\text{set}} = 0.1$. Next, we define two distributions $p_n(\phi)$ and $p_a(\phi)$, where the latter has a larger modulation depth, which describe the locking of non-assembly and assembly spikes, respectively (Fig. 2(a)). Here, $p_n(\phi)$ is modeled as a uniform distribution, whereas for $p_a(\phi)$ the distribution is modeled as a Gaussian as an approximation of a von Mises distribution. The modulation of the Gaussian was chosen to mimic that of the experimentally observed distributions $p_{\text{CC}}(\phi)$, and $p_{\text{UE}}(\phi)$ (compare Figs. 1(b) and 2(a)). Classification of spikes into the groups ISO (taken here as the single spikes in CC windows), CC and UE allows us to calculate the simulated phase distributions $p_{\text{ISO}}(\phi)$, $p_{\text{CC}}(\phi)$, and $p_{\text{UE}}(\phi)$ as mixtures of $p_n(\phi)$ and $p_a(\phi)$.

3.2 Estimating the minimal β from phase distributions

The setup allows us to follow the same analysis steps that we will perform on the experimental data in the following section. From the conceptual model introduced in Section 1.3, which is formally expressed in Eqs. (1–3), we infer the lower bound β_{min}^ϕ of coincidences originating from assemblies during an observed UE period. Substituting the measured population phase distribution of chance coincidences $p_{\text{CC}}(\phi)$ of Eq. (2) into Eq. (3) yields an expression relating the known phase distribution $p_{\text{UE}}(\phi)$ of UE coincidences to the parameter β and the squared phase distribution of assembly spikes $p_a^2(\phi)$:

$$p_{\text{UE}}(\phi) = (1 - \beta) \cdot p_{\text{CC}}(\phi) + \beta \cdot p_a^2(\phi). \quad (11)$$

By systematic variation of the parameter β we thus obtain a corresponding phase distribution $p_a^2(\phi)$ by solving the equation separately for each bin of the respective distributions. However, for small values of β the assembly distribution $p_a^2(\phi)$ must exhibit a strong modulation to compensate for the large difference between $p_{\text{UE}}(\phi)$ and $(1-\beta)p_{\text{CC}}(\phi)$. In the extreme case,

it can become necessary for a bin of $p_a^2(\phi)$ to contain negative values, ruling out that particular choice of the parameter β. From this consideration, we define β_{min}^ϕ as the lowest value of β that leads to a phase distribution $p_a^2(\phi)$ with non-negative entries. This lower bound depends on the choice of the assembly phase distribution $p_a(\phi)$ we initially introduced into our model: the smaller the difference in locking between assembly spikes $p_a(\phi)$ and non-assembly spikes $p_n(\phi)$, the lower are the values that β_{min}^ϕ may assume. In Fig. 3(d) we show that in our spike-LFP model we obtain a value for β_{min}^ϕ well below the mean β^{UE} extracted from the spike analysis (Section 2). We note that when generating the model data with a more strongly modulated Gaussian to model $p_a(\phi)$, the minimum β_{min}^ϕ approaches β^{UE} from below (not shown).

3.3 Estimation of γ from phase distributions

We are now prepared to extract the parameter γ from the simulated phase distributions using the estimate of β^{UE} derived in Section 2. First, using $\beta = \beta^{\text{UE}}$ we extract $p_a^2(\phi)$ from Eq. (3) by solving the equation separately for each bin of $p_a^2(\phi)$. Thus, $p_a(\phi)$ is determined independently of the distribution $p_n(\phi)$ describing the general entrainment of spikes to the LFP. Taking the square root $p_a(\phi)$, we renormalize the distribution to unit area, and insert it in Eq.(1). By variation of the parameter γ, we find the value $\gamma^{\phi,\text{UE}}$ that minimizes the sum of the absolute bin-by-bin differences between the distribution $p_{\text{ISO}}(\phi)$ and the right side $(1 - \gamma) \cdot p_n(\phi) + \gamma \cdot p_a(\phi)$ of Eq. (1).

Using this method, we first analyze separately each of the 32 blocks of data with a fixed number of injected coincidences n_c. Let $\tilde{\beta}$ and $\tilde{\gamma}$ be the estimated values of β^{UE} and $\gamma^{\phi,\text{UE}}$ within a single block. In Fig. 3(e) $\tilde{\beta}$ and $\tilde{\gamma}$ are compared to their theoretical means $\bar{\beta}$ (calculated as the mean of n_c/n_{emp} over all significant time windows from one set of simulations with fixed n_c) and γ^{set}. The estimates are in good agreement for datasets where n_c was set above the significance threshold n_α. Below this threshold, the number of injected coincidences, and hence $\tilde{\beta}$ and $\tilde{\gamma}$, is overestimated. Clearly, in this regime contributing UE periods become significant only due to an unusually high number of coincidences in the background process. Therefore, the mean approximation (Eq. (9)) does no longer hold. In principle it is possible to correct for the bias due to the selection of significant periods, although it is not possible to arrive at an expression in closed form for n_c. Nevertheless, in fact only a small fraction of windows that were created with low values of n_c actually become significant due to such exceptionally high coincidences counts from

the background (the shaded graph in Fig. 3(e) shows the probability for a window to become significant for given n_c). Therefore, in practice, where the true n_c is unknown, only few windows enter the analysis when n_c is low. For this reason, the bias does not affect the estimate when re-sorting the data according to the observed n_{emp} (Fig. 3(a)–(c)), and the original value of γ^{set} is well estimated as $\gamma^{\phi,UE} = 0.098$ from the complete dataset spanning all n_c (black cross in Fig. 3(e)). Consistently, the lower bounds β^{ϕ}_{min} and γ^{ϕ}_{min} are a lower bound on the respective estimates $\tilde{\beta}$ and $\tilde{\gamma}$ for all values of n_c.

To quantify whether our procedure works for other values for the total spike counts n_1 and n_2 we introduce two error measures. First, we note that our prediction of β will be least accurate when the coincidence count n_{emp} is low. Thus, we numerically quantify the absolute error of β for the smallest value of n_{emp} to become significant (Fig. 3(f)). Moreover, we quantify the absolute error of our final estimate γ obtained from the complete spike-LFP model (Fig. 3(g)). Both measures indicate that the parameters are reliably extracted for a wide range of activity levels in both neurons, even in the case of large rate differences between the two. In summary, these calibrations using a simple combined spike-LFP model demonstrate that our method and the approximation (Eq. (9)) are well suited to estimate the parameter γ from the phase distributions in a dataset with a realistic choice of parameters.

4 Analysis of experimental data

In this section we estimate the percentage of coincident spike events reflecting assembly activity (model parameter β) and the percentage of spikes that are part of an assembly activation (model parameter γ) from neuronal data of primary motor cortex of monkeys.

4.1 Experimental procedures

Two rhesus monkeys were trained to perform arm movements in two different tasks involving an instructed delay. One of these tasks requires a self-initiated movement after one of two preset delays, while the other requires the discrimination between two such delays. In this study we exclusively analyze the delay activity during the preparatory period for the upcoming arm movement. Only correct trials were considered, in which the monkey responded within a predefined time window and in which movements were performed in the required movement direction. LFPs and spikes were recorded simultaneously in primary

motor cortex using a multielectrode device of 2–4 electrodes. Spikes of single neurons were detected by an online sorting algorithm. The inter-electrode distance was on the order of 400 μm. LFPs were sampled at a resolution of 250–500 Hz and hardware filtered (band pass, 1–100 Hz). In total, we analyzed 53 recording sessions, which yielded 143 single neurons or 570 combinations of neurons and behavioral conditions. Each of these eight possible behavioral conditions is a combination of the task during which the neuron was recorded, the length of the delay (short or long), and the direction of the arm movement (left or right). This selection included only those neurons which exhibited an average firing rate of 5 Hz or more and exhibited at least 25 spikes in total. On average 33±11 (mean±standard deviation) trials were recorded per experimental condition. In analyses that combine spikes and LFP, each neuron enters only once, and we never combined LFP and spikes that were recorded on the same electrode to exclude the possibility of spike artifacts in the signal. We confirmed that simultaneously recorded LFPs are highly synchronous in the frequency regimes of interest. For experimental details see Roux et al. (2006).

The Unitary Events analysis was applied to all simultaneously recorded pairs of neurons recorded on different electrodes thus totaling 123 pairs of neurons. Data were analyzed separately for the possible behavioral conditions and each experimental session. Defining a neuron by the combination of its identity and the behavioral condition (see above) during which it was recorded, data from the same neuron may enter a population average up to eight times. Based on the results of the UE analyses each individual spike was marked as either ISO, CC or UE (see Section 1.1 and Denker et al. (2010) for details). We observed a dominant component of the β-band (in both monkeys around 17 Hz) in the LFP during the preparatory period (see, e.g., Murthy and Fetz 1996a) and therefore filtered the LFP accordingly before applying the Hilbert transform to extract its instantaneous phase and amplitude (i.e. the envelope). LFPs of both monkeys are filtered with a zero-phase 10–22 Hz band pass filter (Butterworth, 8-pole). Finally, we derived the phase histogram pooled over the whole population for each class of spikes (see Fig. 1(b)). The histogram was pooled after averaging, such that each spike enters the distribution unweighted (as is required for our analysis).

4.2 Estimating the minimal β from phase distributions

As described in Section 3.2, we derive the lower bound of the percentage of coincidences originating from as-

sembly activity $\beta_{\min}^{\phi}$. Again, we make use of the Eq. (11) that relates the phase histogram of the UE to the model parameter β. This equation contains two components that we extract from the data, i.e. the phase distribution of the UE spikes $p_{\text{UE}}(\phi)$ and the phase distribution of the CC spikes $p_{\text{CC}}(\phi)$. We vary β in the interval $[0, 1]$ and retrieve the corresponding $p_a^2(\phi)$. Four examples are shown in Fig. 4 for different values of β. The first two examples demonstrate how choosing a small value of β may lead to distributions $p_a^2(\phi)$ that have negative
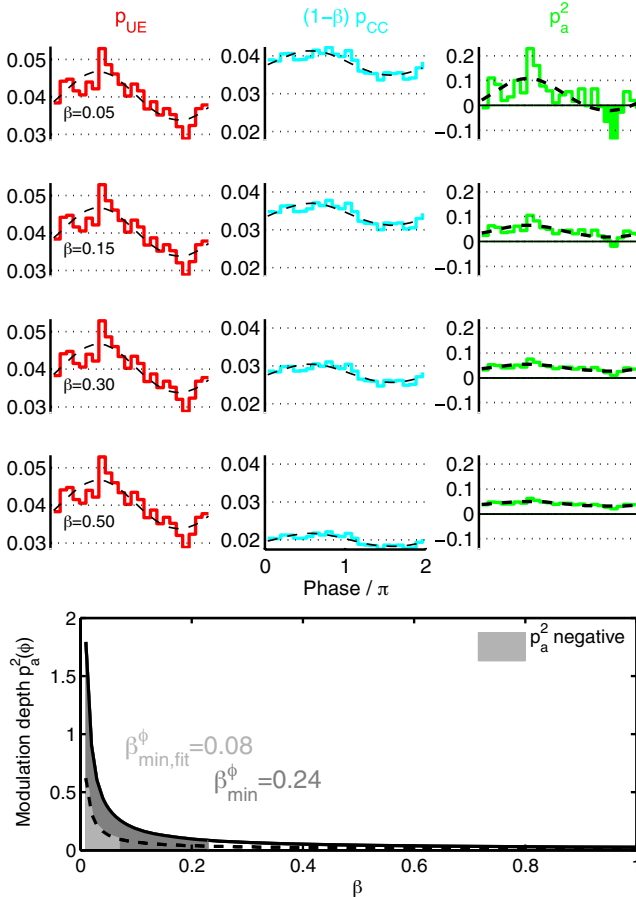
Fig. 4 Determination of $\beta_{\min}^{\phi}$ from experimental data. The *four upper rows* show the measured distributions $p_{\text{UE}}(\phi)$ (*left, red*), $(1 - \beta) p_{\text{CC}}(\phi)$ (*middle, cyan*) and the resulting squared phase distribution of assembly spikes $p_a^2(\phi)$ (*right, green*) entering the third equation in Fig. 2(a), i.e. Eq. (3) in the text, for four different choices of β. The *solid line in the bottom graph* shows the relationship between the choice of β and the resulting modulation depth of $p_a^2(\phi)$ (difference between maximum and minimum). The *dark gray area* indicates invalid choices of β where the corresponding distribution $p_a^2(\phi)$ has negative values (compare *green filled areas* of $p_a^2(\phi)$ in the *top two rows*). Hence, the minimum value for β is determined as $\beta_{\min}^{\phi} = 0.24$. The *dashed lines in the top graphs* show the results for an analogous analysis using fitted von Mises distributions for $p_{\text{CC}}(\phi)$ and $p_{\text{UE}}(\phi)$. The minimum value for β is determined as $\beta_{\min,\text{fit}}^{\phi} = 0.08$ (see *light shaded gray in bottom graph*)

entries and indicate a non-valid solution. From the systematic variation of β we derive the value $\beta_{\min}^{\phi} = 0.24$ that just leads to non-negative entries in all bins of $p_a^2(\phi)$ (see Fig. 4, bottom). This result indicates that at least about a fourth of the coincidences during a UE period are reflections of the observed assembly. In addition, we performed an analogous analysis using von Mises distributions fitted to the experimental distributions $p_{\text{CC}}(\phi)$ and $p_{\text{UE}}(\phi)$ (dashed lines in Fig. 4). This approach is very conservative because the fits neglect the prominent systematic peaks present in the original distributions. In this analysis, the minimum value for β is determined as $\beta_{\min,\text{fit}}^{\phi} = 0.08$.

4.3 Estimate of β^{UE} from coincidence counts

As introduced in Section 2 the parameter β can be estimated by the comparison of the expected number of coincidences n_{exp} and the number of empirical coincidences n_{emp}. Figure 5(a) shows n_{emp} as a function of its respective n_{exp} value for all analysis windows that were detected to contain UE for all pairs and sessions. Due to the selection of windows that contain UE, the empirical coincidence counts are larger than the minimal number of coincidences n_{α} required to be significant given the significance level α. n_{α} does not have a constant difference to n_{exp} but increases non-linearly as a function of n_{exp} since the Poisson distribution used for the evaluation of the significance becomes broader for larger expected number of coincidences (see for details in Grün et al. 2002a).

Based on Eq. (9) we estimate the number of coincidences resulting from the active assembly n_c for each n_{exp}, n_{emp} combination. We use a version of the UE analysis that also allows to detect temporally imprecise coincidences by employing the multiple-shift method. This method detects coincidences of systematically shifted spike trains up to a predefined shift, and sums the coincidence counts from all shifts. For the UE evaluation, the expected number of coincidences then has to be adjusted by a factor of $2s + 1$ with s being the maximal shift (cf. Section 1.1; for details see Grün et al. 1999). Thus to estimate the number of coincidences n_c that generically result from active assemblies, Eq. (9) was adjusted correspondingly (Section 2). Figure 5(b) shows n_c versus its corresponding n_{emp} and we find an average number of assembly coincidences of $n_c^{\text{UE}} = 6.70$.

Next, we can derive for each pair of (n_c, n_{emp}) an estimate of the percentage of coincidences reflecting assembly activity β (Fig. 5(c)). We find these values peaked around a mean of $\beta^{\text{UE}} = 0.51$ with a standard
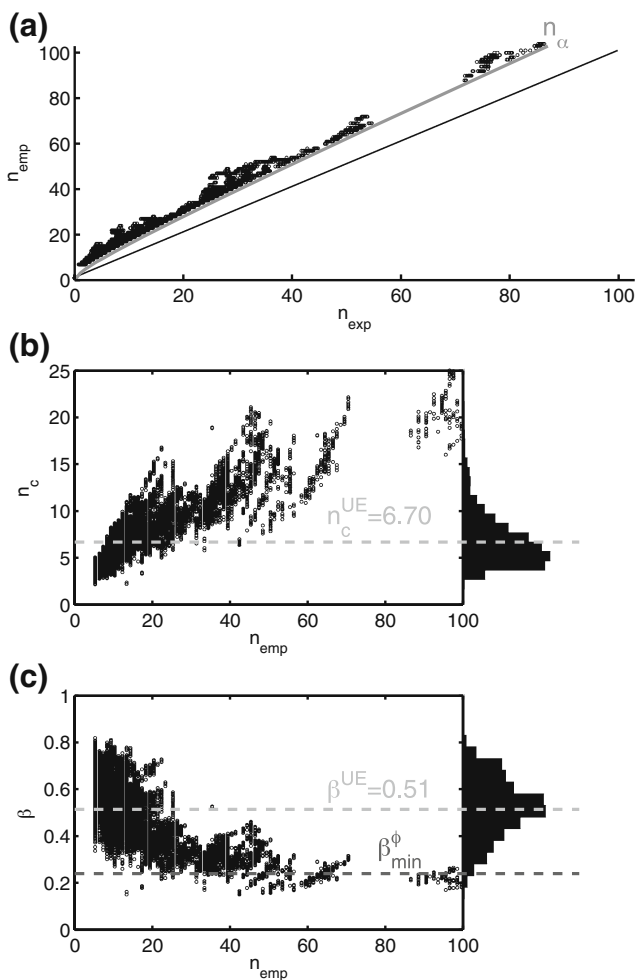
Fig. 5 Distribution of the estimated β and the resulting value of β^{UE} obtained from the data. (**a**) Scatter plot of the expected number of coincidences n_{exp} and the empirical number of coincidences n_{emp}. Here and in the following panels, each *dot* represents the values of one significant 100 ms time window (UE analysis, $\alpha = 0.05$), where each neuron exhibits a minimum average rate of 5 Hz. The *gray curve* indicates the critical number n_α of coincidences to reach significance at a given n_{exp}. The *black line* shows the diagonal. Due to the preselection of datasets with a minimum rate, no data points for small n_{exp} are observed. (**b, c**) Estimated values of n_c (**b**) and β (**c**) as a function of the empirical coincidence count n_{emp}. The total distributions of n_c and β are shown on the *right* (*black*). The averages $n_c^{\text{UE}} = 6.70$ and $\beta^{\text{UE}} = 0.51$ are indicated by the *light gray lines*. The *dark gray line* in (**c**) shows the lower bound β_{min}^ϕ for comparison

only have predictive power on the population mean β^{UE}, but not the estimate β of single windows (compare Fig. 3(d)).

4.4 Estimation of γ from phase distributions

Finally, we estimate the fraction γ of spikes in the network that are part of an assembly activation (compare Section 3.3). Again, by systematic variation of the parameter γ in Eq. (1) we find the best fit between the measured $p_{\text{ISO}}(\phi)$ and the right side of the equation (see Fig. 6). Here, the assembly distribution $p_a(\phi)$ is known from Fig. 5 by setting $\beta = \beta^{\text{UE}}$, while
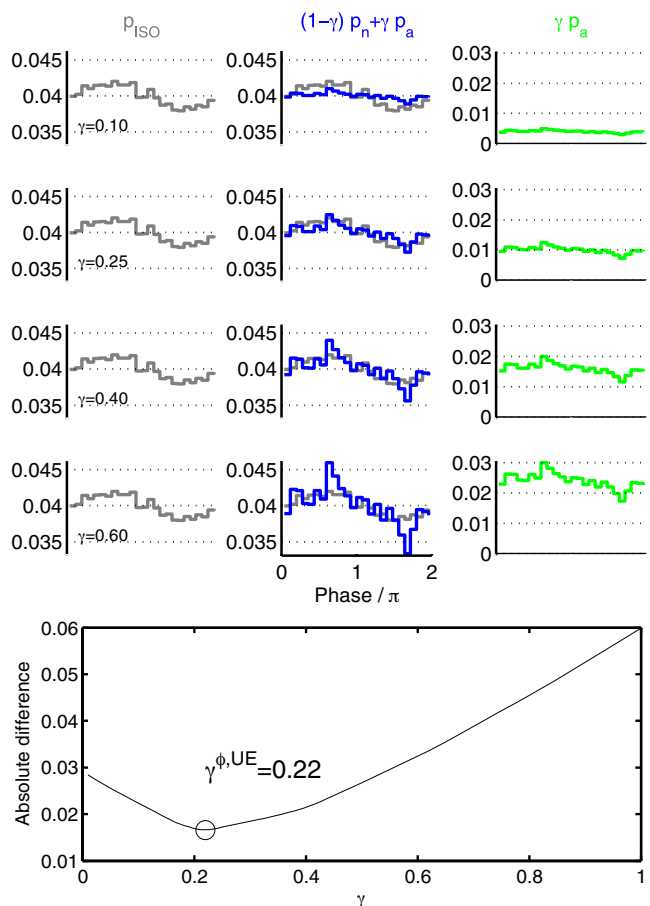
Fig. 6 Calculation of the average assembly participation γ from experimental data. The *top four rows* show the different distributions entering the first equation in Fig. 2(a), i.e. Eq. (1) in the text, for four different choices of γ. *Left*: measured phase distribution $p_{\text{ISO}}(\phi)$ (*gray*) of isolated spikes, repeated in each of the *four rows*. *Middle*: calculated distribution $(1 - \gamma) p_n(\phi) + \gamma p_a(\phi)$ (*blue*) compared to the measured distribution $p_{\text{ISO}}(\phi)$ (*gray*). *Right*: Scaled assembly distribution $\gamma p_a(\phi)$ (*green*). The *bottom graph* shows the absolute sum of the bin-by-bin differences between $p_{\text{ISO}}(\phi)$ and the calculated distribution as γ is varied (compare distributions in the *middle column*). Both distributions match best for $\gamma^{\phi,\text{UE}} = 0.22$

deviation of 0.123. The obtained distribution of β is in good agreement with the minimum β_{min}^ϕ obtained in Section 4.2, as β_{min}^ϕ lies well below the mean β^{UE} of the distribution of β as expected for an estimate of a lower bound. Although β_{min}^ϕ in fact seems to be a bound for the complete distribution of β, this must not necessarily be the case: As a measure derived from the population estimate of phase distributions, it can

the phase distribution of non-assembly spikes $p_n(\phi)$ is taken as the uniform distribution (see Section 5). Taking all analysis steps together, the best fit is derived for $\gamma^{\phi,\text{UE}} = 0.22$ indicating that on average 22% of the spikes of any neuron in the network originate from assembly activity (Fig. 6, bottom). To corroborate this result, we obtain a similar value of $\gamma^{\phi,\text{UE}} = 0.26$ using the same procedure described above to solve (Eq. (2)).

5 Discussion

Despite the complex mechanisms that contribute to the formation of the local field potential, it is well established that a primary contribution to the oscillatory LFP dynamics arises from the superposition of synchronized, slow transmembrane currents of cells close to the recording site (Mitzdorf 1985; Logothetis and Wandell 2004). Nevertheless, how rhythmicity in the LFP is linked to synchrony on the spiking level has remained an open question (Poulet and Petersen 2008; Tetzlaff et al. 2008) due to the unspecificity of the LFP signal and lack of a clear global oscillatory spiking activity. Several authors have interpreted the LFP as reflections of the specific synchronous synaptic activity responsible for the coactivation of neurons in the context of cell assemblies (Eckhorn et al. 1988; Murthy and Fetz 1996b) which, in the simplest case, time their activations to the network rhythm revealed by the LFP (Singer 1999). Our recent experimental findings (Denker et al. 2010) demonstrate that indeed only assembly activity, which is identified as transient periods of significant excess spike synchrony between two neurons (UE analysis, Grün et al. 2002a, b), shows an exceptional phase relationship to LFP oscillation that exceeds expectation (Fig. 1(b)), thus confirming the hypothesis.

This link provides a handle on characterizing the spike synchronization dynamics in the context of the network oscillations. Here, we introduced a simple model that captures the main experimental findings on the spike-LFP relationship in such a way that it includes a parameter γ that measures the overall participation probability of individual spikes in an assembly—independent of whether it is observed or not. However, the estimation of this network parameter from the data requires knowledge of a second model parameter β, the expected relative number of (excess) coincidences stemming from an active assembly (as opposed to chance coincidences) during UE periods. Estimating this number from the spike data alone and integrating it into the conceptual model yields an average participation of single spikes to assembly activity of $\gamma = 22\%$, a parameter that describes precise spike synchronization on the network level.

Our model assumes that the observed excess synchrony is the result of the specific activation of the observed neurons. An alternate hypothesis states that spike synchronization is solely caused by fluctuations of the spiking likelihoods due to the entrainment of neurons to a common LFP oscillation. In this case, no UEs would be observed due to the lack of a spike coincidence count exceeding the expectation based on the rate time course. Only if the analysis would unsuccessfully correct for non-stationarities of the firing rates false-positive detections of UE periods emerged. Consequently, these false-positives would trivially correlate with the LFP. To exclude this possibility, we repeated the analysis presented in this study by replacing the parametric distribution of coincidences used for evaluating the significance of observed coincidence counts by numerically derived distributions based on surrogate data. The employed surrogate method (spike train dithering, see Grün 2009) considers further statistical features of the experimental spike trains (in particular non-stationarity of rates on a short time scale and the inter-spike interval distributions) while destroying precise spike coincidences. Simulations showed that this method is more conservative (Louis et al. 2010) than the analysis used here, but despite the decreased sensitivity confirms the differences in the phase distributions for ISO, CC, and UE. These findings are the essential experimental foundation of our study.

The combination of measurements of synchrony on the local and mesoscopic scales enables us to access parameters of the network dynamics that remain hidden on the two individual levels of observation. The nature of the estimation process via the population phase distributions obtained from the LFP-spike locking requires a vast pooling of the experimental data in different sessions and two different monkeys in order to obtain a large sample of neurons as an appropriate representation of the network activity in motor cortex. The resulting value for γ must therefore be seen as a coarse estimate of the degree of assembly activity. Despite the finding that the parameter β shows a rather narrow distribution (Fig. 5(c)), the extent to which an individual neuron contributes to the assembly dynamics will fluctuate around the network mean γ. By quantifying the quantiles for 2 standard deviations of the experimental distribution of β we find corresponding γ values between $\gamma = 0.14$ and $\gamma = 0.39$ as an estimate of the range of γ across neurons.

In understanding the underlying dynamical structure, more interesting than the precise value of γ itself are two simple observations: first, $\gamma < 1$. Therefore,

not all spikes are part of an assembly activation, and some spikes must be attributed to a complementary mechanism. Intuitively, this is clear from the observation that in individual UE periods, it is always the falling phase of the LFP oscillation where increased locking is observed. Second, $\gamma > 0$, specifically about one fourth of all spikes must originate from assembly activity (with a non-zero lower bound γ_{min}^{ϕ}). Thus, under the assumptions of our model, the activation of assemblies is an ubiquitous phenomenon in the network, providing compelling evidence for the presence of an assembly coding scheme in addition to the correlation of synchrony with behavior (Kilavik et al. 2009). Indeed, we typically observe a fraction of about 26% of neurons that show UEs during a given task. Therefore, combined with the large value of γ this suggests that typically any neuron is part of one or several assemblies. Nevertheless, information about the existence of higher-order synchrony between neurons is required for a further characterization of the assembly, such as an estimate of the number of participating neurons. Moreover, such an analysis could better disambiguate whether spikes that do not originate from assembly activation show an intrinsic degree of phase locking that deviates from our model assumption of a uniform $p_n(\phi)$.

A crucial part of our analysis is the estimation of the average relative amount of excess synchrony β during UE periods directly from the synchrony analysis. By design, all currently available methods to detect the presence of an active assembly activation, such as the UE analysis, rely on a significance test (Grün 2009). Therefore, we must assume that the amount of excess synchrony β is influenced by the explicit choice of the significance level α used for the significance test: For a very restrictive significance criterion, only UE periods with very high excess synchrony n_c will be detected, resulting in higher values of β. Nevertheless, by choosing the same α level for the phase locking analysis, we will in turn also obtain a correspondingly more modulated distribution for $p_{UE}(\phi)$, reflecting that it contains a higher proportion of assembly spikes. Therefore, β yields a consistent phase distribution $p_a(\phi)$ of the assembly spikes independent of α.

The approximate formula (Eq. (9)) does not make any assumptions on preselecting significant periods of synchrony in the first place. However, we apply this estimate specifically in periods that display a significant surplus of synchrony, i.e. UE periods. If the number of coincidences n_c coming from the injection process is small, significance is only reached with an unexceptionally high amount of coincidences originating from the background. Therefore, in our stochastic spike-LFP model, in those very few trials that are significant for a given small n_c, the average background rate will be underestimated and the average injection rate overestimated (Fig. 3(c)). However, restructuring the model data to the realistic case where only n_{emp} is measured (pooling across the data sets with different n_c) automatically incorporates the low frequency of significant windows with small n_c. Therefore the overall estimate of β is not significantly affected.

Our spike-LFP model (Section 3) is created in the spirit of providing a simple abstraction of the experimental findings in order to test the analysis under controlled conditions (equal counts per spike train with fixed injections). Due to this intentional lack of experimental detail in the model, the distribution of β that is obtained in the spike-LFP model naturally deviates from the one found in the data (compare Figs. 3(d) and 5). The model assumes an equal probability for a large range of n_c (number of injected coincidences). In real data, n_c likely follows a much more narrow distribution that does not exploit this range, resulting in lower values of β (compare Fig. 3(e)). Moreover, we assumed fixed counts n_1, n_2 for all neurons in the model. In reality, rates vary considerably across neurons. Especially for higher rates, where a higher proportion of coincidences can be assumed to originate from the background, we would expect a tendency towards lower β values.

In generating the spike-LFP model in Section 3 we did not explicitly model an LFP oscillation to place spikes at specific points of the field potential. In extreme situations this might be an oversimplification, where the constraints placed on the spikes due to the LFP locking influence the probability to detect coincidences. In our model, however, non-assembly spikes are associated with a uniform phase distribution, such that $p_n(\phi)$ does not impose any constraints on the spiking probability in time. The Poissonian spike interval statistics of non-assembly spikes thus remain unaffected. In contrast, the assembly spikes from the background process must in principle be adjusted with respect to a hypothetical LFP so that their phase distribution matches the assumed the distribution $p_a(\phi)$ of assembly spikes. However, as $p_n(\phi)$ is uniform, doing so will only influence the probability of finding a coincidence between two spikes that are both assembly spikes, i.e. that participate by chance in two different assemblies at the same time. Nevertheless, the probability $(\gamma^{set})^2 = 0.01$ of this to happen is negligibly small. Finally, assembly spikes from the injection process are synchronous by definition, and therefore remain unaffected by the choice of $p_a(\phi)$. Due to their small number they can always be freely placed in the time

window T in accordance with $p_a(\phi)$. Taken together, the simple model introduced in Fig. 2 to represent the experimental findings in the context of the spike model includes sufficient detail without the need to explicitly model the actual positions of individual spikes with respect to an artificial LFP.

The analysis we provide in this manuscript focuses on estimating a global parameter that characterizes the network dynamics. Given the overall consistency of our conceptual assembly model, the results suggest that oscillations and assemblies are tightly coupled. However, one may still speculate whether LFPs are in fact the direct cause of the assembly process, or whether they act as a supporting mechanism that coordinates synchronized firing. Whichever the initial cause of the LFP oscillation, a promising next step is to return to the level of the single neuron and integrate our knowledge on the dependence of spike synchrony on the LFP in order to discriminate which of the spikes are likely candidates to originate from the assembly process, and which are not. Two pieces of information aid us in this task: first, the knowledge of the observed UE phase distributions in relation to the ISO and CC distributions determines the likelihood of a single spike to be part of an assembly activation. Second, the parameter γ is informative of the relative number of spikes to assign to the assembly dynamics. Indeed, a third parameter not discussed in this manuscript is the dependence of spike synchrony on the LFP amplitude (Denker et al. 2010). Evaluating these three factors on a spike-by-spike basis, there is reason to believe that analyzing the LFP in relation to the spiking activity may well provide an independent indicator to identify the probabilities with which recorded neurons become transiently synchronized as part of an assembly activation. Moreover, the pulsed activation of assemblies triggered on the LFP oscillations cycle suggested by the experimental data stimulates the idea that maybe even assemblies, where only one neuron is recorded, may be inferred from the phase locking statistics in a probabilistic manner. Thus, incorporating a mesoscopic signals may serve to overcome problems related to the undersampling in multiple single-neuron recordings. The success of the method presented here not only corroborates the assembly hypothesis of neuronal processing, but offers a promising vista on reinterpreting the dynamical implications of observed LFP signals.

List of symbols

T	Duration of each simulated time window
h	Temporal resolution of spike trains
$p_n(\phi)$	Phase distribution of spikes not participating in an assembly
$p_a(\phi)$	Phase distribution of spikes participating in an assembly
$p_{ISO}(\phi)$	Measured phase distribution of ISO spikes
$p_{CC}(\phi)$	Measured phase distribution of CCs
$p_{UE}(\phi)$	Measured phase distribution of UEs
β	Fraction of coincidences part of the observed assembly during a UE period
γ	Fraction of spikes in the network activated in assemblies
β_{min}^{ϕ}	Minimal value of β extracted from the phase distributions alone
γ_{min}^{ϕ}	Value of γ that corresponds to the lower bound β_{min}^{ϕ}
β^{UE}	Average value of β obtained from the UE analysis
$\gamma^{\phi,UE}$	Final estimate of γ obtained from phase distributions using β^{UE}
α	Significance level of the Unitary Event analysis
n_α	Number of coincidences required to reach significance at the α-level
n_i	Spike count of neuron i in a given window across trials
n_{emp}	Empirical number of coincidences in a given time window
n_{exp}	Expected number of coincidences in a given time window based on the firing rates
n_c	Number of spikes part of the observed assembly (i.e., number of injected coincidences in the stochastic spike-LFP model)
M	Number of windows simulated in the spike-LFP model per choice of n_c
γ^{set}	Predefined value of γ in the spike-LFP model

References

Denker, M., Roux, S., Lindén, H., Diesmann, M., Riehle, A., & Grün, S. (2010). The local field potential reflects surplus spike synchrony. arXiv:1005.0361v1 [q-bio.NC].

Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M., et al. (1988). Coherent oscillations: a mechanism of feature linking in the visual cortex? Multiple electrode and correlation analyses in the cat. *Biological Cybernetics, 60*(2), 121–130.

Gerstein, G. L., Bedenbaugh, P., & Aertsen, M. H. (1989). Neuronal assemblies. *IEEE Transactions on Biomedical Engineering, 36*(1), 4–14.

Grün, S. (2009). Data-driven significance estimation for precise spike correlation. *Journal of Neurophysiology, 101*(3), 1126–1140.

Grün, S., Diesmann, M., & Aertsen, A. (2002a). Unitary events in multiple single-neuron spiking activity: I. Detection and significance. *Neural Computation, 14*(1), 43–80.

Grün, S., Diesmann, M., & Aertsen, A. (2002b). Unitary events in multiple single-neuron spiking activity: II. Nonstationary data. *Neural Computation, 14*(1), 81–119.

Grün, S., Diesmann, M., & Aertsen, A. (2010). Analysis of parallel spike trains, chap. unitary events. *Springer Series in Computational Neuroscience*. Springer. ISBN: 978-1-4419-5674-3.

Grün, S., Diesmann, M., Grammont, F., Riehle, A., & Aertsen, A. (1999). Detecting unitary events without discretization of time. *Journal of Neuroscience Methods, 94*(1), 67–79.

Grün, S., Riehle, A., & Diesmann, M. (2003). Effect of cross-trial nonstationarity on joint-spike events. *Biological Cybernetics, 88*(5), 335–351.

Hebb, D. (1949). *The Organization of Behavior: A Neuropsychological Theory*. Wiley, New York.

Katzner, S., Nauhaus, I., Benucci, A., Bonin, V., Ringach, D. L., & Carandini, M. (2009). Local origin of field potentials in visual cortex. *Neuron, 61*(1), 35–41.

Kilavik, B. E., Roux, S., Ponce-Alvarez, A., Confais, J., Grün, S., & Riehle, A. (2009). Long-term modifications in motor cortical dynamics induced by intensive practice. *Journal of Neuroscience, 29*(40), 12653–12663.

Logothetis, N. K., & Wandell, B. A. (2004). Interpreting the bold signal. *Annual Review of Physiology, 66*, 735–769.

Louis, S., Borgelt, C., & Grün, S. (2010). Analysis of parallel spike trains, chap. generation and selection of surrogate methods for correlation analysis. *Springer Series in Computational Neuroscience*. Springer. ISBN: 978-1-4419-5674-3.

Mitzdorf, U. (1985). Current source-density method and application in cat cerebral cortex: Investigation of evoked potentials and EEG phenomena. *Physiological Reviews, 65*(1), 37–100.

Murthy, V. N., & Fetz, E. E. (1996a). Oscillatory activity in sensorimotor cortex of awake monkeys: Synchronization of local field potentials and relation to behavior. *Journal of Neurophysiology, 76*(6), 3949–3967.

Murthy, V. N., & Fetz, E. E. (1996b). Synchronization of neurons during local field potential oscillations in sensorimotor cortex of awake monkeys. *Journal of Neurophysiology, 76*(6), 3968–3982.

Okun, M., Naim, A., & Lampl, I. (2010). The subthreshold relation between cortical local field potential and neuronal firing unveiled by intracellular recordings in awake rats. *Journal of Neuroscience, 30*(12), 4440–4448.

Poulet, J. F. A., & Petersen, C. C. H. (2008). Internal brain state regulates membrane potential synchrony in barrel cortex of behaving mice. *Nature, 454*(7206), 881–885.

Riehle, A., Grün, S., Diesmann, M., & Aertsen, A. (1997). Spike synchronization and rate modulation differentially involved in motor cortical function. *Science, 278*(5345), 1950–1953.

Roux, S., Mackay, W. A., & Riehle, A. (2006). The premovement component of motor cortical local field potentials reflects the level of expectancy. *Behavioural Brain Research, 169*(2), 335–351.

Singer, W. (1999). Neuronal synchrony: A versatile code for the definition of relations? *Neuron, 24*(1), 49–65.

Tetzlaff, T., Rotter, S., Stark, E., Abeles, M., Aertsen, A., & Diesmann, M. (2008). Dependence of neuronal correlations on filter characteristics and marginal spike train statistics. *Neural Computation, 20*(9), 2133–2184.